THE CAMBRIDGE HANDBOOK OF INTELLIGENCE

Edited by Robert J. Sternberg

SECOND EDITION

The Cambridge Handbook of Intelligence

Second Edition

Edited by
Robert J. Sternberg
Cornell University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108485104
DOI: 10.1017/9781108770422

First published 2011
Second edition 2020

Printed in the United Kingdom by TJ International Ltd, Padstow Cornwall

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Sternberg, Robert J. (Robert Jeffrey), 1949– editor.
Title: The Cambridge handbook of intelligence / edited by Robert J. Sternberg, Cornell University.
Description: Second edition. | Cambridge, United Kingdom ; New York, NY : Cambridge University Press, 2020. | Includes index.
Identifiers: LCCN 2019019464| ISBN 9781108485104 (2 volume set : hardback) | ISBN 9781108488433 (Volume II : hardback) | ISBN 9781108488419 (Volume I : hardback)
Subjects: LCSH: Intellect. | Human information processing.
Classification: LCC BF431 .C26837 2020 | DDC 153.9–dc23
LC record available at https://lccn.loc.gov/2019019464

ISBN 978-1-108-48510-4 Set of 2 Hardback Volumes
ISBN 978-1-108-48841-9 Volume I Hardback
ISBN 978-1-108-48843-3 Volume II Hardback
ISBN 978-1-108-71919-3 Set of 2 Paperback Volumes
ISBN 978-1-108-72632-0 Volume I Paperback
ISBN 978-1-108-72634-4 Volume II Paperback

The Cambridge Handbook of Intelligence

Written by the foremost experts in human intelligence, this *Handbook* includes not only traditional topics, such as the nature, measurement, and development of intelligence, but also contemporary research into intelligence and video games, collective intelligence, emotional intelligence, and leadership intelligence. In an area of study that has been fraught with ideological differences, it provides scientifically balanced and objective chapters covering a wide range of topics. It does not shy away from material that historically has been emotionally charged and sometimes covered in biased ways, such as intellectual disability, race and intelligence, culture and intelligence, and intelligence testing. The overview provided by this two-volume set leaves virtually no area of intelligence research uncovered, making it an ideal resource for undergraduates, graduate students, and professionals looking for a refresher or a summary of the latest developments in the field.

ROBERT J. STERNBERG is Professor of Human Development at Cornell University, USA, and Honorary Professor of Psychology at Heidelberg University, Germany. He has won the Grawemeyer Award in Psychology and awards from the Association for Psychological Science (APS), including the William James Award and the James McKeen Cattell Award.

This volume is dedicated to the late Nicholas Mackintosh, formerly professor at the University of Cambridge, whose seminal contributions to the field of intelligence have helped enormously to make the field what it is today.

Contents

Figures

Tables

Contributors

PHILLIP L. ACKERMAN, Georgia Tech

ERIK M. ALTMANN, Michigan State University

SOON ANG, Nanyang Technical University, Singapore

SUSAN M. BARNETT, Northwestern State University

LUCY A. BATES, University of Sussex, UK

DAVID BELLINGER, Harvard University

MARC H. BORNSTEIN, National Institutes for Mental Health

RICHARD E. BOYATZIS, Case Western Reserve University

OLGA BURENKOVA, University of Houston

ALEXANDER P. BURGOYNE, Michigan State University

RICHARD W. BYRNE, University of St. Andrews, UK

NANCY CANTOR, Rutgers University – Newark

PRIYANKA B. CARR, Stanford University

DAVID R. CARUSO, Yale University

STEPHEN J. CECI, Cornell University

ROBERTO COLOM, Autonomous University of Madrid, Spain

ANDREW R. A. CONWAY, Claremont Graduate University

CHRISTINE E. DALEY, Columbus Psychological Associates, Columbus, GA

JIM DAVIES, Carleton University, Canada

PAUL DE BOECK, Ohio State University

JASMINE M. DEJESUS, University of Michigan

COLIN G. DEYOUNG, University of Minnesota

JACQUELINE M. DI SANTO, State University of New York at New Paltz

RON DUMONT, Fairleigh Dickinson University

CAROL S. DWECK, Stanford University

DAVID HENRY FELDMAN, Tufts University

DEBORAH J. FIDLER, Colorado State University

JAMES R. FLYNN, University of Otago

GLENN GEHER, State University of New York at New Paltz

SUSAN A. GELMAN, University of Michigan

GERD GIGERENZER, Max Planck Institute for Human Development, Berlin

JUDITH GLÜCK, University of Klagenfurt, Austria, Austria

ASHOK K. GOEL, Georgia Tech

TRINIDAD GONZÁLEZ, Pontifical Catholic University of Chile

LAURENCE ROBERT GORE, Ohio State University

PATRICIA M. GREENFIELD, University of California, Los Angeles

ELENA L. GRIGORENKO, University of Houston

RICHARD J. HAIER, University of California, Irvine

DIANE F. HALPERN, Claremont-McKenna College

DAVID Z. HAMBRICK, Michigan State University

ISAAC J. HANDLEY-MINER, Stanford University

JENNIFER HEDLUND, Central Connecticut State University

SUSAN HEPBURN, Colorado State University

CHRISTOPHER HERTZOG, Georgia Tech

GABRIELA HOFER, Karl-Franzens University of Graz, Austria

MACIEJ KARWOWSKI, University of Wroclaw, Poland

ALAN S. KAUFMAN, Yale University School of Medicine

JAMES C. KAUFMAN, University of Connecticut

SCOTT BARRY KAUFMAN, Barnard College, Columbia University

HARRISON J. KELL, Educational Testing Service

JOHN F. KIHLSTROM, University of California, Berkeley

MINDY L. KORNHABER, Pennsylvania State University

KRISTOF KOVACS, Eötvös Loránd University (ELTE)

JONI M. LAKIN, Auburn University

DYLAN LARSON-KONAR, University of Florida

CHRISTINA S. LEE, Northeastern University

THOMAS W. MALONE, Massachusetts Institute of Technology

JOHN D. MAYER, University of New Hampshire

RICHARD E. MAYER, University of California, Santa Barbara

MARTHA J. MORELOCK, Tufts University

TED NETTELBECK, University of Adelaide

ALJOSCHA C. NEUBAUER, Karl-Franzens University of Graz, Austria

KOK YEE NG, Nanyang Technical University, Singapore

RAYMOND S. NICKERSON, Tufts University

WEIHUA NIU, Pace University

ANTHONY J. ONWUEGBUZIE, University of Cambridge and University of Johannesburg

JULIE A. PLANKE, State University of New York at New Paltz

JONATHAN A. PLUCKER, Johns Hopkins University

MARIA ÁNGELES QUIROGA, Complutense University of Madrid, Spain

SALLY M. REIS, University of Connecticut

JOSEPH S. RENZULLI, University of Connecticut

HEINER RINDERMANN, Chemnitz University of Technology, Germany

SUSAN E. RIVERS, iThrive Games Foundation

THOMAS ROCKSTUHL, Nanyang Technical University, Singapore

ERNESTO SAN MARTÍN, Pontifical Catholic University of Chile

EMILY SCHWORER, Colorado State University

ELLEN L. SHORT, Long Island University

KEITH E. STANOVICH, Ontario Institutes for Studies in Education

ROBERT J. STERNBERG, Cornell University

CON STOUGH, Swinburne University, Australia

HARISH SUJAN, Tulane University

MITA SUJAN, Tulane University

LISA A. SUZUKI, New York University

MOLLY SWANSON, Colorado State University

MAGGIE E. TOPLAK, York University

JONATHAN WAI, University of Arkansas

ROBERT WALRATH, Rivier University

RICHARD F. WEST, James Madison University

WENDY M. WILLIAMS, Cornell University

JOHN O. WILLIS, Regional Educational Services Center, Amherst, NH

ANITA WILLIAMS WOOLLEY, Carnegie Mellon University

THOMAS R. ZENTALL, University of Kentucky

OLIVER ZWALF, Swinburne University, Australia

Preface

Back in the late 1970s, immediately after I received my PhD, I recognized what I saw as a need for a handbook that would integrate theory, research, and practice in the field of human intelligence. The resulting product was the *Handbook of Human Intelligence*, which also was my first book with Cambridge University Press. At the time, the book seemed enormous, comprising fifteen chapters and more than a thousand printed pages. That book proved to be something of a milestone for the field. Handbooks expire after some number of years: New theory and research result in their replacement. For example, in 1982, Howard Gardner had not yet proposed his theory of multiple intelligences nor had I proposed my own triarchic theory of intelligence. It was time for a replacement. In 2000, I edited for Cambridge the *Handbook of Intelligence*, which had twenty-eight chapters. It was printed in a larger format and came in at about 700 pages. Yet this book had an even shorter shelf life because, by the beginning of the twenty-first century, research on human intelligence, which had moved slowly during the twentieth century, was on the fast track. By 2011, the first *Cambridge Handbook of Intelligence* was published, edited by myself and Scott Barry Kaufman. The book was up to forty-two chapters and more than 900 pages in a large printed format. And now I present what will be, for me, almost certainly the last edition of the handbook I edit: the *Cambridge Handbook of Intelligence* (2nd ed.), which contains fifty chapters and 1250 pages.

The monotonic increase in the number of chapters over the successive versions of the handbook represents the huge expansion of the field since the publication of the original handbook. For example, a chapter on video games would have meant an entirely different thing in 1982, before the advent of the Internet, and there was little or no research to speak of. Concepts such as emotional intelligence, collective intelligence, mating intelligence, practical intelligence, and leadership intelligence had not yet been proposed. Although social intelligence had been studied, the research had been of psychometric tests not all so different from what was to be found on standard intelligence tests. The chapter on biological foundations in the original handbook had almost no overlap with the current chapter, so much has the field changed since then. The Flynn effect had only recently been discovered and the field had not yet realized how important the finding of secular increases in intelligence would be. Chapters on intelligence and wisdom or intelligence and expertise could barely have been written because the literatures at the time were so thin.

In only the near-decade since the first *Cambridge Handbook of Intelligence* came out, the field has again been transformed. And that is the reason why I am editing this last, for me, edition of the handbook, hoping that someone else takes over the next edition. Because, although I may grow old, the field is still a relatively young one, at least in the multiple and diverse forms it takes today. It is hard even to speculate what a handbook of intelligence for 2030 will look like. And that is precisely what makes the field so exciting.

I am grateful to the staff of Cambridge University Press who have helped make this handbook possible – David Repetto and Matthew Bennett, who contracted the volume; Stephen Acerra, who has collaborated from Cambridge on this and other projects; Emily Watton, who worked on putting this monster of a book together; and the rest of the Cambridge staff who have worked so diligently to make the book what you now have before you.

PART I

Intelligence and Its Measurement

1 The Concept of Intelligence

Robert J. Sternberg

The Concept of Intelligence: The Metaphors Underlying How Intelligence Is Understood

> Upon her death bed, Gertrude Stein has been said to have inquired, "What is the answer?" Getting no answer, she said, "In that case, what is the question?"
>
> (Toklas, 1963)

Gertrude Stein's question, "In that case, what is the question?" applies to few fields as well as it does to the field of intelligence (Sternberg, 1985b, 1990). The source of many of the questions asked about intelligence is the model, or metaphor, that drives research on intelligence. Different metaphors of mind give rise to different questions about the nature of intelligence and about what various empirical phenomena relating to it mean (Sternberg, 1990). The field of intelligence has been and continues to be marked by noisy and sometimes vitriolic debates, but often the debates have been more about the best questions to ask rather than about what the answers to particular questions are.

As an example, a persistent debate in the field of intelligence is over the respective roles of nature and nurture in the manifestation of individual differences in intelligence (Mandelman & Grigorenko, 2011; Sternberg & Grigorenko, 1999). But the argument makes sense only if the terms are carefully defined. Where do interactive epigenetic effects fit in? How about gene-environment covariance? What exactly is meant by intelligence? Is the intelligence for which nature and nurture are being studied really the same thing over the life span, or is it a changing entity such that changes in effects may be due in part to changes in the entity of intelligence itself? And perhaps most important, what kind of entity is intelligence anyway – a hypothetical factorial construct, a set of processes going on in the brain, a cultural invention, or what? If it is a factorial construct, then it is perhaps stable across time and space, but if it is a cultural invention, then it certainly is not stable, and so it is not clear how one even could pin down genetic and environmental effects.

Scientists sometimes are unaware or only vaguely aware of the metaphor under which they are operating, even though that metaphor can have a major effect on the way they conceptualize the phenomena they are studying. In this chapter, I consider

This chapter draws and expands upon ideas in Sternberg (1985b, 1990).

seven metaphors: geographic, computational, biological, epistemological, sociological, anthropological, and systems. I further consider mixed metaphors, and how they work in the study of intelligence.

The Geographic Metaphor

The geographic metaphor is based on the notion that a theory of intelligence should provide a metaphorical map of the mind. The use of a map leads to particular questions:

1. What are the underlying sources of individual differences, or psychometric factors, along which people differ, that generate observed individual differences in standardized test scores?
2. How do people differ with respect to their scores on each of these psychometric factors?
3. How does the map of the mind evolve as an individual grows older? For example, do factors grow more differentiated with age, perhaps then becoming less differentiated in old age?
4. How predictive is each of the factors of performances of various kinds, such as school grades or job performance?

If one metaphor has dominated research on intelligence, it almost certainly is the geographic one. Virtually all standardized tests, which are used so widely not only in the United States, but also around the world, are based on the geographic metaphor. Although early editions of the Stanford-Binet and Wechsler tests were largely atheoretical, more recent versions are more closely tied in with geographic theories, such as Carroll's (Roid, 2003; Wechsler, 2008). The most well-known theories of intelligence (e.g., Carroll, 1993; Cattell, 1963; Spearman, 1927; Thurstone, 1938; Vernon, 1950) are based on this metaphor.

Various factorial theories compete to answer the four questions above. The competition today is probably less heated than in the past. Whereas in the past, arguments raged over whether the general factor was predominant (Spearman, 1927) or secondary and even epiphenomenal (Thurstone, 1938), today there is probably pretty widespread agreement over the hierarchical nature of geographically conceived intelligence (Willis, Dumont, & Kaufman, 2011), although the exact nature of the hierarchy is still a matter of dispute (e.g., Carroll, 1993, versus Johnson & Bouchard, 2005).

The competition has always been somewhat dubious because factors can be rotated in an infinite number of different ways, and part of the answer to the question of what geographic representation makes most sense depends on where one places the axes, and on how many orders of factors one decides to extract. There are an infinite number of possible maps that could be correct for any given geographic region, and which map is best turns out to be a question of usefulness rather than of veridicality. (Nowhere is this more clearly shown than in the gerrymandering state legislatures have done to favor one or the other political party, resulting at times in contortionist districts created to favor one political party over the other.) The locations of the points on the map do not change – just how they are labeled (districted). Each map of the mind may be useful, but only for its own particular purposes.

Because geographic theories are structural, they tend more to address questions about structure but less or not at all questions about other issues. If one's goal is to predict school or job performance in general, a single score on a general factor of an intelligence test may be adequate, but if one wishes to do more differentiated prediction, more precise scores, such as of verbal, numerical, and spatial ability, may be more useful (Lubinski, in press). Geographic theories have little or nothing to say about the mapping of factors onto the brain and they do not have a clear way to deal with transition mechanisms in cognitive development (Sternberg, 1984). Geographically based theories tend to be derived on the basis of individual-difference data so that their depiction of intelligence is in terms of sources of individual differences. If intellectual abilities exist that are common in both the nature and level of ability across people, geographic theories tend not to represent them because there are no individual differences from which to derive psychometric factors.

A difficult issue for all of the geographic theorists is how fine a mapping one wishes to create. Carroll's (1993) was very refined, Cattell's (1971) much less so. Just as different kinds of mappings are possible, so are different degrees of differentiation within the regions of those mappings. Once again, then, there is no correct answer. One comes to realize that, within the geographic metaphor, there is room for many theories to be "right," in some sense. They all represent the same ground, but with different borders and different degrees of precision.

The Computational Metaphor

The computational metaphor sees the mind as a computing device analogous to a computer. "Software" in the mind determines the mental processes of thought, just as software in a computer determines the computational processes of the computer. The metaphor has proven to be enormously productive of both theory and research (see, e.g., Hunt, 2010; Sternberg, 1985a, 1988). Perhaps because the computational metaphor in the study of intelligence was generated largely in response to the geographic one (Hunt, 1980; Sternberg, 1977), its strengths and weaknesses tend to be complementary. For example, the geographic metaphor, because of its relatively static nature, tends to be somewhat weak in addressing questions of process. In contrast, the computational metaphor tends to be well able to address questions of process. But in the computational metaphor, inferences about structure are much more indirect than they are in the geographic metaphor, in which the results of a factor analysis directly provide a structural model. Moreover, whereas data from experiments generated by geographically based theories primarily make use of individual-differences data, data from experiments generated by computationally based theories tend to focus on commonalities among people and processing. Normally, in the computational approach, the main source of variation observed is that variation that occurs across stimulus conditions rather than across subjects. The result is that the computational metaphor tends to be well able to point out commonalities in information processing rather than individual differences. Indeed, many information-processing experiments do not even consider individual differences at all, whereas psychometric studies, in contrast, virtually always do. This problem was recognized by Cronbach (1957), who spoke of the difficulty psychologists had had bridging the gap between experimental and differential psychology. Now, more than sixty years later, the problem exists as it did then.

The computational metaphor seemed, when it first was used, to be an answer to the ever-proliferating numbers of factors being posited by geographic theorists. At the time, the worst proliferator was probably Guilford (1967, 1988), who by the end of his career had increased the total number of proposed factors in his theory from 120 to 150 and then to 180. But in fact, the computational metaphor has the same problem as the geographic one. The computational metaphor provides no final answer, however, because just as factors can be subdivided endlessly, so can processes be subdivided endlessly. For example, one can speak of "encoding" stimuli, but certainly there are many subprocesses involved in figuring out what a stimulus is. Again, there is no one correct level of analysis: It depends on what one wishes to do with the theory. But it is important to realize that arguments over how finely processes should be split will be fruitless, because there is no one right answer. For example, one would wish to pay more attention to the details of how stimuli are perceived in a computer program that is designed to simulate visual perception than one would pay in a program that uses visual perception in the service of, say, inductive reasoning. Thus, even computer programs provide no final answer. When a program "encodes," how much information it needs to encode a stimulus depends on the purpose to which the encoding will be put.

Computational theorists tend to be insensitive to individual differences. As a result, these theorists have not always been quick to realize that often there is no one uniformly correct information-processing model that applies to all individuals, with respect either to performance on a given task or to performance on classes of tasks. Rather, there may well be individual differences in the processes and strategies different people use to solve a given problem or class of problems (e.g., Hickendorff et al., 2010; Sternberg & Weil, 1980).

At one time, there was a hope that individual components of information processing would map into factors (Sternberg, 1977), such that each elementary information-processing component would map into its own factor as a source of individual differences. This view was idealistic and also naïve. Information-processing components tend to cluster together into factors, rather than standing as separate factors (Sternberg, 1983). Moreover, the components do not always cluster into the factors they were expected to cluster into. Sternberg (1977) found this out early, when a preparation-response component proved to have high loadings on an inductive-reasoning factor – higher than did the components alleged to measure inductive reasoning. So there is no more a set of "correct" components than there is a set of "correct" factors.

Computational theorists have even argued among themselves as to what constitutes a true information-processing theory. Some early theorists, such as Newell and Simon (1972), viewed their computer programs themselves as theories. This made for extreme specificity in a theory, but perhaps to a fault. Would one really want to argue that even changing a line of code in a computer program resulted in a new theory, or even a serious variant of the original theory? And how general were these theories anyway? A computer program might be able to solve a given type of problem or perform a particular kind of task, but usually it would not have much generalization beyond the problem or task. Designating a program as a theory results in at least two criteria for good theories perhaps going by the wayside – parsimony and generalizability.

Other theorists, such as Schank (1972) and Anderson (2015), have viewed computer programs as operationalizations, and imperfect ones at that, of theories. Perhaps the biggest danger with computational theories is that they will fail to distinguish the forest from the trees. Because they tend to deal with information processing at a very basic level, it is easy for computational theorists to become extremely focused on the details of information processing and, at times, to lose sight of their deep motivation for studying a particular task. They may even lose sight of how performance on that task fits into a larger scheme of things. Cognitive psychology can get quite wrapped up in trees and be quite oblivious of the forest in which the trees grow.

The Biological Metaphor

Biologically based theories generally study intelligence in terms of the functioning of the central nervous system and especially the brain. Because our understanding of the brain is still rather rudimentary, biological theories are largely works in process. But some of them have come quite far (see, e.g., Haier, 2017). Many of the theories tend to be based on one or more of five types of data.

The first kind of data, briefly mentioned earlier, is actually of two subtypes – behavior-genetic and molecular-genetic data (see Plomin et al., 2012). These models typically use geographically derived theories to determine the structure of intelligence. That is, they are not models of structure or process but rather of the origins and development of intelligence. In the past, genetic and environmental factors were viewed as somehow opposed to each other, and theorists such as Jensen (1998) and Kamin (1974) saw their respective roles as arguing either for the higher (Jensen) or lower (Kamin) heritability of intelligence as defined by IQ. Hans Eysenck and Leon Kamin even wrote a book together that expressed their debating positions of higher or lower heritability (Eysenck & Kamin, 1981). Those arguing for higher heritability, such as Eysenck and Jensen, believed the heritability of IQ to be around 0.80, whereas those at the other extreme, such as Kamin, questioned whether there was any heritability at all. A consensus estimate was around 0.50 (Mandelman & Grigorenko, 2011).

The debate was not one of the more productive ones in the history of psychology. For one thing, the debate falsely assumed that genetics and environment work in opposition to, rather than in coordination with, each other. Today, there is good reason to believe that gene-environment covariation is extremely important, and it is hard cleanly to assign such covariation effects to either genes or the environment since they work in coordination (Flynn, 2016). For another thing, as Herrnstein (1973) pointed out, there is no single true heritability of intelligence or of anything else. Heritability is determined as a ratio of genetic to phenotypic variance, and the level of the ratio will depend on the amount of variation there is in a given gene pool or in a particular set of environments. Where genetic variance is low – in genetic pools that are largely homogeneous – environmentality will have to be relatively high. Where environmental variance is low – in environments where everyone, say, has a poor or a good environment – heritability will have to be relatively high. These conclusions are deductively true – that is, they have to be correct because they are a function of the mathematics of the situation. What is odd, therefore, is that Herrnstein received such a vitriolic reaction to his book and other writings of the time. He was merely stating what mathematically had to be true, and what suggested that an effort to find "true heritability" was a waste of time.

Eric Turkheimer and his colleagues (Turkheimer et al., 2003) showed that the situation is even more complex, in that the heritability of intelligence appears to vary with social class. In particular, lower social class is associated with lower heritability of intelligence, presumably because there is more variability in the environments of individuals from lower socioeconomic status (SES) than is found for individuals of higher SES. The fact that SES affects heritability of intelligence suggests, of course, that there probably are other variables that affect heritability as well. The bottom line is that simply looking for the value of heritability seems to be a fruitless search. Understanding the mechanisms aside from variability that affect heritability, however, seems like an entirely worthwhile pursuit.

A second path to understanding the biology of intelligence is via brain-scanning mechanisms, such as positron emission tomography (PET) scanning and functional magnetic resonance imaging (fMRI) analysis. Some of the earliest pioneers of this method, such as Richard Haier and colleagues (1992a, 1992b), used PET scans to measure glucose consumption in the brain while subjects performed complex tasks, such as the game of Tetris. One might have expected that the more able subjects would become deeply involved in the game and show higher levels of glucose consumption; however, the opposite was true. The more able subjects showed lower levels of glucose consumption, presumably because they found the tasks easy and did not have to work very hard on them. For more difficult tasks, the pattern was reversed, presumably because the less able subjects gave up and the more able ones worked hard to solve the problems.

More recent studies have used fMRI to isolate portions of the brain involved in particular tasks. As a result of PET and fMRI analysis, Jung and Haier (2007) proposed a new theory of intelligence, parieto-frontal integration theory (P-FIT). This theory, described in more detail in Chapter 19, argues that the most important parts of the brain for the development and execution of intelligence are in the frontal and parietal lobes of the brain. Some of the relevant areas are in the left hemisphere, others in both hemispheres. The theory emphasizes the importance of the integration of the different parts of the brain in producing intelligence. This theory, in a sense, is opposite to Howard Gardner's (2011), which emphasizes the modularity of the various aspects of intelligence.

A third type of data is the specific use of patients with various kinds of brain damage, such as HM (now known to be Henry Molaison), who had a bilateral medial temporal lobectomy in the hope of mitigating symptoms of epilepsy. As a result, HM lost most of his ability to acquire new information. Thus, he could remember much of what happened in his past, but he could not form new memories. Many of the early brain-based theories of intelligence relied heavily on knowledge gained from such patients (e.g., Gazzaniga, 1970; Levy, Trevarthen, & Sperry, 1972). These theories represent pioneering early work on brain-based theories of intelligence. The advantage of such theories was that it was possible to study what effects specific lesions had on intellectual functioning. At the same time, the methodologies raised some challenging questions. First, brain-damaged patients are scarcely typical, and it has never been clear how generalizable results from these patients are to other individuals. Second, to the extent that intelligence depends on interconnections among various brain systems, these patients are not the best individuals to study, because their brain interconnections are disrupted. One's conclusion may or may not generalize to normal brains. Third, the N's in studies of brain-damaged patients tend to be extremely small, leading to questions about how generalizable the results can be.

A fourth method has been somewhat superficial but nevertheless popular among some investigators, namely, studying head size or brain size (Pietschnig et al., 2015; Witelson, Beresh, & Kigar, 2006). There seems to be little doubt that there is a correlation between brain size and intelligence, as well as between brain integrity and intelligence. The larger question is what to make of the correlation. For example, is a larger brain a cause of greater intelligence, an effect, or both? In what specific areas does size matter? It appears greater interconnectivity between hemispheres can compensate in some degree for lesser size (Gur et al., 1999). How much do size and interconnectivity actually matter, and to what extent can they trade off? At the very least, the technology has come a long way since people simply were measuring head size, as they did in earlier centuries.

A fifth approach has been to use electrophysiological data. Electrodes are generally attached to a person's skull, and event-rated potentials (ERPs) or electroencephalogram (EEG) measurements are taken while subjects are at rest or while they are performing some task. The idea here is usually to look at patterns in the electrophysiological data or to convert the data into one or more scores and then relate the patterns or scores to measures that are believed to assess intelligence, such as the Wechsler Adult Intelligence Scales or Raven's Progressive Matrices (Neubauer, Freudenthaler, & Pfurtscheller, 1995). Such studies have revealed interesting aspects of performance, such as the relevance of certain event-related potentials (such as P300 – a positive deflection in an ERP occurring about 300 ms after the presentation of a stimulus) to alertness in task performance.

The Epistemological Metaphor

The epistemological metaphor is largely attributable to Jean Piaget. That said, Piaget was himself influenced by others, such as James Mark Baldwin. Piaget referred to his own approach to theorizing as "genetic epistemological," reflecting its joint influence by biology and philosophy. Although Piaget's theory is multifaceted, the theory as it applies to intelligence has two main parts.

One part of Piaget's account of human intelligence is his theory of equilibration, according to which the absorption of new information is achieved by a dynamic equilibrium between two complementary cognitive processes, assimilation and accommodation. The other part of Piaget's theory is his account of periods of development, starting with the sensorimotor period and ending with the formal-operational period. Piaget's theory has been enormously influential in developmental psychology and psychology in general, although today it has been superseded, at least in parts, by more modern cognitive theories (see Goswami, 2013).

Piaget's theory is not only formal itself, but draws heavily on formal logic and other aspects of the philosophy of knowledge in its development. As a result, it sometimes has been viewed as a theory of competence rather than of performance, describing the formal structures that underlie development rather than the way these structures are put into practice.

First, the theory is much more nearly complete in accounting for the development of formal and logical thought than it is in accounting for the development of intuitive and aesthetic thought. But much of intelligence is not understandable in terms of scientific modes of thought. Moreover, current cognitive research questions much of the account of scientific thinking (see Goswami, 2013). That said, the only theories that have been more influential in psychology probably are those of Freud. The goal of psychological theories probably is not to be correct in all details, but rather to have heuristic value in generating further research, and few theories have generated more research than have Piaget's.

Second, there seem to be problems with strictly stagelike theories, whether Piaget's or anyone else's. Several critiques have been written of stage theories (e.g., Brainerd, 1978), but the biggest problem is that intellectual development just does not appear to show the strictly stagelike properties that it is supposed to show according to Piaget's theory. Piaget himself was aware of the problem and introduced the concept of "horizontal decalage" to account for the fact that not all operations within a given period seem to develop at the same time. But naming a problem does not in itself solve or somehow get rid of the problem.

Third, it now appears that children can accomplish many tasks at earlier ages than Piaget thought possible (Galotti, 2016). In some cases, children defined problems differently from the way Piaget and his coworkers defined them, resulting in their performance looking poorer than it might have if they had better understood what the examiner intended to ask. In other cases, the sources of difficulty that Piaget and his coworkers attributed as critical to the problems now appear to be quite correct. The most famous example perhaps is that pointed out by Bryant and Trabasso (1971), where difficulty in transitive inference problems that Piaget attributed to reasoning proved instead to be attributable to memory.

Fourth, formal operations, as Piaget called the last period, which begins at roughly age eleven or twelve, no longer seem to be the end of the line with respect to intellectual development. A variety of research efforts have suggested periods of intellectual development that go beyond formal operations (Arlin, 1975; Case, 1984). At the same time, it has become clear that not everyone even reaches the formal-operations stage.

The Sociological Metaphor

The sociological metaphor owes as much to Lev Vygotsky (1978) as the epistemological metaphor does to Piaget. Whereas Piaget viewed intelligence as making its way from the inside, outward, Vygotsky saw it as making its way from the outside, inward. As children develop, they internalize the socioemotional processes they observe in the environment. The sociological metaphor, consequently, focuses on how socialization processes affect the development of intelligence and related constructs.

The sociological metaphor is a fairly popular one today in developmental psychology, perhaps partly in reaction to Piaget. However, its being in vogue is certainly due in large part to the importance of enculturation and socialization processes to the young. That said, there exists nothing even resembling a complete theory of intelligence that is based on the sociological metaphor. The closest is the theory of Reuven Feuerstein (1979), the late Israeli psychologist, who believed that human intelligence is fully modifiable. However, his theory is more one of cognitive modifiability than of intelligence per se.

Vygotsky emphasized the notion of a *zone of proximal development*, according to which learning occurs best with guidance from an experienced teacher at a level just beyond that at which an individual feels comfortable. The idea is that, with intervention, people can learn things they could never learn themselves. Intelligence is measured via dynamic testing (Grigorenko & Sternberg, 1998; Sternberg & Grigorenko, 2002).

Dynamic testing is like conventional static testing, in that people are tested and then inferences are made about their abilities. But dynamic tests differ from static tests in that children are provided with feedback that will help them to improve their performance. Vygotsky (1978) argued that children's ability to profit from the guided instructional feedback the children received during a dynamic testing session could function as a measure of children's zone of proximal development, or the difference between their developed abilities and their underlying capacities. Put another way, testing and instruction are treated as being of a single kind rather than as being distinct processes. This integration of testing and instruction makes sense in terms of conventional definitions of intelligence, which emphasize the importance of the ability to learn. A dynamic test directly measures processes of learning in the context of assessment rather than measuring these processes of learning indirectly as the product of past learning experiences. Such dynamic measurement is especially important when not all children have had equal opportunities in the past to learn what they need to know to perform well on tests.

The Anthropological Metaphor

Adherents to the anthropological metaphor view intelligence as, at least in part, a cultural invention (Berry, 1974; Berry & Irvine, 1986; Sternberg, 2004). On this view, intelligence is a somewhat different thing from one culture to another, because the knowledge and skills it takes to adapt to one culture will be quite different from those needed to adapt to another. Thus we may learn relatively little about the intelligence of one culture from studying intelligence in another culture, and indeed, our attempts to transfer knowledge actually may be harmful, because we may make generalizations that are likely not to be correct. From the anthropological viewpoint, the best example of this is IQ testing. In such testing, a test that is developed in one culture is often brought directly into another culture. Often, the translation does not even adequately convey the meanings of the test items to the individuals in the distinct culture. Not all adherents of the anthropological view are radical cultural relativists like Berry, but all of them believe that in order fully to understand intelligence within a culture, one needs to study that culture in its own right and not assume that generalization can be made from one culture to another.

The anthropological metaphor provides a needed counterbalance to the metaphors considered earlier, especially the geographic, computational, and biological, because it views intelligence in terms of the external world, not just the internal world of the individual. So whereas Haier (2017) argued that intelligence is entirely biological, those who believe in the anthropological metaphor might say that cognitive processes have some biological origins, but how they are manifested in the everyday world to display intelligence can vary widely from one culture to another.

A first possible problem with this metaphor is that it can go from the biological extreme to a cultural one. Intelligence probably represents some interaction of biology with culture. It is not clear that either a purely biological or a purely cultural viewpoint is as enlightening as some kind of integration of the two.

A second possible problem is that we do not have good theories of context, at least as it applies to intelligence. Context may well matter for the manifestation of intelligence, but we have nothing with the precision of the factor models in order to say just how these effects come about or even what exactly they are.

A third problem is that if intelligence is really so different from one culture to another, the implication is that any model must be painfully culturally specific. There have been attempts to remedy this apparent difficulty. For example, Sternberg (2004) has argued that the mental processes that underlie intelligence – processes like recognizing the existence of a problem defining the problem, formulating a strategy to solve the problem representing information about the problem, etc. – are universal.

But the ways in which these processes are manifested in the environment, and the ways in which they are best tested, can differ radically from one culture to another. On this view, some things are common, others are not, and part of understanding intelligence is understanding what is specific and what is not.

The Systems Metaphor

The systems metaphor is perhaps the vaguest of all of the metaphors that have been considered. I group under "systems theories" those theories that seek to understand intelligence in terms of multiple systems of intelligence or even multiple intelligences (Gardner, 2011; Sternberg, 2003). These theories are more complex, in some respects, than theories in the past generally have been, although probably no more complex than P-FIT theory. One goal of these theories is to understand intelligence in a way that transcends a single metaphor and that combines aspects of at least several of the metaphors that have been considered above.

The dangers inherent in systems theories are not readily dismissed. For example, Gardner's theory was first presented in 1983 (Gardner, 1983), but as yet there is no adequate empirical test or set of tests that provides solid empirical support for its main claims – that intelligence is modular, that there are eight multiple intelligences, and that the various intelligences are uncorrelated with each other. On the contrary, the existing evidence suggests that these claims are questionable (Haier, 2017; Visser, Ashton, & Vernon, 2006). Although there has been a lot of evidence collected in favor of Sternberg's theory of successful intelligence (Sternberg, 2003), many would question whether elements such as creative, practical, and wisdom-based skills should be included in the definition of intelligence (Hunt, 2010). Hunt, among others (see Sternberg & Kaufman, 2011), also considers the general factor of intelligence (*g*) to be more powerful than do I. So one might argue that systems theories, in attempting to be more encompassing, may go too far in stretching the concept of intelligence. In large part, whether one thinks so depends on the metaphor of mind that underlies one's thinking about intelligence. Systems theories also are broad but when one looks broadly, one sometimes misses important local details.

Conclusions

In this chapter, I have reviewed different metaphors of mind for understanding human intelligence. Battles among intelligence researchers have either been between competing theories within metaphors, or equally often, between advocates for different metaphors. Some researchers appear to believe that there is a "correct" or most basic metaphor. For example, Haier (2017) seems to believe that all of intelligence is reducible to the study of the human brain. Lest it sound like I am casting aspersions on others, I hasten to add that I once believed that I had found the "correct" metaphor. At the beginning of my career, I wrote an entire book arguing that the computational metaphor was the basic one, and that eventually, other approaches would be reduced to the computational one (Sternberg, 1977). For example, I believed that factors of intelligence recovered through the geographic metaphor were nothing more than systematic combinations of information-processing components. Thus, a factor such as "inductive reasoning" could be accounted for in terms of information-processing components such as inference (seeing relations between items), higher-order mapping (seeing relations between relations), and application (applying relations from one item to another). I was able to collect data that showed that the information-processing components of inductive reasoning that I identified accounted for a lot (usually over 80%, sometimes over 90%) of the variation in reaction-time data across item types. But there were problems suggesting that the particular theory needed revision. The regression constant – preparation-response time – was more highly correlated with measures of fluid intelligence than were the inference, mapping, and application components!

Here is the problem: One can call into question and even disconfirm theories generated within a metaphor but there is no way of disconfirming a metaphor. Metaphors are not in any absolute sense right or wrong, but rather, more or less useful for particular purposes. If someone is committed to a metaphor, over time that someone may come to believe that the metaphor really is somehow the "right" one, even though it is just one of many possible windows into the world of intelligence (or anything else). If one looks at scientific careers, people change theories with at least some degree of regularity. They rarely change metaphors. As discussed in Chapter 2 of this volume, there are some luminaries who have done so. Both Spearman and Thurstone did so (Spearman, 1923 – computational – versus Spearman, 1927 – geographic; Thurstone, 1924 – biological – versus Thurstone, 1938 – geographic), but such flexibility across metaphors is the exception rather than the rule. Investigators, once they embark on the use of a metaphor of mind, change it only relatively rarely. There are exceptions, such as but not limited to Deary (2000), who has varied in his work across geographic, computational, and biological approaches.

Metaphors are like languages: They are different ways of expressing ideas. And even languages can employ different symbol systems, such as those that use the Latin alphabet, or the Chinese hanzi symbol system of Chinese and some other Asian languages, or the gestural symbols of American Sign Language, or the Cyrillic alphabet of Russian, or simply the only vaguely specifiable code of the nonverbal communication we all use with each other. There is no one "correct" symbol system. Each symbol system can be useful for different purposes. So is it with metaphors of mind. There is no one "right" one, although different ones can be more useful for different purposes. If one wants to understand the role of the brain in intelligence, one might turn to the biological metaphor, but if one wants to understand the role of culture, the anthropological metaphor will be more useful. We maximize our learning about human intelligence when we recognize that the metaphor that serves us best is the one that best serves our theoretical and research purposes. Sometimes, combining metaphors – even mixing them – is best of all.

References

Anderson, J. R. (2015). *The adaptive character of thought*. New York: Psychology Press.

Arlin, P. K. (1975). Cognitive development in adulthood: A fifth stage? *Developmental Psychology*, *11*, 602–606.

Berry, J. W. (1974). Radical cultural relativism and the concept of intelligence. In J. W. Berry & P. R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 225–229). London: Methuen.

Berry, J. W., & Irvine, S. H. (1986). Bricolage: Savages do it daily. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 271–306). New York: Cambridge University Press.

Brainerd, C. J. (1978). The stage question in cognitive-developmental theory. *Behavioral and Brain Sciences*, *1*, 173–182.

Bryant, P. E., & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature*, *13*, 456–458.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Case, R. (1984). The process of stage transition: A neo-Piagetian view. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 20–44). New York: Freeman.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1–22.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford: Oxford University Press.

Eysenck, H. J., & Kamin, L. (1981). *Intelligence: The battle for the mind*. New York: Macmillan.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. Baltimore, MD: University Park Press.

Flynn, J. R. (2016). *Does your family make you smart? Nature, nurture, and human autonomy.* New York: Cambridge University Press.

Galotti, K. (2016). *Cognitive development: Infancy through adolescence* (2nd ed.). Thousand Oaks, CA: Sage.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gazzaniga, M. S. (1970). *The bisected brain*. East Norwalk, CT: Appleton-Century-Crofts.

Goswami, U. (Ed.) (2013). *The Wiley-Blackwell handbook of child cognitive development* (2nd ed.). New York: Wiley-Blackwell.

Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, *124*, 75–111.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Guilford, J. P. (1988). Some changes in the structure-of-intellect model. *Educational and Psychological Measurement*, *48*, 1–4.

Gur, R. C., Turetsky, B. I., Matsui, M., Yan, M., Bilker, W., Hughett, P., & Gur, R. E. (1999). Sex differences in brain gray and white matter in healthy young adults: Correlations with cognitive performance. *Journal of Neuroscience*, *19*, 4065–4072.

Haier, R. J. (2017). *The neuroscience of intelligence*. New York: Cambridge University Press.

Haier, R. J., Siegel, B., Tang, C., Abel, L., & Buchsbaum, M. S. (1992a). Intelligence and changes in regional cerebral glucose metabolic-rate following learning. *Intelligence*, *16*, 415–426.

Haier, R. J., Siegel, B. V., Jr., MacLachlan, A., Soderling, E., Lottenberg, S., & Buchsbaum, M. S. (1992b). Regional glucose metabolic changes after learning a complex visuospatial/motor task: A positron emission tomographic study. *Brain Research*, *570*, 134–143.

Herrnstein, R. (1973). *IQ in the meritocracy*. New York: Little-Brown.

Hickendorff, M., van Putten, C. M., Verhelst, N. D., & Heiser, W. J. (2010). Individual differences in strategy use on division problems: Mental versus written computation. *Journal of Educational Psychology*, *102*, 438–452.

Hunt, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology*, *71*, 449–474.

Hunt, E. (2010). *Human intelligence*. New York: Cambridge University Press.

Jensen, A. R. (1998). *The* g *factor*. Westport, CT: Praeger.

Johnson, W., & Bouchard, T., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.

Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, 30(2), 135–154; discussion 154–187. https://doi.org/10.1017/S0140525X07001185

Kamin, L. (1974). *The science and politics of IQ*. Hillsdale, NJ: Erlbaum.

Levy, J., Trevarthen, C., & Sperry, R. W. (1972). Perception of bilateral chimeric figures following hemispheric disconnection. *Brain*, *95*, 61–78.

Lubinski, D. (in press). Mapping the outer envelope of intelligence: A multidimensional view from the top. In R. J. Sternberg (Ed.), *The nature of human intelligence*. New York: Cambridge University Press.

Mandelman, S. D., & Grigorenko, E. L. (2011). Genes, environments, and their interactions. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 85–106). New York: Cambridge University Press.

Neubauer, A., Freudenthaler, H. H., & Pfurtscheller, G. (1995). Intelligence and spatiotemporal patterns of event-related desynchronization (ERD). *Intelligence*, *20*, 249–266.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Pietschnig, J., Penke, L., Wicherts, J. M., Zeiler, M., & Voracek, M. (2015). Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean?" *Neuroscience & Biobehavioral Reviews*, *57*, 411–432.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. (2012). *Behavioral genetics* (6th ed.). New York: Worth.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales* (5th ed.). Itasca, IL: Riverside Publishing Company.

Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, *3*, 552–631.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition*. London: Macmillan.

Spearman, C. (1927). *The abilities of man: Their nature and assessment*. New York: Macmillan.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1983). Components of human intelligence. *Cognition*, *15*, 1–48.

Sternberg, R. J. (Ed.) (1984). *Mechanisms of cognitive development*. San Francisco: Freeman.

Sternberg, R. J. (Ed.) (1985a). *Human abilities: An information-processing approach*. San Francisco: Freeman.

Sternberg, R. J. (1985b). Human intelligence: The model is the message. *Science*, *230*, 1111–1118.

Sternberg, R. J. (Ed.) (1988). *Advances in the psychology of human intelligence*, vol. 4. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. New York: Cambridge University Press.

Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist*, *59*(5), 325–338.

Sternberg, R. J., & Grigorenko, E. L. (1999). Genetics and intelligence. *Journal of American Academy of Child and Adolescent Psychiatry*, *38*, 486–488.

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing*. New York: Cambridge University Press.

Sternberg, R. J., & Kaufman, S. B. (Eds.) (2011). *Cambridge handbook of intelligence*. New York: Cambridge University Press.

Sternberg, R. J., & Weil, E. M. (1980). An aptitude–strategy interaction in linear syllogistic reasoning. *Journal of Educational Psychology*, *72*, 226–234.

Thurstone, L. L. (1924). *The nature of intelligence*. New York: Harcourt Brace.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago.

Toklas, A. B. (1963). *What is remembered?* New York: Holt.

Turkheimer, E., Haley, A., Waldron, M., d'Onofrio, B., & Gottesman, I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 14, 623–628. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x

Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen.

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2006). Beyond g: Putting multiple intelligences theory to the test. *Intelligence*, *34*(5), 487–502.

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale – Fourth Edition*. San Antonio, TX: Psychological Corporation.

Willis, J. O., Dumont, R., & Kaufman, A. S. (2011). Factor-analytic models of intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 39–57). New York: Cambridge University Press.

Witelson, S. F., Beresh, H., & Kigar, D. L. (2006). Intelligence and brain size in 100 postmortem brains: Sex, lateralization and age factors. *Brain*, *129*, 386–398.

2 A History of Research on Intelligence

Part 1: Pre–Twentieth-Century Origins in Philosophy

Robert J. Sternberg

Human Intelligence: The History of Theory, Research, and Measurement

The history of theory, research, and measurement of intelligence goes back a long way. Although histories often start with Galton's (1883) work in the nineteenth century, the history of research on intelligence goes back much further. What you are sure to learn in this chapter is that there is relatively little "new" in modern theories of intelligence. Many times our theories of intelligence have borrowed from the distant past, whether we are aware of it or not.

We will begin our journey into historical understanding with ancient times. Then we will work our way up to the pre-nineteenth century.

Ancient Conceptions of Intelligence

As early as the Han dynasty in China (206 BC–AD 220), an imperial examination was given to candidates for civil service jobs. Our current ideas about testing, therefore, are hardly recent: They go back into the third century before the birth of Christ. But actually, for the history of thinking about intelligence, those are relatively modern times! Here's why (see also Sternberg, 1990, on which this chapter draws and expands).

Homer

Homer, who lived in ancient Greece somewhere between the twelfth and the sixth centuries BC, already was thinking about the nature of intelligence. Homer recognized intelligence as an entity and distinguished it from other skills. In the *Odyssey*, Ulysses criticizes Euryalus, who has accused Ulysses of lacking skill in any sports:

> You are an insolent fellow – so true is it that the gods do not grace all men alike in speech, person, and understanding. One may be of weak presence, but heaven has adorned this with such a good conversation that he charms everyone who sees him; his honeyed moderation carries his hearers with him so that he is leader in all assemblies of his fellows, and wherever he goes he is looked up to. Another may be

> as handsome as a god, but his good looks are not crowned with discretion. This is your case. No god could make a finer looking fellow than you are, but you are a fool. (Adler, 1987, 5, 223)

From time to time, people have suggested that society simply get rid of the word "intelligence." The use of the concept of intelligence, or something similar to it, by Homer suggests why it is so hard for the concept to die: The concept has been around for a long time, and whether we call it "intelligence" or "discretion" or "scholastic aptitude" or "academic preparedness" or something else, the concept is likely to remain, whatever it may be called. What Homer believed to be a gift from heaven or a gift of the gods, today we would call a genetic gift. We merely attempt to express in a scientific way what Homer recognized well before the birth of Christ – that intelligence has at least some genetic (heritable) component.

Plato

Plato made his seminal contributions to philosophy in the fourth century BC, although he was born during the fifth century BC. Plato had many thoughts about intelligence. It would be impossible to review all he wrote here. One aspect of intelligence, Plato believed, is the ability to learn. In Book 5 of the *Republic*, Socrates asks Glaucon:

> When you spoke of a nature gifted or not gifted in any respect, did you mean to say that one man will acquire a thing easily, another with difficulty; a little learning will lead the one to discover a great deal; whereas the other, after much study and application, no sooner learns than he forgets; or again did you mean, that the one has a body which is a good servant to his mind, while the body of the other is a hindrance to him? – Would not these be the sort of differences which distinguish the man gifted by nature from the one who is ungifted? (Adler, 1987, 5, 359)

Glaucon agrees with Socrates that Socrates' observations are correct. Socrates further shows Glaucon that part of human intelligence is a love of learning and knowledge. Truthfulness is key and so is an unwillingness to accept falsehoods. Many of us wish the love of the truth showed up in intelligent people today as it did in Plato's philosophy!

Today, love of truth more likely might be classified as wisdom than as intelligence (see Sternberg & Jordan, 2005), but the idea that people can infer the truth and distinguish it from falsehood is very much alive in the notion of inductive reasoning as a key part of intelligence, ranging back to Charles Spearman (1923) and continuing with modern theories of intelligence (Sternberg, 1977).

In the *Theaetetus*, Theaetetus imagines that in the mind of every person exists a block of wax, which is of different sizes in different people. The block of wax further can differ in moistness, hardness, and purity. Socrates, referring back Homer, suggests that when the wax is pure and clear and also sufficiently deep, the mind easily will learn and retain and thus will not be likely to become confused. The mind will think only about things that are true, and because the impressions in the wax are clear, those impressions will be quickly distributed into their appropriate places on the block of wax. But when the wax is clearly imperfect – muddy or impure or very soft or very hard – there will be defects in a person's intellect. People with soft wax will be good learners but also be apt to forget things quickly. People with hard wax will be slow learners, but will retain longer what they learn. People whose wax is rugged or shaggy or gritty, or whose wax contains a mixture of earth and dung, will have only indistinct impressions and thus lack clarity of thought. Those individuals who have hard wax also will lack clarity of thought, because their thought will lack depth. If a person's wax is very soft, the impressions left in it will be indistinct, because they so easily can be confused or remolded (based on Adler, 1987, 7, 540).

Plato's theory, as expressed in the dialogue, is a rather primitive metaphor of mind. But is it really much weaker than any of the metaphors discussed in Chapter 1? Scientists still speak of brains as modifiable in varying degrees (Haier, 2017), and although they may not think of the brain as a block of wax, they simply have replaced that concept with more modern biology, recognizing the roles of neurons, synapses, and interconnectivity in place of Plato's wax.

Aristotle

Aristotle, the second giant of Greek philosophy after Plato, also had some fairly sophisticated views on the nature of intelligence. Aristotle lived during the fourth century BC. In the *Posterior Analytics Book I*, he conceived of intelligence in terms of "quick wit":

> Quick wit is a faculty of hitting upon the middle term instantaneously. It would be exemplified by a man who saw that the moon has a bright side always turned towards the sun, and quickly grasped the cause of this, namely that she borrows her light from him; or observes somebody in conversation with a man of wealth and defined that he was borrowing money, or that the friendship of these people sprang from a common enmity. In all these instances he has seen the major and minor terms and then grasped the causes, the middle terms. (Adler, 1987, 8, 122)

Aristotle essentially was recognizing the importance of deductive reasoning to intelligence. Indeed, syllogisms of the form "All men are mortal. Socrates is a man. Therefore, Socrates is mortal" are referred to as Aristotelian syllogisms. In the early form of Louis L. Thurstone's (1938) geographic (psychometric) theory of intelligence, deductive reasoning was identified as a factor, although it later was subsumed under other factors. And modern information-processing work has recognized deductive reasoning as an important part of intelligence (e.g., Sternberg, Guyote, & Turner, 1980; Sternberg & Weil, 1980). Aristotle's discussion of the conversation with the man of wealth also shows a sensitivity to the concept of social intelligence (Kihlstrom & Cantor, 2011; Sternberg & Smith, 1985) or what Gardner (2011) calls "interpersonal intelligence."

Augustine

St. Augustine was born in the fourth century AD and died in the fifth century AD. Earlier, in the work of both Plato and Aristotle, and in the work of most of their successors, intelligence was seen as a good. Indeed, in the writings of some of the twentieth-century psychologists, it can become difficult to distinguish intelligence from some kind of overall judgment of a person's quality or value (e.g., Herrnstein & Murray, 1994). But this view, it turns out, is not universal. It is useful to see another side of the "coin," if only for balance. In Book 4 of the *Confessions*, St. Augustine both communicated his conception of intelligence and raised questions about the value of intelligence:

> Whatever was written, either on rhetoric, or logic, geometry, music, and arithmetic, by myself without much difficulty or any instructor, I understood, Thou knowest, O Lord my God; because both quickness of understanding, and acuteness in discerning, is Thy gift: yet did I not thence sacrifice to Thee. So then it served not to my use, but rather to my perdition, since I went about to get so good a portion of my substance in my own keeping; and I kept not my strength for Thee, but wandered from Thee into a far country, to spend it upon harlotries. For what profited me good abilities, not employed to good uses? (Adler, 1987, 18, 26)

Augustine goes on to question whether those with less intelligence might be better off, in that they would be less susceptible to departing from the will of God. Augustine's point is that intelligence and its concomitants can have a dark side, not just a bright one.

Again, modern theorists have made similar points, often without realizing that they were following in the tradition of Augustine. For example, in an essay on foolishness, I have argued that intelligent people can be especially susceptible to foolish behavior, precisely because (like Augustine) they believe that their intelligence somehow insulates them from foolish behavior (Sternberg, 2005). In particular, they are more likely to overestimate the quality of their ideas, to be egocentric, and to have illusions of omniscience, omnipotence, and invulnerability, as well as to be ethically disengaged. In more recent work, I have argued that our preoccupation with intelligence has made us technologically more sophisticated, but also experts at developing destructive technologies (Sternberg, 2017). Cropley and colleagues (2010) have made a similar point with regard to creativity: Creativity is as readily used for ultimately destructive ends (think atomic bombs, air pollution, global warming) as for good ones.

Aquinas

St. Thomas Aquinas lived during the thirteenth century. Aquinas, in the first part of his classic work, *Summa Theologica*, presents his views on the intellect. He suggests that although people do not understand all things, God does. Individuals of superior intellect have understanding that is more universal and deeper, whereas those with inferior intellects have less universal understanding and lesser comprehension: "Those who are of weaker intellect fail to acquire perfect knowledge through the universal conceptions of those who have a better understanding, unless things are explained to them singly and in detail" (Adler, 1987, 19, 474).

Again, Aquinas adumbrated more modern conceptions of intelligence, according to most of which intelligence involves the ability to learn better and with lesser guidance ("Intelligence and Its Measurement," 1921; Sternberg & Detterman, 1986).

Today, learning ability is seen as central to intelligence, and those who are not intelligent are sometimes referred to as "slow learners," picking up on the idea of Aquinas that speed of learning is a distinguishing feature of intelligence.

Montaigne

Today we look at some of the choices people make – in their personal lives, in their voting behavior, on their jobs – and wonder how people can be so stupid. Montaigne addressed this question, years before we were ever born to think about it: "In the average understandings and the middle sorts of capacities, the error of opinion is begotten; they follow the appearance of the first impression, and have some color of reason on their side to impute our walking on in the old beaten path to simplicity and stupidity, meaning us who have not informed ourselves by study" (Adler, 1987, 25, 150).

Montaigne recognized that with stupidity come some other undesirable traits, such as dogmatism. Researchers today have recognized the same thing, suggesting the eagerness of some people to get stuck in old ways of thinking (Ambrose & Sternberg, 2012; Ambrose, Sternberg, & Sriraman, 2012). They also may be less likely to be inclined to think in new ways that would challenge their own beliefs (Sternberg & Grigorenko, 2001). Lipman-Blumen (2006) has taken things one step further, arguing that the reason nations so often end up with toxic leaders is not that the leaders impose themselves on the populace, but rather that the ignorant and in many ways not very bright populace seeks them out. People in the United States and every other country could learn a lot from her book.

Hobbes

Thomas Hobbes published his work *Leviathan* in 1651. In *Leviathan*, Hobbes expressed some of his views regarding the nature of intelligence. Hobbes distinguished between natural and acquired "wit." But Hobbes did not mean by the distinction between natural and acquired that which is innate versus that which is learned. Rather, by *natural*, he meant intellectual skills that are acquired by use and experience rather than through direct instruction, whereas by *acquired*, he referred to that which is instilled by culture and instruction. He believed that natural wit consists principally of two skills: "celerity of imagining," or the swiftness with which one moves from one thought to another; and "steady direction," or one's ability to move toward some approved end. He viewed intelligent people as ones who are slow in thought and who cannot easily move toward an approved end (see Adler, 1987, 23, 66).

The idea of intelligence as speed of thought has had much support in modern theorizing about intelligence (see, e.g., Jensen, 1998; Sternberg, 1985; Vernon, 1988). Current information-processing theories are in good agreement that potential speed of thought, or neuronal conduction, is one important element of intelligence. Hobbes also emphasizes what I and others have come to call practical intelligence, which draws on tacit knowledge, or what Hobbes referred to as "natural" intellectual skills – those that are not taught (Sternberg et al., 2000; Wagner, 2011). Sternberg, Wagner, and others have emphasized the role of natural intellectual skills in practical intelligence. What they instead call "tacit knowledge" is what one learns from experience, usually without direct instruction. It is knowledge that not only usually is not taught, but that often is hidden. These researchers have argued that those who are successful in their lives are people who can acquire and utilize this tacit knowledge to advantage – to learn how the world works, including those aspects of it that are often hidden from view (Sternberg, 1997).

Hobbes further recognized that emotions and motivations can interfere with the utilization of intelligence: "The causes of this difference of wits are in the passions, and the difference of passions proceedeth partly from the different constitution of the body, and partly from different education" (Adler, 1987, 23, 68).

Scholars today could say the same thing differently but it is not clear they could say it better. As all of us come to recognize, emotions can help or, at least as likely, interfere with the utilization of intelligence. Irving Janis (1972) made this point when he showed how extremely intelligent men in a variety of presidential administrations (and, at that point, they all were men) made poor decisions because they allowed their emotions and motivations to interfere with the utilization of their intelligence. On the more positive side, Carr and Dweck (2011) showed how motivation, in particular a "success mindset," can lead people to utilize their abilities more effectively. In either case, Hobbes foreshadowed modern theorists in his recognition of the importance of the passions to how intelligence actually is put into practice.

Hobbes further believed that people differ little in the intelligence with which they are born. Whereas he acknowledged that one person might be better in one skill than in another, when all was said and done, the other person would be better than the first in some other skill. Thus, men are born basically equal in ability (see Adler, 1987, 23, 24). Although this view is not widely accepted among psychologists today, it has had many proponents in modern times, including among eminent scientists (e.g., Kamin, 1974; Lewontin, Rose, & Kamin, 2017). Again, early thinkers foreshadowed many of the ideas that we have debated in recent times as though they were new ideas that people in times past never would have reflected on.

Pascal

Blaise Pascal lived during the seventeenth century. In his *Pensées*, Pascal presented his thoughts on the nature and structure of intelligence. He suggested that there are two kinds of intellect:

> The one able to penetrate acutely and deeply into the conclusions of given premises, and this is the precise intellect; the other able to comprehend a great number of premises without confusing them, and this is the mathematical intellect. The one has force and exactness, the other comprehension. Now the one quality can exist without the other; the intellect can be strong and narrow, and can also be comprehensive and weak. (Adler, 1987, 33, 1972)

Pascal believed that some people are intelligent within a broad array of fields, whereas others show their intelligence only within a narrow band of fields. Pascal further suggested that people can have either a mathematical/logical mind or an intuitive mind. In some respects, he was foreshadowing the twentieth-century distinction that arose in some theories of cognitive styles, such as category width (see Zhang, Sternberg, & Rayner, 2012 for a review of cognitive styles). Pascal was likely right about one thing: People have different patterns of intellectual abilities, with some depending more on logic and others on intuitive thinking. All modern and not so modern theories recognize individual differences in such profiles of abilities (e.g., Carroll, 1993; Davidson, 2011; Mackintosh, 2011a, 2011b).

Locke

John Locke was a well-known empiricist philosopher during the seventeenth century. Locke, in his *An Essay Concerning Human Understanding*, believed, like Pascal, in the existence of two kinds of intelligence. But Locke had a different conception from that of Pascal regarding what the two kinds of intelligence are. He distinguished between *wit*, one the one hand, and *judgment*, on the other. He argued that individuals who have a lot of the one kind of intelligence do not necessarily have a lot of the other.

> For *wit* lying most in the assemblage of ideas, and putting those together with quickness and variety, wherein can be found any resemblance or congruity, thereby to make up pleasant pictures and agreeable visions in the fancies; *judgment,* on the contrary, lies quite on the other side, and separating carefully, one from another, ideas wherein can be found the least different, thereby to avoid being misled by similitude, and by affinity to take one thing for another. (Adler, 1987, 35, 144)

Locke's distinction sounds very much like the distinction Kahneman (2013) and others have made between fast and slow thinking. One set of processes arrives at quick judgments, often with our being only minimally consciously aware of what we are doing. The other set of processes is slower and more deliberate, occurring with full, or at least what feels like full, consciousness. Locke also foreshadowed later ideas about the importance of mental speed and intelligence, which made up much of the effort of cognitive psychologists studying intelligence in the latter part of the twentieth century (see Hunt, 2010; Mackintosh, 2011b; Nettelbeck, 2011). Locke suggested that bright people are those who have their ideas in memory quickly ready at hand, a view also consistent with recent research on working memory and intelligence (Conway et al., 2011). Locke also suggested that bright people keep their ideas in an unconfused fashion, and that they are well able to distinguish one thing or idea from another. Where there are just small differences, brighter people can see them, whereas duller ones cannot. This idea is similar to that behind tests such as the Raven Matrices and to the ideas of investigators like Spearman (1923) and Sternberg (1977), who have viewed seeing analogy and disanalogy as central to intelligence.

Locke is probably most well known for his concept of the *tabula rasa.* This is the idea that individuals are born with a blank slate in their mind that they then proceed to fill up. This idea became the basis for much of the empiricist movement in philosophy, but also the associationist and later behaviorist movement in psychology (e.g., Watson, 1930/1970), according to which all behavior is a result of learning. Although today most psychologists believe that people are born with more than a "blank slate," they are indebted to Locke for recognizing the importance of learning in human behavior, an importance that is recognized today as it was so many years ago by Locke (see Schultz & Schultz, 2015).

Smith

Up to now, the predominant view described has been that there are different kinds or perhaps inborn aspects of intelligence. People are different with respect to their degrees of these various kinds or aspects of intelligence. Adam Smith, in his book *The Wealth of Nations*, presented a different point of view. Smith, best known as an economic philosopher of the nineteenth century, argued that the differences are not really in intelligence or in natural talents, but instead in how the environment affects the work people do.

> The difference of natural talents in different men is, in reality, much less than we are aware of; and the very different genius which appears to distinguish men of different professions, when grown up to maturity, is not upon many occasions so much the cause as the effect of the division of labor. The difference between the most dissimilar characters, between a philosopher and a common street porter, for example, seems to arise not so much from nature as from habit, custom, and education. When they came into the world, and for the first six or eight years of their existence, they were perhaps very much alike, and neither their parents nor playfellows could perceive any remarkable difference. About that age, or soon after, they come to be employed in very different occupations. The difference of talents comes then to be taken notice of, and widens by degrees, till at last the vanity of the philosopher is willing to acknowledge scarce any resemblance. (Adler, 1987, 39, 7–8)

Adam Smith's view is certainly different from that of most psychologists who study intelligence, who believe in at least some level of innate differences (Asbury & Plomin, 2013; Mandelman & Grigorenko, 2011). But his views are quite similar, actually, to those of some expertise theorists, most notably, Ander Ericsson (Ericsson & Pool, 2017). Ericsson has argued that virtually all differences in attained levels of expertise are a result of what he refers to as deliberate practice. For deliberate practice, (1) an individual needs to be motivated to improve their performance; (2) the practice must take into account preexisting knowledge and build on that knowledge; (3) the individual must receive informative feedback immediately in order to improve their performance; and (4) the practice needs to be repeated on a frequent basis so as to allow improvement. People improve on their jobs through deliberate practice. And Ericsson, like Smith, believes that deliberate practice, not prior abilities, is the principal basis for individual differences in expertise. So Smith, like so many other earlier theorists, was prescient in adumbrating research that would come many years after his death.

Kant

Immanuel Kant, arguably the most famous philosopher of the eighteenth century, believed that intelligence, or what Kant referred to as "the higher faculties of cognition," comprises three parts: understanding, judgment, and reason. Kant's *Critique of Pure Reason* is, at least in part, a further elaboration of this point of view. Kant suggested that the structure of logic mirrors the structure of the mind, and hence Kant may be viewed as a forerunner of the genetic epistemology of Piaget, discussed in Chapter 1.

Kant further distinguished between creative and imitative intelligence, which he referred to as *genius* in opposition to the *spirit of imitation.*

> Genius (1) is a *talent* for producing that for which no definite rule can be given, and not an aptitude in the way of cleverness for what can be learned according to some rule; and that consequently *originality* must be its primary property. (2) Since there may also be original nonsense, its products must at the same time be models, i.e., be *exemplary;* and, consequently, though not themselves derived from imitation, they must serve that purpose for others, i.e., as a standard or rule of estimating. Everyone is agreed on the point of the complete opposition between genius and the *spirit of imitation.* Now since learning is nothing but imitation, the greatest ability, or aptness as a pupil (capacity), is still, as such, not equivalent to genius. (Adler, 1987, 39, 42, 525–526)

Kant was distinguishing between the sheer rote-learning aspect of intelligence and its creative aspect, which many modern theorists have done as well (see Kaufman & Plucker, 2011; Sternberg & O'Hara, 1999). Others, such as Bruner (1977), have emphasized the importance of discovery learning as opposed to mere rote or expository learning. The twentieth-century scholar most identified with discovery learning and the importance of Kant's distinction for education is John Dewey (1938/1997), whose work helped transform much of thinking about education in the early twentieth century. What is most interesting is that some of the fundamental ideas of Dewey and Bruner were not so new, dating back to Kant in the eighteenth century!

Mill

John Stuart Mill, writing in the nineteenth century, had a dim view of the capabilities of ordinary people. Mill was especially interested in one aspect of the intellect – originality.

> Originality is the one thing which unoriginal minds cannot feel the use of . . . In sober truth, whatever homage may be professed, or even paid, to real or supposed mental superiority, the general tendency of things throughout the world is to render mediocrity the ascendant power among mankind . . . Those whose opinions go by the name of public opinion are not always the same sort of public . . . But they are always a mass, that is to say, a collective mediocrity. And what is a still greater novelty, the mass do not now take their opinions from dignitaries in Church or State, from ostensible leaders, or from books. Their thinking is done for them by men much like themselves, addressing them or speaking their name . . . I am not complaining of all this. I do not assert that anything better is compatible, as a general rule, with the present low state of the human mind. But that does not hinder the government of mediocrity from being mediocre government. (Adler, 1987, 39, 43, 298)

This is perhaps a good place to end our tour of the history of philosophical thought about intelligence – with the warning that the intelligence of humankind is not sufficient to keep humankind from descending into mediocrity and mediocre governments. Mill provided us all with a warning that without the wisdom to use our intelligence well (see also Sternberg, 2019), we doom ourselves to whatever uncertain but inauspicious fate that may await us.

In this chapter, I have briefly reviewed the thoughts of some of the pre–nineteenth-century philosophers whose ideas about intelligence would end up being important for the scientific work of their scientific successors. These thoughts provide a backdrop and foundation for modern thinking, and indeed, many of the views expressed by the philosophers noted in this chapter could as well have been expressed in the twentieth or even twenty-first century. Many of the key elements of contemporary views of intelligence – understanding, judgment, knowledge, direction, and so on – have their origins in views of intelligence that date back many centuries. If there is a lack of agreement among the philosophers, one cannot attribute it to the failures of philosophical methods. There is not all so much more agreement today among psychological scientists than there has been among philosophers of the past 3,000 years.

Science is in many respects like a relay race, wherein the scientists who race to understand the world pick up the torch from their predecessors in order better to understand whatever they are studying. Imagine, however, a relay race where, instead of picking up the race where the last racer left off, each successive racer started over at the start line, or someplace close to it. That is the situation scientists are in when they fail to educate themselves regarding the contributions of their predecessors. They may think they are picking up where others have left off, but because they are unaware of the contributions of the past, they end up remaking them. No one wants to start the race over in each successive generation, and that is why it is so important to understand the history of the field.

Modern thinking about intelligence is actually quite far advanced over some of the ideas of early philosophical thinkers. But the one thinker who perhaps most stands out, at least for me, is Mill. John Stuart Mill reminds us that the intelligence of humankind is not enough to save it from descending into mediocrity and worse. If one is a true believer in IQ, then one can point to the work of Flynn (2012), who reported that IQs showed an astonishing gain of thirty points during the twentieth century. But the twentieth century was also the century of two world wars, of the invention of thermonuclear bombs with the capability of destroying all of humankind, and of genocides that occurred not only at the beginning of the century, but all throughout it (Totten & Parsons, 2008). If we all do not learn how to use our intelligence more wisely, our intelligence in and of itself may not be enough to take our civilization to the heights that we all might have hoped it was capable of attaining.

References

Adler, M. J. (Ed.) (1987). *Great books of the Western world*, 54 vols. Chicago: Encyclopedia Britannica.

Ambrose, D., & Sternberg, R. J. (Eds.) (2012). *How dogmatic beliefs harm creativity and higher level thinking*. New York: Taylor & Francis.

Ambrose, D., Sternberg, R. J., & Sriraman, B. (Eds.) (2012). *Confronting dogmatism in gifted education*. New York: Taylor & Francis.

Asbury, K., & Plomin, R. (2013). *G is for genes: The impact of genetics on education and achievement*. New York: Wiley-Blackwell.

Bruner, J. (1977). *The process of education*. Cambridge, MA: Harvard University Press.

Carr, P. B., & Dweck, C. S. (2011). Intelligence and motivation. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 748–770). New York: Cambridge University Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Conway, A. R. A., Getz, S. J., Macnamara, B., & Engel de Abreu, P. M. J. (2011). Working memory and intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 394–418). New York: Cambridge University Press.

Cropley, D. H., Cropley, A. J., Kaufman, J. C., & Runco, M. A. (Eds.) (2010). *The dark side of creativity*. New York: Cambridge University Press.

Davidson, J. E. (2011). Contemporary models of intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 58–82). New York: Cambridge University Press.

Dewey, J. (1938/1997). *Experience and education*. New York: Free Press.

Ericsson, K. A., & Pool, R. (2017). *Peak: Secrets from the new science of expertise*. New York: Eamon-Dolen/Mariner Books.

Flynn, J. R. (2012). *Are we getting smarter? Rising IQ in the twenty-first century*. Cambridge, UK: Cambridge University Press.

Galton, F. (1883). *Inquiry into human faculty and its development*. London: Macmillan.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. New York: Basic.

Haier, R. J. (2017). *The neuroscience of intelligence*. New York: Cambridge University Press.

Herrnstein, R., & Murray, C. (1994). *The bell curve*. New York: Free Press.

Hunt, E. (2010). *Human intelligence*. New York: Cambridge University Press.

"Intelligence and its measurement: A symposium" (1921). *Journal of Educational Psychology*, *12*, 123–147, 195–216, 271–275.

Janis, I. L. (1972). *Victims of groupthink*. Boston: Houghton Mifflin.

Jensen, A. R. (1998). *The* g *factor: The science of mental ability*. Westport, CT: Praeger-Greenwood.

Kahneman, D. (2013). *Thinking, fast and slow*. New York: Farrar, Straus, Giroux.

Kamin, L. J. (1974). *The science and politics of IQ*. Hillsdale, NJ: Erlbuam.

Kaufman, J. C., & Plucker, J. A. (2011). Intelligence and creativity. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 771–783). New York: Cambridge University Press.

Kihlstrom, J. F., & Cantor, N. (2011). Social intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 564–581). New York: Cambridge University Press.

Lewontin, R. C., Rose, S., & Kamin, L. J. (2017). *Not in our genes: Biology, ideology, and human nature* (2nd ed.). Chicago: Haymarket Books.

Lipman-Blumen, J. (2006). *The allure of toxic leaders: Why we follow destructive bosses and corrupt politicians . . . and how we can survive them*. New York: Oxford University Press.

Mackintosh, N. J. (2011a). History of theories and measurement of intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 3–19). New York: Cambridge University Press.

Mackintosh, N. J. (2011b). *IQ and human intelligence*. New York: Oxford University Press.

Mandelman, S. D., & Grigorenko, E. L. (2011). Intelligence: Genes, environments, and their interactions. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 86–106). New York: Cambridge University Press.

Nettelbeck, T. (2011). Basic processes of intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 371–393). New York: Cambridge University Press.

Schultz, D. P., & Schultz, S. E. (2015). *A history of modern psychology*. Belmont, CA: Wadsworth Publishers.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition*. London: Macmillan.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (Ed.) (1985). *Human abilities: An information-processing approach*. San Francisco: Freeman.

Sternberg, R. J. (1990). *Metaphors of mind*. New York: Cambridge University Press.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.

Sternberg, R. J. (2005). Foolishness. In R. J. Sternberg & J. Jordan (Eds.), *Handbook of wisdom: Psychological perspectives* (pp. 331–352). New York: Cambridge University Press.

Sternberg, R. J. (2017). ACCEL: A new model for identifying the gifted. *Roeper Review*, *39* (3), 139–152.

Sternberg, R. J. (2019). Why people often prefer wise guys to guys who are wise: An augmented balance theory of the production and reception of wisdom. In R. J. Sternberg & J. Glueck (Eds.), *Cambridge handbook of wisdom* (pp. 162–181). New York: Cambridge University Press.

Sternberg, R. J., & Detterman, D. K. (Eds.) (1986). *What is intelligence?* Norwood, NJ: Ablex Publishing Corporation.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J., Snook, S., Williams, W. M., et al. (2000). *Practical intelligence in everyday life*. New York: Cambridge University Press.

Sternberg, R. J., & Grigorenko, E. L. (2001). A capsule history of theory and research on styles. In R. J. Sternberg & L. F. Zhang (Eds.), *Perspectives on thinking, learning and cognitive styles* (pp. 1–21). Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J., Guyote, M. J., & Turner, M. E. (1980). Deductive reasoning. In R. E. Snow, P. A. Federico, & W. E. Montague (Eds.), *Aptitude, learning, and instruction: Cognitive process analyses of aptitude* (vol. 1, pp. 219–245). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J., & Jordan, J. (Eds.) (2005). *Handbook of wisdom: Psychological perspectives*. New York: Cambridge University Press.

Sternberg, R. J., & O'Hara, L. A. (1999). Creativity and intelligence. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 251–272). New York: Cambridge University Press.

Sternberg, R. J., & Smith, C. (1985). Social intelligence and decoding skills in nonverbal communication. *Social Cognition*, *2*, 168–192.

Sternberg, R. J., & Weil, E. M. (1980). An aptitude–strategy interaction in linear syllogistic reasoning. *Journal of Educational Psychology*, *72*, 226–234.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Totten, S., & Parsons, W. S. (Eds.) (2008). *Century of genocide: Critical essays and eyewitness accounts*. London: Routledge.

Vernon, P. A. (Ed.) (1988). *Speed of information processing*. Westport, CT: Praeger.

Wagner, R. K. (2011). Practical intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 550–563). New York: Cambridge University Press.

Watson, J. (1930/1970). *Behaviorism*. New York: W. W. Norton.

Zhang, L.-F., Sternberg, R. J., & Rayner, S. (Eds.) (2012). *Handbook of intellectual styles*. New York: Springer.

3 A History of Research on Intelligence

Part 2: Psychological Theory, Research, and Practice in the Nineteenth and Twentieth Centuries

Robert J. Sternberg

Forefathers of the Testing Movement in Intelligence

Francis Galton

Sir Francis Galton (1883), one of the earliest researchers on human intelligence, proposed a theory of the "human faculty and its development." Because Galton also proposed techniques for measuring the "human faculty," his theory could be applied directly to human behavior. The combination of psychological theory and measurement techniques created a new way of understanding human intelligence.

Galton proposed two general qualities that distinguish the more from the less intelligent. The first quality is energy, which Galton viewed as tantamount to the capacity for labor. Galton suggested that intellectually gifted people in a wide variety of fields are characterized by high levels of energy. The second quality is psychophysical sensitivity. Galton suggested that the only information that can reach us with regard to external events passes through the senses. The more perceptive the senses are of differences in sensory qualities, such as luminescence, pitch, odor, or whatever, the greater would be the range of information on which intelligence could act. Galton spoke crudely and with dubitable validity of the lower range of the intellectual spectrum. In his 1883 book, he claimed, "The discriminative facility of idiots is curiously low; they hardly distinguish between heat and cold, and their sense of pain is so obtuse that some of the more idiotic seem hardly to know what it is. In their dull lives, such pain as can be excited in them may literally be accepted with welcome surprise" (p. 28).

The quotation probably tells us much more about Galton than it does about people with low intelligence. It is unclear whether he had any evidence to support his analysis. But we have evidence of his clear disdain for such people. Certainly, there is no compassion or even pity in his statements. Were it only Galton, perhaps one could write off the man as idiosyncratic. But too much of the history of intelligence work reveals similar attitudes toward people of low intelligence (e.g., Goddard, 1912, 1914; Terman, 1916).

During a seven-year period (1884–1890), Galton managed an "anthropometric laboratory" at the South Kensington Museum in London. Here, for a small amount

This chapter draws and expands upon ideas in Sternberg (1990).

of money, visitors could have themselves assessed on a number of psychophysical tests. One of these tests was weight discrimination. The weight-discrimination apparatus consisted of a number of cases of gun cartridges, each filled with alternating layers of shot, wool, and wadding. The cases all were identical in appearance; they differed only in their weight. The weights, from least to most heavy, together formed a geometric series of heaviness. The examiner recorded the smallest interval that an examinee could discriminate. Galton noted that similar geometric sequences could be used for testing other senses as well, such as touch and taste. With touch, Galton proposed the use of wirework of varying degrees of fineness. For taste, Galton proposed the use of stock bottles of solutions of salt of various strengths. For olfaction, Galton suggested the use of bottles of attar of rose in varying degrees of dilution.

Galton also created a whistle for ascertaining the highest pitch that various individuals could perceive. Tests using the whistle enabled Galton to discover that people's ability to hear high notes declines considerably with age. Galton also discovered that people are inferior to cats with regard to their skill in perceiving tones of very high pitch. This finding is problematic for any psychophysically based theory of intelligence that propounds a notion of evolutionary continuity. It suggests that, in at least this one respect, cats are superior in intelligence to humans. Although one may grant this superiority to cats, one then will be obliged to grant superiority to various animals in many other psychophysical characteristics, leaving humans in a mediocre position to which they are not accustomed!

The critical thing to note about Galton is that he was the first to study intelligence in anything that reasonably could be labeled a scientific way. Even if his theory seems a bit quaint today – does a person who is nearsighted deserve to be labeled as less intelligent than a person with 20/20 vision? – Galton recognized the need to formulate a theory and test it with empirical data. That was a major and, up to that time, unique contribution.

James McKeen Cattell

James McKeen Cattell carried many of Galton's seminal ideas across continents to the United States. As head of a major psychological laboratory at Columbia University, Cattell was in an excellent position to publicize Galton's psychophysical approach to intelligence. Cattell (1890) believed that psychology, as a science, could not attain the certainty and exactness of the physical sciences unless it was built on a foundation of experimentation and measurement. A step in this direction was to devise mental tests based on Galton's notions.

Cattell (1890) proposed a series of fifty psychophysical tests, ten of which he described in detail, for example:

1. *Dynamometer pressure.* The dynamometer pressure resulting from the greatest possible squeeze of one's hand.
2. *Rate of movement.* A measure of the fastest possible movement of the right hand and arm from rest through a distance of 50 cm. Cattell measured rate of movement by the use of an apparatus that opened up an electric current when the hand was first moved and closed down the current when the hand came to rest.

3. *Sensation areas.* A measure of the distance on the skin by which two points must be separated in order for them to be felt as distinct and separate points. Cattell recommended that the back of the closed right hand between the first and second fingers be used as the basis for measurement of these areas.
4. *Pressure causing pain.* The point at which the sensation of pressure first causes noticeable pain may be an important constant. Cattell used the tip of hard rubber 5 mm in radius as his pain-producing instrument. He suggested the center of the forehead as the place at which to apply the instrument.
5. *Least noticeable difference in weight.* Measured with small wooden boxes. Subjects were handed two such boxes and then were asked to indicate which box was heavier.

Cattell thus took Galton's ideas, refined them, and made them into what arguably was the first truly structured and carefully constructed test of intelligence. Again, the theory may have left something to be desired, but Cattell set, in the United States, a standard for scientific research on intelligence. He moved the weight of intelligence research from the realm of philosophy to the realm of the budding discipline of psychology.

Psychophysical tasks, though, find little or no place in modern-day tests of intelligence as administered in schools and in industry. Indeed, such tests ceased to play an important role in practical mental measurement by the turn of the twentieth century, although they are coming back into vogue among certain theoreticians. The tests were ultimately discarded because they were found to display a chance pattern of correlations both with each other and with the external criteria used to validate the tests.

The coup de grace for Cattell's measurements was administered by a student in Cattell's own laboratory, Clark Wissler. Wissler (1901) proposed that Cattell's tests should correlate with each other and with external criteria of academic success, such as grades in the undergraduate program at Columbia University. In fact, Wissler found they correlated with neither. His findings seemed to invalidate the Galton-Cattell theory. It was time for a new view on intelligence to come to the attention of scientists. That view proved to be the view of Alfred Binet.

Alfred Binet

In 1904, the minister of public instruction in Paris named a commission charged with studying or creating tests that would ensure that mentally challenged children received an adequate education. The commission decided that no child suspected of mental challenges should be placed in a special class for such children unless it could be certified that the child truly was mentally challenged. Binet and Simon (1916a) devised tests to meet this placement need.

Thus, whereas theory and research in the tradition of Galton grew out of purely scientific concerns, research in the tradition of Binet grew out of practical education concerns. There was one other key difference between Binet and Galton. Whereas Galton showed obvious disdain for intellectually challenged children, Binet was intent on helping them. As we shall see, however, the tests he helped create could be used by people with a view toward challenged children similar to Binet's, or by people with a view toward such children similar to Galton's.

At the time, definitions of various degrees of subnormal intelligence lacked both precision and standardization. Personality, on the one hand, and intellectual deficits, on the other, were seen as being of the same ilk. So a child with a perceived personality issue might be placed in a class for intellectually subnormal children. The risk, of course, was that schools would dump children whom they perceived to be problematic in those classes simply to reduce their own workload. Binet and Simon presented a different point of view. The believed that the core of intelligence is:

> judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting one's self to circumstances. To judge well, to comprehend well, to reason well, these are the essential activities of intelligence. A person may be a moron or an imbecile if he is lacking in judgment; but with good judgment he can never be either. Indeed, the rest of the intellectual faculties seem of little importance in comparison with judgment. (Binet & Simon, 1916a, pp. 42–43)

Binet cited Helen Keller as an example of someone of known extraordinary intelligence whose scores on psychophysical tests would be poor and, yet, who could be expected to perform at a high level on Binet's tests of judgment.

Binet is sometimes thought of as test-driven or as atheoretical (Hunt, Frost, & Lunneborg, 1973), but he was neither. Binet and Simon's (1916a) theory of intelligent thinking in many ways foreshadowed the research being done many years later on the development of metacognition (e.g., Brown & DeLoache, 1978; Flavell, 1981; Flavell & Wellman, 1977; Sternberg, 1984a). Binet and Simon (1916b) proposed that intelligent thought comprises three distinct but interrelated elements: direction, adaptation, and control.

Direction involves knowing what it is that has to be done and how it is to be accomplished. For example, when a person is asked to add two numbers, the person gives themselves a series of prompts on how to proceed with the task. These prompts then constitute the direction of thought. The prompts need not always be conscious. In many respects, these ideas about the development of direction over time foreshadow much later theorizing regarding automaticity of information processing (see, e.g., Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). According to Binet and Simon (1916b), mentally challenged individuals show an absence or weakness of direction that manifests itself in two different forms: "Either the direction, once commenced, does not continue, or it has not even been commenced because it has not been understood" (p. 138).

Adaptation involves one's selection and monitoring of one's strategies while performing a task. Thoughts consist of a series of selections. One can think this way, or one can think that way. Brighter children are more likely to select and monitor their strategies in an efficient way. This idea is directly relevant to metacognitive ideas, such as Sternberg's theory of metacomponents in learning and problem-solving (Sternberg, 2001).

Finally, *control* is the ability to criticize one's own thoughts and actions. Binet and Simon (1916b) believed that much of this ability is exercised below the conscious level. Less intelligent individuals show a lack of control. Their actions often are inappropriate to the task at hand. For example, a mentally challenged child "told to copy an 'a' scribbles a formless mass at which he smiles in a satisfied manner" (p. 149).

Binet and Simon further distinguished between two types of intelligence, ideational and instinctive intelligence. Ideational intelligence functions through words and ideas. It makes use of logical analysis and verbal reasoning. Instinctive intelligence functions by means of feeling. It refers not to the instinct attributed to other animals and to simple forms of human behavior, but rather to the "lack of a logical perception, of a verbal reasoning, which would permit of explaining and of demonstrating a succession of truths" (Binet & Simon, 1916b, p. 316). It seems to be similar to an intuitive sense. Thus, Binet and Simon, like some of the philosophers who preceded them, had a sense of the kinds of fast and slow processing that would be important to cognitive theory many years later (Kahneman, 2013).

Binet and Simon's tests were early attempts to operationalize Binet's conception of intelligence. Binet came up with the concept of mental age, or the comparable age level at which a child performs on a test, regardless of their chronological age. Binet, unlike Galton, believed that intelligence is modifiable, and even designed what he called "mental orthopedics" to help children improve their cognitive skills. But it remained for Lewis M. Terman, his successor in the United States, to operationalize the tests in a standardized way. Terman never operationalized the mental orthopedics, nor is it likely he would have been interested, as he was strongly hereditarian.

Lewis M. Terman

Lewis M. Terman was a world-renowned professor of psychology at Stanford University in California. He created the earliest versions of what since have come to be called the Stanford-Binet Intelligence Scales (Terman & Merrill, 1937, 1973). I describe in this chapter Terman's historical versions of the test.

In the 1973 version of the tests, testing starts at age two. Examples of tests at this level are a three-holed form board, which requires children to put circular, square, and triangular pieces into holes on a board of appropriate shape; identification of parts of the body, which requires children to identify body parts on a paper doll; block building, which requires children to build a four-block tower; and picture vocabulary, which requires children to identify pictures of common objects.

Six years later, by chronological age eight, the nature of the tests changes considerably. However, the tests still are measuring the kinds of complex cognitive processes that the tests appropriate for age two attempt to tap. For age eight, the tests comprise vocabulary, requiring children to define words; verbal absurdities, requiring children to recognize why each one of a set of statements is foolish; similarities and differences, requiring children to indicate how each of two objects is the same as but also different from the other; comprehension, requiring children to solve practical problems of the kind found in everyday life; and naming the days of the week.

 https://doi.org/10.1017/9781108770422.004

Six years later, when subjects reach age fourteen, the tests are similar, although overall, they are more difficult. The tests comprise vocabulary; induction, in which the test administrator makes a notch in an edge of some folded paper and asks test-takers how many holes the paper will have when it is unfolded (note that the test appears to be more spatial than inductive in character); reasoning, involving solving of an arithmetic word problem; ingenuity, requiring individuals to specify the series of steps that one could use to pour a given amount of liquid; spatial directions; and reconciliation of opposites, requiring individuals to say in what ways two opposites are actually alike. The most difficult level of the Stanford-Binet "Superior Adult III" includes assessments of vocabulary, interpreting proverbs, spatial orientation, reasoning, repeating main ideas of a story, and solving of analogies.

IQ on the Stanford-Binet formerly (although no longer) was calculated based on the ratio of mental age to chronological age, multiplied by 100. So if a child had a mental age of twelve (i.e., was capable of doing the mental work of a twelve-year-old) and a chronological age of ten (was ten years old physically), the child's IQ would be (12/10) x 100 = 120.

Terman is probably better known for his applied work rather than for his theoretical work. Beyond the intelligence test, he is most well known for a longitudinal study of the gifted that he and his successors conducted (e.g., Terman, 1925; Terman & Oden, 1959). In his sample of the gifted, Terman included California children under age eleven with IQs over 140 as well as children in the 11–14 age bracket with slightly lower IQs (to allow for the lower ceiling at this age in test scores). The mean IQ of the 643 subjects selected was 151; only twenty-two of the subjects had IQs of under 140. The accomplishments in later life of the selected group were notable. By 1959, there were seventy listings among the group in *American Men of Science* and three memberships in the National Academy of Science. Moreover, thirty-one men were listed in *Who's Who in America* and ten appeared in the *Directory of American Scholars.* There were numerous highly successful businessmen as well as individuals who were succeeding well in all of the professions. The sex bias in these sources is obvious. Because most of the women became housewives, it is impossible to make any meaningful comparison between the men on the one hand (none of whom were reported to be househusbands) and the women on the other.

The Terman study was a landmark contribution in that it showed a clear association between IQ and various real-world measures of life success. Moreover, it helped dispel the notion that people with high IQs are oddballs, or physically infirm, or even tending toward the mentally ill. All of these were fairly common stereotypes at the time the study was commenced. At the same time, it was a correlational study, so it could not show causal effects. How much of any effect was due to confounding factors – such as socioeconomic status or parental investment – just is not clear. Moreover, although there were many successes, one might argue there were regression effects: Few of the individuals seem to have reached percentiles of success comparable to the percentiles of their IQs. Indeed, the study generally has been touted as having large percentages of success, but would not the same be true of children who simply had high school marks at an early age? All that said, the study was tremendously important in suggesting that investment in the gifted could pay serious dividends. It was also important in a way that some, including myself, would argue was not so positive (Sternberg, 1984b, 1985, 2003; Sternberg & Davidson, 1982, 1983): It led to many gifted programs emphasizing IQ as a basis for identification at the expense of serious consideration of other measures as well.

David Wechsler

David Wechsler, a twentieth-century American psychologist, followed in the tradition of Alfred Binet in emphasizing higher cognitive processes in his test of intelligence (e.g., Wechsler, 1939, 1958, 1974). But his notion of intelligence, although related to Binet's, was somewhat different: "Ultimately, intelligence is not a kind of ability at all, certainly not in the same sense that reasoning, memory, verbal fluency, etc., are so regarded. Rather it's something that is inferred from the way these abilities are manifested under different conditions and circumstances" (Wechsler, 1974, pp. 5–6); "Intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment" (Wechsler, 1944, p. 3).

The Wechsler scales, like the Stanford-Binet scales, are wide ranging in their content. However, these scales do not do full justice to the breadth of Wechsler's original conception of the nature of intelligence. Indeed, it is unlikely that, even today, any scale could be created that would fully do justice to the broad conceptions of intelligence of Binet and Wechsler.

In the 1974 revision of the Wechsler Intelligence Scale for Children (WISC-R), appropriate for children ranging in age from six to sixteen, there are twelve subtests. Ten of these subtests are considered mandatory and two are optional. The content of the revised test is nearly identical to that of the adult scale, except that items are somewhat easier, as befits their lower targeted age range. The tests are divided into two distinct parts, verbal tests and performance tests. Each part yields a separate deviation IQ, meaning that IQ is determined by converting percentile scores into an IQ using normal-distribution IQ equivalents of percentile scores. As with the Stanford-Binet, the test must be administered individually. In both the Wechsler and Stanford-Binet tests, an examiner administers only items that are appropriate to the age and ability of the subject. Subjects begin with easier items appropriate for their age, and work their way up. They then end with items that are difficult enough to result in repeated failure to get the items correct.

The verbal part of the test includes as subtests the following: information, which requires demonstration of knowledge about the world; similarities, which requires an indication of a way in which two different objects are alike; arithmetic, which requires the solution of arithmetic word problems; vocabulary, which requires definition of common English words; comprehension, which requires understanding of societal customs; and optionally, digit span, which requires recall of strings of digits forward in one section and of digits backward in another section. The performance part of the tests includes as subtests the following: picture completion, which requires recognition of a missing part of a picture of an object; picture arrangement, which requires rearrangement of a scrambled set of pictures into an order that tells a coherent story from beginning to end; block design, which requires children to reproduce a picture of a design, constructed from a set of red, white, and half-red /half-white blocks, by actually building the design with physical blocks; object assembly, which requires children to manipulate jigsaw puzzle pieces to form a picture of a common object in the real word; coding (the analogue of digit-symbol at the adult level), which requires rapid copying of symbols that are paired with pictures of objects according to a prespecified key that links the pictures with the objects; and mazes, which requires tracing a route through each of a set of mazes from beginning to end.

Early Theories of Intelligence

My goal here is merely to provide a brief historical overview of the development of theories of intelligence in the twentieth century, with emphasis on the early part of the century.

Charles Spearman and Two-Factor Theory

English theorist Charles Spearman, the father of psychometric theorizing about intelligence, proposed a two-factor theory of intelligence (1904, 1927). Spearman believed that intelligence comprises two kinds of factors – *general intelligence*, which is measured by all subtests in a test battery, and *specific intelligence*, which is specific to each particular test in a test battery. Actually, the name of the theory is something of a misnomer. Spearman was not claiming that there are two factors of intelligence, but rather, two kinds of factors. Because the former is more general, Spearman was far more interested in general intelligence.

What exactly is *g*, the general factor? Spearman considered several different alternative explanations, including attention, will, plasticity of the nervous system, and the state of the blood. In the end, his preferred explanation (Spearman, 1927) was mental energy.

Spearman suggested that the concept of mental energy dated back to Aristotle. But the concept of mental energy had a different meaning for Aristotle than it had for Spearman. For Aristotle, energy was triggered by any actual manifestation of change. For Spearman, one's total output and the mental energy on which it is based is a constant. Constancy in mental energy, and thus in output, is achieved by a kind of universal mental competition: The beginning of one mental activity causes some other activity to stop, whereas conversely, the ending of that other activity would allow the first one to begin. Put another way, we each have a fixed amount of mental energy, which we can allocate at different times to different tasks. Differences in general intelligence are due to our differences in amounts of mental energy.

Spearman (1923) also had another, less well known theory. Spearman proposed three qualitative principles of cognition, which formed the basis for the computational metaphor of intelligence. *Apprehension of experience* refers to what today might be referred to as "encoding." It is the taking in of information that later can be used in reasoning and problem-solving. *Eduction of relations* refers to inferring relations between two concepts or objects (see Sternberg, 1977). And *eduction of correlates* refers to the application of inferred relations in new situations (see Sternberg, 1977). An analogy, A : B : C : ?, is paradigmatic for the implementation of these operations. Encoding is involved in figuring out what the terms are saying. Eduction of relations is used to figure out the relation between A and B, and eduction of correlates is used to apply that relation to generate the missing D term.

Spearman (1927) also suggested five quantitative principles of cognition. *Mental energy*, also considered as the basis of *g*, is used to activate the mind. *Retentivity* refers to the occurrence of a cognitive event producing a tendency for that same event to occur again later. *Fatigue* refers to any cognitive event producing a tendency opposed to its occurring afterward. *Conative control* refers to the intensity of cognition as controlled by motivation. And *primordial potencies* refers to manifestations of the preceding four quantitative principles on action.

Godfrey Thomson and the Theory of Bonds

Spearman's theory was criticized by a number of other theorists. One of Spearman's chief critics was an English psychologist who spent much of his time working in Edinburgh, Scotland. Thomson (1939) suggested that it was possible to have a general factor of intelligence in the absence of any real general ability in the mind. To Thomson, *g* was not a psychological reality but rather a statistical artifact, with the artifact stemming from the fallacious view that a single factor necessarily means that there is a single psychological construct corresponding to it.

Thomson suggested that the general factor might result from the workings of an extremely large number of what he called "bonds," many of which are sampled simultaneously in intellectual tasks. Imagine, for example, that each of the intellectual tasks found in the test batteries of Spearman and other psychologists requires particular mental skills. Thomson never actually indicated what these mental skills or bonds might be. He simply linked them to the functioning of the neurons in the brain. The general factor is the result of a massive simultaneous sampling of overlapping bonds on multiple tests.

Thorndike and the Theory of Connections

A theory in some respects similar to that of Thomson was proposed by Edward L. Thorndike, who is much better known for his work on learning theory than for his work on intelligence. Thorndike and colleagues (1926) suggested that:

> in their deeper nature the higher forms of intellectual operations are identical with mere association or connection forming, depending upon the same sort of physiological connections but requiring *many more of them.* By the same argument the person whose intellect is greater or higher or better than that of another person differs from him in the last analysis in having, not a new sort of physiological process, but simply a larger number of connections of the ordinary sort. (p. 415)

Thorndike and colleagues actually proposed a measure, *c*, the number of connections a person has, which they suggested is a measure of intelligence. Scores on intelligence tests were hypothesized to be indirect measures of this number of connections (*c*). The link between Thorndike's and Thomson's theories is shown by the fact that Thorndike referred to these connections in some instances as "bonds." Thorndike's view also anticipated the later view of Cattell (1971), whose conception of crystallized intelligence (measured by tests such as vocabulary) was rather similar to Thorndike's conception of bonds.

Thorndike examined many kinds of tests for assessing intelligence. He concluded that four particular tests work best. His view on testing is sometimes referred to as "CAVD theory." CAVD is an acronym for Completions, Arithmetic problems, Vocabulary, and Directions. Completions involve supplying words so as to make a statement true and sensible. Arithmetic problems simply require solution of quantitative word problems. Vocabulary is the measurement of knowledge of single words (which is how Cattell's "crystallized intelligence" is usually measured). Directions provides an assessment of the person's ability to understand connected discourse, as in oral directions or paragraph reading.

Louis L. Thurstone

Louis L. Thurstone, like Charles Spearman, ardently believed in the usefulness of factor analysis. He proposed a theory of primary mental abilities (Thurstone, 1938), according to which intelligence can be understood in terms of seven distinct although interrelated abilities:

1. *Verbal comprehension.* Involves a person's ability to understand verbal material. This ability is assessed by tests such as vocabulary and reading comprehension.
2. *Verbal fluency.* Involves rapidly producing words, sentences, and other verbal material. This ability is assessed by tests that require the examinee, for example, to produce as many words as possible in a brief period of time beginning with a particular letter, such as "n."
3. *Number.* Involves rapid arithmetic computation and solving simple arithmetic word problems.
4. *Memory.* Involves remembering strings of words, or symbols such as letters or numbers, or remembering people's faces.
5. *Perceptual speed.* Involves proofreading and rapidly recognizing letters and numbers.
6. *Inductive reasoning.* Involves generalization – reasoning from the specific to the general. It is assessed by tests such as letters series, number series, and analogies.
7. *Spatial visualization.* Involves visualizing shapes, rotations of objects, and how pieces of a puzzle fit together.

In general, Thurstone was antagonistic to Spearman's theory of *g*. Thurstone suggested that Spearman's general factor was an artifact of Spearman's failure to rotate his factorial axes on obtaining an initial solution. Spearman, predictably, believed the opposite, suggesting that Thurstone's primary mental abilities were subsidiary to *g*. Indeed, the primary mental abilities are correlated, so that if one analyzes these factors, one obtains a *g*-factor. It is clearly not meaningful to ask whether the lower-order factors or the higher-order factor is somehow more "basic."

Holzinger and Bifactor Theory

Holzinger (1938) proposed a bifactor theory, which retained the general and specific factors of Spearman, but also permitted the group factors of Thurstone. The theory was thus an attempt to integrate Spearman's and Thurstone's theories and to show that they are not conflicting but rather, ultimately, compatible. This theory became the basis for hierarchical theories that followed Spearman and Thurstone.

Early Hierarchical Theories: Burt, Vernon, Cattell

A number of early hierarchical theories of intelligence were proposed. Burt (1940) suggested a five-level hierarchical model. At the top of the hierarchy is what Burt referred to as "the human mind." At the second level of the hierarchy, the "relations level," are *g* (the general factor) and a practical factor. At the third level of the hierarchy are associations, at the fourth level is perception, and at the fifth level is sensation.

More sophisticated than Burt's model is the hierarchical model that was suggested by Vernon (1950). At the top level of the hierarchy is *g*. At the second level of the hierarchy are two major group factors, called *v:ed* and *k:m*. Here, *v:ed* refers to verbal-educational abilities, of the kinds assessed by conventional tests of academic skills. In contrast, *k:m* refers to spatial-mechanical (kinesthetic) abilities. The third level of the hierarchy contains minor group factors, and the fourth level contains specific factors.

Raymond B. Cattell (1971) suggested two major abilities, crystallized and fluid. Crystallized ability, which is not dissimilar to Vernon's *v:ed*, are measured by tests such as vocabulary, reading comprehension, and general information. Fluid abilities are used to cope with novel kinds of problems and tasks, and are measured by tests such as number series, abstract analogies, and matrix problems. Gustafsson (1984) has suggested that *g* (general intelligence) and G*f* (fluid intelligence) are pretty much the same thing – that is, that there is no real difference between general intelligence and fluid abilities.

Arthur Jensen (1970) proposed that there are two broad classes of abilities. According to Jensen, Level I abilities involve simple encoding, storage, and retrieval of sensory inputs. These abilities are used in rote-learning tasks, for example, memorizing a string of unrelated words. Level II abilities are used in reasoning tasks and appear to be similar to what Spearman referred to as *g*.

Guilford and the Structure-of-Intellect Model

The Structure of Intellect (SI) theory of J. P. Guilford (1956, 1967; Guilford & Hoepfner, 1971) was one of the most ambitious theories of intelligence ever proposed. It was also one of the least successful, as will be explained.

Guilford originally posited 120 distinct abilities (increased to 150 by Guilford, 1982, and later to 180), organized along three dimensions. The dimensions in Guilford's model are operations, products, and contents. In all, there are five operations, six products, and four contents. Because these dimensions are crossed with each other, there are a total of 5 x 6 x 4 = 120 abilities in an earlier version of the theory.

Operations are of five different types, namely: cognition, memory, divergent production, convergent production, and evaluation. The six types of products in Guilford's model are units, classes, relations, systems, transformations, and implications. And the four kinds of contents are figural, symbolic, semantic, and behavioral.

The set of abilities is commonly represented by a cube. The three dimensions of the cube are content, product, and operation, and within the cube are little cubes corresponding to each of the 120 ways in which the contents, products, and operations can be crossed. Guilford devised several tests for each of the constituent cubes. For example, a matrix problem would measure cognition of figure relations.

Guilford's theory is viewed as a relic of the past. Horn (1967) showed that random data would provide as good a fit to Guilford's theory as did real data. Horn and Knapp (1973) showed that when the form of factorial rotation that Guilford used, Procrustean rotation, was applied to randomly generated theories used on Guilford's data, the fit to the model was just as good as for Guilford's actual theory. So that was pretty much the end of Guilford's theory. It survived as long as it did only because it capitalized on chance through its method of optimally rotating factors into congruence with the pre-specified theory.

Guttman and the Radex Model

One of the more interesting models of intelligence was proposed by Louis Guttman (1954). The model is what Guttman called a radex, or radial representation of complexity. The radex uses polar coordinates rather than Cartesian ones. The radex involves two parts.

The first part is what Guttman referred to as a simplex. If one imagines a circle, the simplex signifies the distance of a given point (representing an ability) from the center of the circle. The more proximal an ability is to the center of the circle, the more central it is to human intelligence. From this standpoint, general ability, or *g*, could be seen as being at the center of the circle; the more peripheral abilities such as perceptual speed would be closer to the periphery of the circle. A simplex relation is one in which later elements in a sequence completely contain earlier elements in the sequence. Thus, the abilities nearer to the periphery of the circle are completely contained by those nearer the center.

The second part of the radex is referred to as a circumplex. A circumplex refers to the angular orientation of a given ability with respect to the circle. Therefore, abilities are seen as being arranged around the circle with those abilities that are more highly related (correlated) closer to each other in the circle. Thus, radex functions through a system of polar coordinates. Snow, Kyllonen, and Marshalek (1984) used nonmetric multidimensional scaling on Thurstone's (1938; Thurstone &

Thurstone, 1941) data in order to demonstrate that the Thurstonian primary mental abilities could actually be mapped into a radex.

Modern Hierarchical Theories

The end of the twentieth century saw the advent of two modern hierarchical theories. The one that has gotten the most attention was proposed by John B. Carroll (1993). Carroll proposed a model with general intelligence at the top (stratum III), major group factors at the middle level (stratum II), and minor group factors at the lowest level (stratum I). This three-stratum model is the result of Carroll's reanalysis of many of the previous factor analyses in the literature. At the important middle level (stratum II) are fluid intelligence (G*f*), crystallized intelligence (G*c*), general memory and learning (G*y*), broad visual perception (G*v*), broad auditory perception (G*u*), broad retrieval ability (G*r*), broad cognitive speediness (G*s*), and processing speed (G*t*).

Johnson and Bouchard (2005) proposed a g-VPR model, where they propose that, in addition to *g*, there are three main factors of intelligence: verbal, perceptual, and rotation. This rotation factor is more related to Vernon's (1950) *k:m* (kinesthetic-mechanical) ability perhaps than to any of the others. Johnson and Bouchard have separated out mental rotation from other aspects of perceptual processing. They claim that their model provides a better fit to empirical data than any of the other models.

Conclusions

By the end of the twentieth century, traditional theories of intelligence had in some respects come a long way, and in other ways not come so far at all. Most theories have been modifications of Spearman's (1927) original theory of general intelligence. In all that time, few psychometrically oriented theorists moved much from the theory of *g*. The main source of progress is that theorists came to see Spearman's and Thurstone's theories as complementary rather than as opposed. Hierarchical models basically expanded on the theories of Spearman and Thurstone to explain how the two kinds of theories might be integrated.

Testing also has, in some respects, come far and in other ways hardly at all. The tests used today are very different from Galton's original tests, but little different from Binet's. Certainly the cosmetic appearance has changed, but the fundamental item types have changed little. Perhaps Binet just hit on the best way to measure intelligence. Or perhaps test constructors became entrenched, and simply failed to move beyond where Binet brought them so long ago. Only the future will tell.

References

Binet, A., & Simon, T. (1916a). *The development of intelligence in children* (trans. E. S. Kite). Baltimore, MD: Williams & Wilkins.

Binet, A., & Simon, T. (1916b). *The intelligence of the feeble-minded* (E. S. Kite, trans.). Baltimore, MD: Williams & Wilkins.
Brown, A. L., & DeLoache, J. S. (1978). Skills, plans, and self-regulation. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 3–35). Hillsdale, NJ: Erlbaum.
Burt, C. (1940). *The factors of the mind*. London: University of London Press.
Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
Cattell, J. M. (1890). Mental tests and measurements. *Mind*, *15*, 373–380.
Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.
Flavell, J. H. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills* (pp. 35–60). New York: Academic Press.
Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3–33). Hillsdale, NJ: Erlbaum.
Galton, F. (1883). *Inquiry into human faculty and its development*. London: Macmillan.
Goddard, H. H. (1912). *The Kallikak family: A study in the heredity of feeble-mindedness*. New York: Macmillan.
Goddard, H. H. (1914). *Feeble-mindedness: Its causes and consequences*. New York: Macmillan.
Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, *53*, 267–293.
Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
Guilford, J. P. (1982). Cognitive psychology's ambiguities: Some suggested remedies. *Psychological Review*, *89*, 48–59.
Guilford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence*. New York: McGraw-Hill.
Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*, 179–203.
Guttman, L. (1954). A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258–348). New York: Free Press.
Holzinger, K. J. (1938). Relationships between three multiple orthogonal factors and four bifactors. *Journal of Educational Psychology*, *29*, 513–519.
Horn, J. L. (1967). On subjectivity in factor analysis. *Educational and Psychological Measurement*, *127*, 811–820.
Horn, J. L., & Knapp, J. R. (1973). On the subjective character of the empirical base of Guilford's structure-of-intellect model. *Psychological Bulletin*, *80*, 33–43.
Hunt, E. B., Frost, N., & Lunneborg, C. (1973). Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), *The psychology of learning and motivation* (vol. 7, pp. 87–122). New York: Academic Press.
Jensen, A. R. (1970). Hierarchical theories of mental ability. In W. B. Dockrell (Ed.), *On intelligence* (pp. 119–190). Toronto, ON: Ontario Institutes for the Study of Education.
Johnson, W., & Bouchard, T., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.
Kahneman, D. (2013). *Thinking, fast and slow*. New York: Farrar, Straus, Giroux.
Schneider, W., & Shiffrin, R. W. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: 2. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol. 2, pp. 47–103). Hillsdale, NJ: Erlbaum.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition*. London: Macmillan.

Spearman, C. (1927). *The abilities of man*. New York: Macmillan.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1984a). Toward a triarchic theory of human intelligence. *Behavioral and Brain Sciences*, *7*, 269–287.

Sternberg, R. J. (1984b). What should intelligence tests test? Implications of a triarchic theory of intelligence for intelligence testing. *Educational Researcher*, *13*, 5–15.

Sternberg, R. J. (Ed.) (1985). *Human abilities: An information-processing approach*. San Francisco: Freeman.

Sternberg, R. J. (1990). *Metaphors of mind*. New York: Cambridge University Press.

Sternberg, R. J. (2001). Metacognition, abilities, and developing expertise: What makes an expert student? In H. J. Hartman (Ed.), *Metacognition in learning and instruction: Theory, research and practice* (pp. 247–260). Boston: Kluwer Academic Publishers.

Sternberg, R. J. (2003). WICS: A model for leadership in organizations. *Academy of Management Learning and Education*, *2*, 386–401.

Sternberg, R. J., & Davidson, J. E. (1982). The mind of the puzzler. *Psychology Today*, *16*, 37–44.

Sternberg, R. J., & Davidson, J. E. (1983). Insight in the gifted. *Educational Psychologist*, *18*, 51–57.

Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.

Terman, L. M. (1925). *Genetic studies of genius: Mental and physical traits of a thousand gifted children*, vol. 1. Stanford: Stanford University Press.

Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence*. Boston: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1973). *Stanford-Binet intelligence scale: Manual for the third revision*. Boston: Houghton Mifflin.

Terman, L. M., & Oden, M. H. (1959). *Genetic studies of genius*, vol. 4, *The gifted group at midlife*. Stanford: Stanford University Press.

Thomson, G. H. (1939). *The factorial analysis of human ability*. London: University of London Press.

Thorndike, E. L., Bregman, E. D., Cobb, M. V., & Woodyard, E. I. (1926). *The measurement of intelligence*. New York: Teachers College Press.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Thurstone, L. L., & Thurstone, T. G. (1941). *Factorial studies of intelligence*. Chicago: University of Chicago Press.

Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1944). *The measurement of adult intelligence* (3rd ed.). Baltimore: Williams & Wilkins.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (5th ed.). Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children, Revised*. New York: The Psychological Corporation.

Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review, Monograph Supplement*, *3*(6).

4 An Alternative View on the Measurement of Intelligence and Its History

Paul De Boeck, Laurence Robert Gore, Trinidad González, and Ernesto San Martín

The history of intelligence measurement has evolved in important ways since the first, Victorian-era efforts, without, however, arriving at a consensus or a generally satisfying approach. One of the earliest initiatives was from James McKeen Cattell (1890), who borrowed his approach from Sir Francis Galton (1883). Cattell proposed simple tasks, called "mental tests," to measure intelligence, for example, tasks based on the perception of size and perceptual discrimination. Later approaches are based on more complex cognitive tasks such as quantitative reasoning in the Stanford-Binet test (Roid, 2003) and word and concept similarities in the Wechsler Adult Intelligence Scale (WAIS) (Wechsler, 2008), and these are two of the most popular individual intelligence tests as well as two of the most enduring in common use. However, there is no consensus that cognitive tasks are sufficient to capture the most important aspects of intelligence (Sternberg, 1997).

While the history of intelligence measurement may be well known, we will throw a somewhat different light on it, based on two types of perspectives, the first of which is a rather abstract, philosophical approach in an attempt to define contours of what has been done in intelligence measurement and to anticipate limitations on what might be done.

1. One perspective is derived from Aristotle and concerns the way a thing can be identified. It is only through the *identification* of intelligence that its *measurement* becomes possible, though identification does not necessarily imply a definition. Just as different methods exist to identify a person, such as from a person's face, voice, fingerprint, handwriting, or DNA, there are different approaches to identify intelligence (e.g., through its manifestations or results).
2. The other perspective is that all measurement is based on comparison and that different *bases of comparison* are possible. Intelligence has traditionally found a home in individual differences research – enabling comparisons across people – but other comparisons are possible, such as comparing a person's intelligence measure with the problems persons with a similar score can solve.

From our discussion of the history of intelligence measurement, we conclude that a variety of meaningful ways to measure intelligence exists, that there is no consensus on a best measure, and that much depends on one's perspectives. Our conclusion is that intelligence is many things, that it neither has a unitary nor a multiple essence, but that, instead, each time it is measured it is a somewhat different, though, one hopes, a *not too much different*, construction on the part of the investigator. We believe that there are not different approaches to the same intelligence but rather that each approach corresponds with a somewhat different construction (notion) of intelligence. Still, at the same time the corresponding measures – those developed by different investigators with different notions – are useful for statistical prediction and for the explanation of other variables. Although there is no essence to intelligence, there must be some form of overlap among different notions for people to understand each other when they use the term, and for researchers to read each other's publications and to organize scientific meetings.

Metaphorically speaking, intelligence is like the Italian dish "saltimbocca." If you google the term or consult a cookbook, you will find a variety of definitions, and if you order the dish in Italian restaurants around the world, it turns out, amazingly, that there is not one single common ingredient in the dishes served as "saltimbocca." Still, there are experts who claim they know and other experts who admit they do not know for certain, but they believe nevertheless there must be a definition (and they may still be working on one). Unfortunately – or fortunately, perhaps, for gourmands who appreciate variety – definitions cannot be enforced, and the reality of served saltimbocca contradicts any beliefs about its essence. Of course, saltimbocca is food, but the equivalent statement for intelligence is that intelligence is a potential of the mind. We do realize that this nonessentialist view on intelligence is only an interpretation, but it may help the reader start our chapter with that interpretation clearly in mind.

Ways of Identification: Aristotle's Causae

The first perspective we take on the history of intelligence measurement relates to the way psychologists have chosen to identify intelligence. In his theory of *causae*, Aristotle attempted to enumerate a complete list of possible ways of identifying a thing – *any* thing – that might form the object of interest. Prior philosophers used the term "aiton," and neither term refers to causation as we think of it. Instead, both terms referred to all those substances, events, and forces believed to be responsible for an object, phenomenon, or observable change in the world. Aristotle's notions can be interpreted as ways to identify something and distinguish it from similar things that it is not. Suppose you were trying to explain to a friend who a person you are both familiar with is, and your friend does not recognize her name. To identify a person, one can refer to (1) the person's appearance (e.g., "she wears glasses"), (2) what kind of person they are (e.g., "she is smart"), (3) the person's origin (where they came from) (e.g., "she was X's student"), (4) what the person's intentions or goals are and what they have produced or are producing or what their contributions are (e.g., "she is the author of the novel I am reading"). Similarly, there exist a variety of ways to identify intelligence, with possible consequences for how intelligence is measured.

An identification criterion does not necessarily imply a definition (De Boeck, 2013) (although it is a common belief that a definition is crucial, e.g., Hunt & Jaeggi, 2013), just as you can know who a person is without knowing their name and unique identification number, such as a social security number in the United States. Perhaps unfortunately, efforts to define intelligence have produced a huge diversity of definitions (Thorndike et al., 1921; Legg & Hutter, 2007; Neisser et al., 1996; Sternberg & Detterman, 1986). A very popular definition is from Gottfredson (1997, p. 13):

> Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings – "catching on," "making sense" of things, or "figuring out" what to do.

It is a composite and complex definition. Another, much simpler but still composite definition is from Plomin and von Stumm (2018, p. 148): "the ability to learn, reason, and solve problems."

We believe that a definition is just one way to identify, a way that corresponds best with identifying someone indicating what kind of person they are. Ironically, the diversity of definitions is in contrast with the similarity of the most popular methods to measure intelligence (the IQ), which illustrates that a definition is not necessary to measure. To illustrate the difference between a scientific definition and identification, think of warmth. Warmth is kinetic energy of molecules, but we do not need to know that to infer that it is hot outside when most people wear short sleeves. Another example from psychology is psychopathology. There is no single, commonly accepted definition that unifies everything from schizophrenia through depression to caffeine withdrawal. Nevertheless, people can come to reasonable agreement most of the time as to whether a syndrome represents pathology. The fact that some syndromes are still debatable as to whether they represent forms of psychopathology or not along with the fact that some things originally on the official lists of pathological phenomena (such as homosexuality) have been removed from those lists testify to the incompleteness and imperfection of the definition of psychopathology. Nevertheless, progress has been made even without a complete definition. Likewise, progress has been made in understanding intelligence without a complete and universally accepted definition.

Different authors have explained Aristotle's notions differently, depending on the context of the discussion and the time in history. We take the freedom in this chapter to interpret the notions so that they can be used for our discussion. The Latin terms are from Thomas Aquinas (Aquinas, 1961).

1. *Causa formalis* is the form or shape something takes – how it appears. For intelligence, the *causa formalis* is how intelligence appears – in what kind of behaviors and performances it shows. For example, we may attribute intelligence to someone because they successfully solve cognitive problems or successfully deal with practical issues (Sternberg et al., 2000), social issues (Barnes & Sternberg, 1989; Goleman, 2006), or emotional issues (Goleman, 1995; Mayer & Salovey, 1993). What we mean by *causa formalis* has often been labeled the *domain of intelligence*, and we will adopt that usage here.

2. *Causa materialis* is the material something is made of. For intelligence, these are the *components of intelligence* underlying its appearance. The components are factors and cognitive processes. Factors are derived from factor analysis, a correlation-based technique to find the basic differences between people, called factors or dimensions, that create variation in a set of test item responses and test scores. Processes are the mental activities when intelligence is at work. As an example, numerical ability and verbal ability are two of the seven primary abilities (factors) Thurstone (1938) found in his analysis of a set of tests. As an example of cognitive processes, Sternberg (1977a) has developed a model with processes such as, for example, encoding (framing the question), inference (of a relationship), and justification (of a tentatively chosen response). They are three of the processes people use to solve analogy problems. They are cognitive processes within a person's mind and not (necessarily) dimensions of differences between persons. In general, factors are derived from a statistical analysis of data. By and large, factors have been developed from a tradition of analyzing test data collected to measure intelligence, while cognitive processes have been developed by cognitive psychologists who attempt to understand how cognition functions. There is no necessary reason why factor analysis should provide the same list of *causae materiales* as analyses of cognitive processes do.
3. *Causa efficiens* is what leads to something and can be interpreted as the origin, source, or basis of something. Although factors and cognitive processes can be interpreted as the origin, in our discussion we will consider only external elements as *causa efficiens*. For example, genes, brain activation, and brain connectivity are external to the actual problem-solving activities where intelligence shows in the sense that they are not derived from those activities and that they require separate observations regarding aspects that can be interpreted as the distal or proximal basis of the activities. Therefore, they are candidates to be a *causa efficiens*. In fact, speed of processing and the efficiency of information accumulation can be interpreted as a *causa efficiens*, on the condition that the speed is not observed in or derived from the intelligence tests under consideration. Being inherent versus extrinsic to the intelligence domain that is typically considered will be used as the criterion to differentiate between *causa materialis* (inherent) and *causa efficiens* (extrinsic). Something cannot be the efficient cause of itself.

Some readers may be surprised that genes and brain activation are not treated here as *causa materialis*. Two examples may help. Einstein undoubtedly used his brain to create highly intelligent approaches to problems in physics, but his brain was not what his intelligent thinking consisted of. Similarly, if an alien built a spaceship and landed on Earth with nothing remotely resembling a human brain or human carbon-based genetic code, we would not deny that the alien is intelligent. Our common notions of intelligence, including those of artificial intelligence, suggest that carbon-based brains and genes are external to intelligent problem-solving. The really central, or material, aspect of intelligence has to do with the processes needed to resolve problems.

4. *Causa finalis* is the purpose or result of something. For example, we would identify a scientist as the inventor of something, or we would describe the president of a country through her legacy. For intelligence, one historically and still important *causa finalis* concerns successful academic achievements. A *causa finalis* is also implied by notions such as "successful intelligence" (Sternberg, 1997) and in Thurstone's (1924) definition of intelligence as "the ability to get along well in society . . . " This view implies that the things measured in intelligence assessment need to lead to or reliably predict the "finis" or the end one has in mind.

To summarize with just one, arbitrarily chosen, example of each of the four *causae*, intelligence shows in successful performances on cognitive tasks, it consists among other things of inference processes, it is partly caused by the speed of neural connections, and it helps to do well in school. Each of these four qualifications helps to some extent to identify intelligence. The notion of intelligence as explained in definitions often is an abstract formulation of the *causa materialis*, and thus to what constitutes intelligence, although sometimes other *causae* are also mentioned. Aristotle proposed that no analysis of a phenomenon is complete until all four *causae* have been examined.

Ways to Compare

The second perspective from which to discuss the history of intelligence measurement is the necessary comparison all measurement requires. Measurement is always interpreted in terms of a comparison. For example, the meter has been defined as the length of the path of light traveling in a vacuum for 1/299,792,458 of a second (Giacomo, 1984), whereas one of the earliest definitions of a meter was the French National Assembly's definition in 1791 as "one ten-millionth of the length of a quadrant along the Earth's meridian through Paris, that is the distance from the equator to the north pole along that quadrant" (Agnoli & D'Agostini, 2005, p. 2). The process of setting up a basic definition of a physical unit is known as standardization. For psychological variables, such as intelligence, there is not a well-defined standard unit of measurement to compare with. What would represent a basic unit of thought or problem resolution? There are two basic approaches to comparison that substitute for the kind of standardization available in the physical sciences: person comparison and task comparison.

One can compare a person's score on an intelligence test with the scores of other persons, which is what the modern IQ score does. For example, an IQ of 120 is among the 10 percent highest IQs in the population. However, an IQ among the 10 percent highest IQs does not tell us what a person with such an IQ is able to do. Alternatively, one can interpret a test score in terms of the tasks the person with that score can successfully perform. A special type of task comparison occurs when the comparison is not made with tasks from the test but with external tasks such as those required for a job. Following the old way of calculating the IQ developed by

Lewis M. Terman, sets of tasks were defined for each age group such that most subjects (e.g., two-thirds) from the age group in question succeed in the tasks. A subject is assigned a mental age depending on the tasks they can perform correctly, and the IQ is determined as the ratio of mental age divided by the chronological age times 100. For example, if an eight-year-old child succeeds in the tasks for age group nine but not in those for age group ten, the child's mental age would be nine and IQ would be 9/8 times 100, which is 112.5. In that way, individual performances are compared to a set of standard tasks believed to be achievable by most people in an standard reference age group.

Early Stages in the Measurement of Intelligence

The earliest initiatives trace back to Galton's work (e.g., Galton, 1883) and his anthropometric project to improve the quality of the human race, also known as eugenics. Galton distinguished among three types of qualities: bodily qualities, energy, and sensitivity. He considered energy as the most important from an eugenic point of view ("In any scheme of eugenics, energy is the most important quality to favor," p. 19), and in his view energy was also related to intelligence, as was sensitivity: "The only information that reaches us concerning outward events appears to pass through the avenue of our senses and the more perceptive the senses are of difference, the larger is the field upon which our judgment and intelligence can act. Sensation mounts through a series of grades of 'just perceptible differences'" (p. 19). The term "just perceptible difference" is a term from psychophysics to indicate how well people can perceptually differentiate between two stimuli, for example, two different lengths. Galton was of the belief that discriminative power was a preliminary condition to be intelligent, for the reasons explained in the above quote. Therefore, he suggested discrimination of weights and audibility of shrill notes as tests to measure sensitivity and thus intelligence.

Galton's work was continued by James McKeen Cattell (1890), the first psychology professor in the United States and long-time editor of the now highly prestigious journal *Science*. Cattell had studied with Wilhelm Wundt and had worked at the University of Cambridge in the UK with Galton. Cattell's motivation was, in the first place, to establish psychology as a quantitative science. The most prestigious quantitative results in the psychology of his time came from psychophysics: the measurement of sensations (e.g., perceived loudness) and the mathematical laws specifying how the magnitude of sensations a person experiences depends on the magnitude of the corresponding physical attribute (e.g., sound pressure). The just noticeable difference (jnd) between two physical attribute values (e.g., two sound pressures, two lengths) was used as the unit of sensation measurement, and it turned out that measured sensations show a clear logarithmic relationship with the corresponding physical attribute (Fechner, 1860), a universal or near-universal fact of human experience that can still be demonstrated easily today. Psychology at that time was conceived as the empirical science of immediate experiences, for example, sensations of the same objects which are investigated in the natural sciences:

> The point of view of natural science may, accordingly, be designated as that of *mediate experience*, since it is possible only after abstracting from the subjective factor present in all actual experience; the point of view of psychology, on the other hand, may be designated as that of *immediate experience*, since it purposely does away with this abstraction and all its consequences. (Wundt, 1902, p. 3)

The combination of Wundt's definition of psychology in terms of the achievements of psychophysics and Galton's view on the importance of the senses for judgment and intelligence inspired Cattell's design of mental tests. Accuracy and time of differentiation were among his measures of intelligence. Although one may think of discriminatory power as a preliminary condition (as Galton did) and thus as a *causa efficiens*, James McKeen Cattell was in the first place a pragmatist and an instrumentalist; he selected mental tests from experimental tasks of his time based on their ability to produce a quantitative measure of performance. He did not give a definition of intelligence, but he was evidently of the belief that there is a unitary mental power, a concept expressed in the quantitative experimental tasks of his time and thus in the mental tests he proposed. In sum, his way of identifying intelligence was of the *causa formalis* type: "if a performance can be quantified it must reflect intelligence" (our interpretative summary). His approach is now considered simplistic, but it was inspired by his scientific norms and the early stage of experimental psychology as conceived by Wundt. It is worth reflecting on the limits imposed on Cattell by his time. Wundt sought to understand how the nervous system translates basic physical magnitudes into sensory experiences. Complex experiences, such as the inference of a person's motives based on a pattern of behavior, or the formulation of an effective strategy in a politically tense situation, lay entirely outside the scope of what was being measured by early psychologists such as Wundt. As a result, Cattell's early efforts to quantify cognitive ability replicated the limited nature of available performances that Wundt's choices dictated. It is interesting to wonder how the course of intelligence testing might have changed if early psychologists had begun with the study of social perception rather than physical perception.

The idea that sensory discriminative power is an important factor underlying intelligence may seem simplistic, but it also shows in more recent work on intelligence. For example, Ratcliff, Thapar, and McKoon (2010) relate intelligence to efficiency of information accumulation in simple binary perceptual decisions, and Melnick and colleagues (2013) link sensory discrimination to intelligence through the process of sensory suppression of irrelevant aspects. In the years since Cattell carried out his research, it has become practical to measure responses more accurately (in terms of response time) and to present stimuli more consistently by programming cognitive experiments on computers. It is conceivable that Cattell might have come closer to succeeding in establishing elementary cognitive tasks (the term of art in psychology for simple decisions) as the basis of intelligent behavior if he had access to the tools investigators can now use.

Wissler (1901), who studied with Cattell, took this work one step further and investigated individual differences, using correlations in line with the approaches of Karl Pearson and Francis Galton in the UK. Academic achievement was already considered to be an important variable (a *causa finalis*) with which a test should correlate to have credibility as a measure of intelligence; Wissler also calculated correlations between his measurements and academic achievement. Unfortunately, Wissler's correlations were too low to support any claim of credibility for Cattell's measures as a basis for assessing intellectual potential.

Binet and a Switch to Another Domain

The next step in the history of the measurement of intelligence was inspired by the concern that intelligence be relevant for academic performance, something Cattell's tests did not seem to be able to accomplish. Instead of the simple tasks Cattell borrowed from experimental psychophysics, Alfred Binet based measurement on more complex cognitive problems. He first explained that all psychological phenomena, including sensation and perception, imply intelligence (he is kind enough to expand beyond rather than contradicting Cattell's formulation). He then posed the rhetorical question "Devrons-nous faire intervenir dans nos examens la mesure de la sensation à l'exemple des psycho-physiciens. Devenons-nous en tests mettre toute la psychologie?" (Binet & Simon, 1905, p. 196) (Do we need to include sensation measures following the psychophysicists? Do we need to include the whole of psychology in our tests?) It was a rhetorical question with a suggested negative response. After this diplomatic and rhetorical reasoning, he concluded that judgment is the crucial element in intelligence. However, right before he singled out "judgment" as the concept to focus on, he asserted that judgment is just one among many concepts relevant to intelligence. Others are common sense, practical sense, initiative, adaptation, understanding, and reasoning (Binet & Simon, 1905, pp. 196–197). He did not differentiate between those terms. One of his reasons for the change in how intelligence should be measured and for refuting psychophysics as an approach to testing intelligence was that people such as Helen Keller and Laura Bridgman were very intelligent and would nevertheless fail the psychophysical tests because of their sensory handicaps. These two persons were both manifestly intelligent and well-educated American women who were also blind and deaf. Binet and Simon proposed for inclusion in their test those tasks requiring judgment, a very broad concept, given that it was not differentiated from common sense, practical sense, etc.

In practice, the problems included in what came to be known as the Binet-Simon test were similar to the kinds of problems with which students are confronted in school, except that solving the test problems did not require the knowledge acquired in school and the training given in schools. Although Binet wanted to employ tasks from a broad and diverse domain, he also wanted to reduce any effects of the amount and quality of prior schooling experience. This concern was inspired by the practical need (in the Paris public school system, which commissioned the work of Binet and Simon) to identify students who required special education. A child previously taught in low-quality schools might perform poorly on an achievement test and might need better schooling to remediate past educational deficits but not necessarily special education to compensate for a learning disability. The difference between the measurement of intelligence and of intelligence plus schooling is still considered by some as the basis for the differentiation between intelligence tests and achievement tests (Cianciolo & Sternberg, 2004), respectively. Although the presentation of the first Binet-Simon test looks as if it is based on the central component of judgment, Binet was instead inspired by the domain of problems that are similar to those presented in academic education, but without requiring knowledge specifically acquired in school. The context for Binet was education whereas the context for Cattell was early experimental psychology. The attempt to separate the effects of past educational enrichment from current intellectual potential has eluded the best efforts of scholars in intelligence for over a century since Binet and Simon initiated it. The popular understanding of intelligence still encodes Binet's idea of a generic cognitive potential separable from past educational experiences.

Another interesting feature of Binet's approach was that his first series of thirty problems developed in 1905 in collaboration with Theodore Simon were a set of problems with increasing difficulty so that the derived score would basically express the level of difficulty a child was able to achieve. This approach allows for but does not require a comparison between persons. In their first article (Binet & Simon, 1905), on their first test, Binet and Simon used the term "intelligence level." In a later version of the test, the 1908 version (Binet & Simon, 1908), the items were grouped into age clusters, so that one could determine what the "mental age" of a child was. In 1912, William Stern launched the idea of an intelligence quotient as the division of mental age by chronological age (Stern, 1912), which then yields an IQ (intelligence quotient) if multiplied by 100.

The Binet-Simon tests were introduced into the United States by Lewis M. Terman (1916) as the Stanford-Binet test, but in contrast with Binet's inspiration to find the best method of instruction for every child, workers in the United States deployed intelligence testing for eugenic purposes, just as Galton foreshadowed (Black, 2003). The Stanford-Binet, currently in its fifth edition, remains a popular intelligence test, particularly in the assessment of giftedness in children and general assessment of cognitive performance of preschoolers (Roid & Tippin, 2009). Later, two other tests were partly inspired by Binet's test but with a totally different purpose: the Army Alpha and Army Beta, the first group-administered intelligence tests, developed under the direction of Robert Yerkes at the end of World War I to test recruits before they were sent to Europe (Yerkes, 1921). Testing had to be fast and efficient, given the urgency of the war, which explains why the tests were administered in groups. The alpha test consists, for example, of verbal and numeric problems and was presented in written English, whereas the beta test was a nonverbal test with problems that could be understood and solved without reading or writing English, so that illiterate and foreign language–speaking recruits could also be tested. A rather large variety of questions was included, inspired by the type of tasks Binet was using, in line with Binet's *causa formalis* approach. The test score was expressed in terms of the job or position it was predicted one could fulfill in the army. After the war ended, scores were tabulated across groups defined by race and ethnicity and used to advocate for immigration controls in order to protect the quality of the intellectual gene pool in the United States (Black, 2003).

Spearman's Search for Invariance and the Foundation of Factor Analysis–Based Measurement

Around the same time early in the twentieth century, Charles Spearman started his own research (Spearman, 1904a): empirical studies to test his theory of general intelligence using an early form of factor analysis. Spearman's work has become the foundation of psychometrics (Spearman, 1904a, 1904b), for example, in that Spearman was the first to use the notion of a factor or latent variables. A latent variable is one that cannot be directly measured or observed but which explains the covariations in observed scores.

Spearman today is best known for his *g* theory, which says that all measurements of intelligence measure the same general intelligence. This view was first described in his 1904 article on general intelligence, without using the lowercase *g* (Spearman, 1904a). Later he started using the lowercase letter *g* of the term "general" because words come with connotations, whereas *g* is formally speaking nothing more than the only source for correlations between different intelligence tests (Spearman, 1927). His *g* theory fits the eugenic belief that there are general qualitative genetic differences between people. Here, we will highlight a few of the more intellectual and less well known motivations behind Spearman's work.

Spearman, like Cattell, had been a student of Wilhelm Wundt, and he was also heavily influenced by Galton. The early scholars of intelligence had a background in both experimental psychology and measurement because measurement was, at that time, a primary interest of experimental psychologists. To earn respect, psychology had to describe human behavior and perception in a precise, quantitative form. Spearman's work on intelligence extended the quantification to individual differences and how they are correlated. After Spearman, the experimental tradition of precise measurement diverged from the correlational approach based on the study of individual differences. It took many decades for the two traditions, one correlational and one experimental, to be again extensively combined in the study of intelligence, specifically in the early work of Robert J. Sternberg (1977a), a topic we will return to later in this chapter.

Spearman's (1904a) early work on intelligence was an interesting and wonderfully up-to-date mix of psychophysics and psychometrics. He was able to establish a substantial correlation of psychophysical discrimination performances (discrimination of pitch, luminosity, and weights) with academic performance, intelligence, and common sense. For intelligence and common sense, he relied on judgment by teachers (intelligence) and peers (common sense). The correlations between the sensory discrimination tasks seemed strong evidence for a general discrimination notion and, more importantly, a combination of these task performances were extremely highly correlated with a general factor underlying the intelligence and common sense ratings and academic performances. This was for Spearman the proof that intelligence could be measured in an objective way using the most prestigious experimental and mostly psychophysical tasks of his time, i.e., discrimination tasks, and that there was one general factor underlying everything.

The reasons why Spearman (1904a) was more successful than Wissler (1901) in showing a strong relationship between simple discriminatory tasks and between such tasks with academic achievement were his carefulness in conducting the experiments and more importantly his formulae to correct for sources of unreliability and confounding and his factor-analytic approach. Whereas Spearman's (1904b) formulae have become the basis for the psychometric literature on reliability and validity, the discrimination tasks have mostly been given up as measures of intelligence. After being almost completely abandoned, elementary cognitive tasks involving simple decisions or perceptual judgments are experiencing a comeback in cognitive experimental approaches to measure cognitive processing efficiency, which is seen as related to intelligence (Ratcliff et al., 2010; Schmiedek et al., 2007).

Spearman was inspired by Wundt in his search for ways to achieve measurement invariance in psychology, just as they exist in physics. For example, "heat" is kinetic energy of molecules wherever and whenever it is observed. In Wundt's and Spearman's views, for a concept (intelligence, in Spearman's case) to be valuable, one had to find invariance, an unchanging identity, in the same sense as kinetic energy is for warmth. Whatever the measure is, if intelligence exists, it should show in that measure. Spearman chose correlation as a method to find invariance, which he called "identity" and "functional uniformity." After corrections for unreliability and types of confounding, the correlation between all the measures should be perfect if they measure the same construct. For psychology to be a real science, in Spearman's view, the most important work was to uncover the invariance of a concept, in an analogous way as physics was able to achieve for warmth. For Spearman it happened to be intelligence, which was also called "mental energy" at that time. In his own view, his search for invariance was successful. After making the corrections he considered necessary, all measures of intelligence, from teacher ratings of common sense to academic task performance (such as solving math problems) to psychophysical discriminations were highly or even perfectly correlated.

Although *g* theory is controversial, there was a clearly intellectual motivation for Spearman's endeavors in search of invariance, inspired by Wundt and physics as a model for psychology, and Spearman's research to validate the theory has bought us important formulae (as just mentioned) that are still being used, and factor analysis, a method that is still highly popular to find the basic dimensions along which people differ. Invariance as a principle has survived in the field of measurement, for example, in the ideal of measurement invariance, which means that a test must measure the same concept in all groups of people where it is used to measure.

Spearman was not satisfied with just quantitative evidence for *g*, void of any meaning. He also published a theory about the processes underlying *g*. He conjectured that these processes were the "apprehension of experience," "eduction of relations," and "eduction of correlates" (Spearman, 1923), terms used to indicate understanding the problem, inference, and application of relationships. For example, if one is asked the following question "grandfather (A) is to grandson (B) as grandmother (C) is to . . . (D)?" then one must be aware of the constituents of the question and understand what is being asked (apprehension of experience), then "educt" the generational relationship between A and B and "educt" "granddaughter" as a correlate of C by applying the generational relationship. Raven, a student of Spearman, later developed the Progressive Matrices test (Raven, 1938), a well-known measure of intelligence still in research and clinical use, in which three-by-three matrices were presented with figures in eight of the nine cells with the ninth cell left blank. To fill the ninth cell, one needs to find the relationships between the other cells and apply these relationships to select the correct response. Almost four decades later, Sternberg's first topic of intelligence research (modeling analogy-solving processes) was based on Spearman's three processes (Sternberg, 1977a, 1977b): encoding (understanding the terms of the analogy), finding how A and B are related (eduction of a relation) and then applying the relationship to C to find D (eduction of a correlate).

In sum, Spearman was less a *causa formalis* thinker than Cattell and Binet. He was a *causa materialis* thinker, in two ways. Under the influence of Galton, he was looking at factors of individual differences, and under the influence of Wundt, he was looking for evidence of measurement invariance. In his view, he found the essential component of individual differences and the essential components of the corresponding processes. Raven's Progressive Matrices test is a *causa materialis* type of intelligence measurement, based on an identification of intelligence as *g* and the processes we have discussed.

Later Factorial Approaches to Intelligence

Factor analysis became a very popular approach, indeed, one of the few approaches to investigating the *causa materialis* of intelligence. Intelligence theories have been and still are heavily influenced by the result of factor analysis. Originally, factor analysis supported exploration of patterns of correlations among test scores. Later, when more theories had been developed, a more confirmatory approach was followed testing hypotheses derived from those theories. Factor analysis focuses on individual differences and the structure of the correlations (covariance) of these differences. Factor analysis shows that people who achieve high scores on one intelligence test also tend to achieve high scores on others, and people who receive low scores on one such test tend to receive low scores on others.

It is true that the components of individual differences have often been interpreted in terms of cognitive processes, but such an interpretation does not logically follow. The interpretation is necessarily a post hoc interpretation based on the assumptions that processes are directly reflected in individual differences in performances and that correlation between performances defining a factor indicates that a common process is involved. Factor analyses of intelligence are exclusively based on individual differences in performances and correlations between these performances. Factor analyses are easily mistaken as indicating more than they really show.

Two issues are at stake regarding factor-analytic theories of intelligence: (1) the kind of model for the structure of individual differences in intelligence that is used, and (2) the kinds of intelligence, assuming that intelligence has multiple facets in contrast with Spearman's single factor *g* theory. In line with the factor-analytic results, the common belief is that intelligence does have multiple facets, even though there also may be room for a notion of general intelligence.

Kinds of structure. The kinds of structure can be roughly categorized into two types: (1) nested factors and (2) correlated factors. Using a metaphor, think of provinces as nonoverlapping entities within a country and of cooperation between provinces, respectively. *Nested factors* are initiated by Holzinger and Swineford (1937) as an extension of Spearman's *g* theory (they are also called bifactor models). Holzinger had worked with Spearman and has collaborated with him on the development of nested factors. The concept of *g* is preserved, but separate clusters of tests also have their own "cluster-specific *g*-factors" so to speak. In the literature they are called group factors. They can be considered intermediate organizational units. Using the metaphor of provinces and a country, each town (each test) belongs to a province (a group factor) that in turn belongs in a country (*g*). The country is the whole domain of intelligence. In Spearman's *g* theory there is only the country (all tests measure *g*) and specific towns (the specificity of each test, called *s* by Spearman). The nested factor model has intermediate levels of organization, either just one level (like provinces) or two (counties within provinces), and in theory more levels are possible.

The nested factor model was supported by Burt (1949) based on a review of factor analysis results, and Vernon (1950) developed a well-known factor theory of intelligence with two intermediate levels between general intelligence and the specific factors. General intelligence is first subdivided into two large provinces, verbal-education ability (*v:ed*) and practical-spatial-mechanical ability (*k:m*); these are further subdivided into counties (narrower factors), and, finally, at the lowest level there is a place for towns, Spearman's specific factors. For a recent description of the nested factor view, see Kell and Lang (2017). A typical feature of the nested factors approach is that the factors within each level and between levels are independent. They each represent different nonoverlapping aspects of intelligence: pieces of the intelligence puzzle and conglomerates of pieces of the intelligence puzzle.

An alternative approach, called *correlated factors*, was initiated by Thurstone (1938). Based on his way of factor-analyzing correlations, Thurstone claimed that Spearman's theory was too simple. In his famous book on the primary mental abilities, Thurstone described the results of his factor-analytic work as yielding seven different but correlated abilities without a general factor (Thurstone, 1938). He gave up on the invariance ideal of one common intelligence and was of the belief that invariance should be sought in distinct kinds of intelligence. He followed the data, whereas Spearman tried to use data to confirm his theory, which in his view was necessary for the notion of intelligence to make sense. Starting with his belief that the data could be misleading due to sampling error and confounds, for Spearman there was one and only one intelligence, while for Thurstone there were more, similar to what have later come to be called multiple intelligences by Howard Gardner (1983), who has not relied on factor analysis or focused on specific test development as much as conceptualization. Spearman saw intelligence as one country, while Thurstone saw only provinces that worked together (i.e., correlated factors).

Thurstone's results were not a definite verdict on the absence of a general factor, because the general factor has been reintroduced to explain the correlation between the factors as found by Thurstone (e.g., Gustafsson, 1984). The positive correlation between the factors of intelligence can be explained by a common underlying factor. In terms of the metaphor we used for the nested factor approach, the correlated factors are like cooperating provinces and the general factor as the collaborative body. Unity can be defined in terms of positive relationships (correlated factors) or in terms of being subsets of the same superset (nested factors).

One can reconcile the two views, nested and correlated factors, as follows. Suppose there are three provinces and they share a common set of laws which can be interpreted as *g*. They each also have their own province-specific set of laws: A, B, and C. In the correlated factor approach the total set of laws valid in a province is $A^+ = g + A$, $B^+ = g + B$, and $C^+ = g + C$. The factors A^+, B^+, and C^+ are correlated group factors and as such they are *integrated factors*, in that the two kinds of laws are combined. On the contrary, the nested group factors A, B, and C are *deviation factors*, $A = A^+ - g$, $B = B^+ - g$, $C = C^+ - g$. This has important consequences for the interpretation of factor scores, either as deviations from the general factor score (nested factors), or as the result of an integration with the general factor (correlated factors). In practice, the two approaches are two workable perspectives on the same structure of individual differences. The observant reader will note that if $A^+ = g + A$, it follows from a little simple algebra that $A = A^+ - g$. They are two ways of looking at the same country and two ways of thinking about the individual differences in the domain of intelligence. The differences between nested factors and correlated factors are somewhat more complicated, but that would lead us too far in this chapter.

Kinds of intelligence. Starting with Thurstone's work, an interest has grown in the kinds of intelligence. We will not discuss the history of factor-analyzing intelligence, but only briefly touch on the most popular view, one that is aligned with the correlated factor approach: the Cattell-Horn-Carroll (CHC) theory (McGrew, 2009), referring to work by Cattell (1963), Horn (1968), Horn and Cattell (1966), and by Carroll (1993). The factor-analytic quest for the kinds of intelligence has been answered in terms of broad factors. They are factors that explain the correlation between narrower factors and whose correlation is in turn explained by *g*. Cattell and Horn have investigated the two most important ones, fluid and crystallized intelligence: the intelligence of pure reasoning in the context of novel problems (fluid intelligence, designated G*f*) and the intelligence of accumulated experience, knowledge, and skills (crystallized intelligence, or G*c*). In an influential publication Gustafsson (1984) describes the result of a higher-order factor analysis, with, among others, these two higher-order factors and a small number of additional ones, and with general intelligence (*g*) at the apex. An interesting result was that fluid intelligence seemed to be a pure representative of *g*, suggesting that fluid intelligence captures Spearman's invariant notion of intelligence. In 1993, Carroll (1993) published a book with a collection of factor analyses and the results of his own analyses of available data sets. He summarizes the factor-analysis results thus far in the history of factor-analyzing intelligence tests. His conclusions are that intelligence has three strata (a large collection of narrow factors, a smaller number of broad factors above the narrow factors and driving their variability, and *g*, the single unitary factor helping to explain or drive variation in all measures of intelligence) and that general intelligence is primarily defined by the ability to solve complex cognitive problems. The broad factors are fluid and crystallized intelligence, and, for example, visual intelligence and cognitive speed. It is an integrative rather than an innovative model. As far as individual differences are concerned, Carroll's three-stratum theory and McGrew's (2009) integration in the CHC theory are the state-of-the-art theories but see Johnson and Bouchard (2005) for a defense of Vernon's theory against the CHC theory. Both are *causa materialis* theories about the nature of individual differences.

However, the end result of the factor-analytic approach to ascertaining what intelligence is can only be a restatement of patterns of intercorrelation among test scores, such as "the factors of individual differences in various intelligence tests are . . . " followed by a list. Such a result does not necessarily speak to the underlying particulars of the problem-solving processes through which intelligent decisions are generated. The individual differences in scores may not reveal the processes through which cognitive ability test skills are executed. For example, having comparatively high reading skills does not tell us what the reading processes are. To understand processes underlying performance, theories must be created and tested to explain each particular kind of cognitive performance.

Still, some of the factor-analytic scholars do also have an a priori definition of intelligence, independent of the factor-analytic result. For example, Thurstone (1924) first admits that, for the measurement of intelligence, one needs to rely on the products of intelligence. In our terminology this is how intelligence appears in the performance of tasks from the domain of intelligence (*causa formalis*). Next, he speculates that intelligence implies covert trial and error before deciding on a behavior. This requires inhibition of overt trial-and-error behavior, in order to think things through before deciding on an action, and in even higher forms of intelligence trial-and-error thoughts are inhibited in favor of more abstract thinking using categories of problems and solutions. This has led to definitions of intelligence as abstract reasoning more generally. How well people can inhibit overt trial and error has never been measured. In Thurstone's work, as elsewhere, we often observe a discrepancy between a general definition of intelligence and its actual measurement, with no close connection between the two. It is often impossible to derive someone's definition of intelligence from the intelligence tests that are used and vice versa.

Because factor-analytic approaches focus on individual differences rather than cognitive processes, it is not surprising that Guilford (1956) formulated a theory that was indeed based on cognitive processes. In his effort to help the U.S. Air Force understand how best to identify and train personnel for complex duties, he developed an elaborate (some might say formidable) a priori framework for the measurement of intelligence, called the Structure of Intellect (SI) theory. The processes are defined by cognitive operations, the content on which they operate, and their product. This model provided a clear basis on which to develop *causa materialis* tests: develop a measure for each cell of the resulting (large) array. The problem was that an a priori theory about differences between tasks is not necessarily valid with respect to people, especially when tested with an individual-differences method such as factor analysis, which was used to test SI theory. Contents and products are rather easy to determine, but operations are more hypothetical. The fact that different operations can be formulated in theory does not imply that they are part of the actual processes in practice and it is far from certain that contents, products, and operations or their combinations are each (separate) sources of individual differences. Although laudable in principle, Guilford's approach was not successful, and was even characterized by Carroll (1993, p. 60) as "a somewhat eccentric aberration in the history of intelligence models." It was an understandable initiative to fill a gap and it was therefore based on a quite different view, one that may have seemed an outlier. The only surviving operations from his theory are convergent and divergent production, solving problems with just one correct solution (convergent production) versus solving problems with multiple alternative good solutions (divergent production), which is often seen as a form of creativity. An important lesson from Guilford's endeavor (and similar endeavors in other domains) is that the components of individual differences are often lesser in number than the fine conceptual distinctions experts come up with when asked to make their views explicit. Conceptual taxonomies are appealing, which explains why Guilford's SI theory and framework have been widely cited in the past. They create order in people's thinking, but conceptual structures do not necessarily fit with the empirical reality of individual differences. Horn and Knapp (1973) have documented that the fit some have claimed between the theory and the empirical reality is in fact subjective.

Extension of the Domain and Concept, and a Reformulation of the IQ Measure

The range of views on intelligence is wide, from a narrow cognitive view to a view that intelligence is the ability to deal successfully with the issues of life. David Wechsler, creator of the most popular series of intelligence tests in common use, was an adherent of the broader view. While his predecessors were of the belief that the consequences of intelligence (*causa finalis*) are broad, Wechsler was of the belief that the *causa materialis* was broad and perhaps even relativistic. His writings are ambiguous, perhaps because of his relativistic views. In an article from 1975 (Wechsler, 1975), he takes five pages to explain the impossibility of defining intelligence. Wechsler does not believe in intelligence as one or more underlying invariants: "Intelligence has no invariants, but its components sometimes act as if it does" (p. 138). Among the most common tests in clinical psychology practice and training (depending on age of people being tested) are the Wechsler Adult Intelligence Scale, Fourth Edition (WAIS-IV) (Wechsler, 2008), the Wechsler Intelligence Test for Children, Fifth Edition (WISC-V) (Wechsler, 2014), and the Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV) (Wechsler, 2012). Wechsler was an entrepreneurial psychologist who saw in the Army Alpha and Beta tests postwar surplus of potential value and cleverly created an enduring clinical enterprise out of it. Wechsler died in 1981, but the descendants of his tests still bear his name, just as the Stanford-Binet has memorialized Alfred Binet.

Early on in his career, Wechsler had been sent to England by the US Army to learn from Spearman, but as a clinical psychologist, he thought that Spearman's views were too narrow: "Intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment" (Wechsler, 1944, p. 3). His contribution to the measurement of intelligence is large, primarily because of his creation of the tests that bear his name for use, first in clinical work at the Bellevue psychiatric hospital in Manhattan attempting to see whether patterns of intelligence subtest scores could help with differential diagnosis and treatment planning in psychiatry. He was of the belief that how well test-takers solve the problems presented in intelligence tests depends not only on intellect but also on noncognitive factors, similarly to the factors that also play a role in effectively dealing with one's environment, and therefore, he included these factors in his concept of intelligence (as *causa materialis*). He went even so far as claiming that if one tried to control for these factors, the result would not be a purer measurement of intelligence but a less effective measurement of general intelligence (Wechsler, 1944, p. 9). This is an ambiguous position, although in line with his relativistic thinking. The ambiguity is that he rejects the idea that intelligence is what intelligence tests measure while he at the same time defines intelligence as all the factors that affect performance on intelligence tests.

The innovation in his contribution is that he made use of intelligence test behavior to make inferences that are broader (and more clinically relevant) than those inferences other psychologists had made before him. Although from his interests one may have expected that he would have broadened the domain to include a variety of nonintellectual tasks to broaden the *causa formalis*, he did not do so. Instead he broadened the *causa materialis*. This was partly because, at that time, other measures, such as projective tests like the Thematic Apperception Test, were utilized as purer measures of motivational states and personality traits. Some of the items from the subtest "Picture Arrangement" (understanding stories presented visually as disarranged frames of a comic strip or storyboard) may be interpreted as forms of social intelligence, and he also included performance subtests such as those used in the Army Beta to measure the intelligence of individuals who had not learned the English language. His reason for including these subtests was that they offered better opportunities to make inferences on nonintellectual aspects of general intelligence. Coincidentally, however, the different tests and test clusters (verbally based descendants of the Army Alpha and performance-based descendants of Army Beta) appeared to map well onto broad aspects of brain function, such as right-brain/left-brain asymmetries of function (Doehring, Reitan, & Kløve, 1961; Todd, Coolidge, & Satz, 1977), leading early clinical psychologists to make inferences about brain damage based on patterns of scores on the subtests of the WAIS and WISC (many of which have been debunked by subsequent research; Lezak, Howieson, & Loring, 2004). Clinical practice, therefore, drifted into *causa efficiens*–type explanations based on integrity of different brain regions.

Somewhat in the line of Wechsler's interests, others have proposed even broader extensions of the intelligence domain, to measure social intelligence (Barnes & Sternberg, 1989; Goleman, 2006) and practical intelligence (e.g., Wagner & Sternberg, 1986), although with tasks that are more clearly in line with these types of intelligence than are the tasks from the Wechsler tests. Finally, another important extension of the domain and concept is the emotional intelligence approach (Goleman, 1995; Mayer & Salovey, 1993), with an extension of the *causa formalis* and the *causa materialis*. Already Wechsler was interested in a similar kind of intelligence, but he stayed with an assessment through traditional intelligence test tasks. This shows again the ambiguity of his position. However, the hybrid combination of traditional tasks with a broader and clinically attractive concept may be at the basis of the success of his tests. For example, the WAIS was included in the once popular diagnostic testing approach initiated by Rapaport (Rapaport, Gill, & Shafer, 1945, 1946) together with projective techniques.

A final important contribution of Wechsler is that he moved intelligence scores away from the mental-age concept. He could not believe that the kind of intelligence of a person with a chronological age of, for example, twenty and a mental age of six could be the same as the intelligence of a child of six with a mental age of six. This is a realistic belief in correspondence with the changing nature of intelligence through psychological development. And it is not surprising, given that in his view, intelligence is broader than intellect. Even if the intellect of an adult is perhaps that of a six-year-old, the noncognitive factors of intelligence may have developed further through personality maturation.

Instead of the mental age–based IQ, as explained earlier in this chapter, Wechsler proposed the deviation IQ, expressing how much someone's performance score deviates from the mean performance score of the person's age group. When expressed on a standard scale with mean 100 and standard deviation 15, this is still the current notion of IQ. It is a purely person-comparative measure, a measure that no longer refers in any way to the kind and difficulty of the problems involved. Although moving away from the mental-age concept was necessary for the assessment of adults, it was at the same time an unfortunate development in the sense that IQ lost its reference to tasks and therefore became a less substantive and even more purely individual differences–based measurement. Intelligence became not only "what intelligence tests measure," but also "how people differ on intelligence test scores," without even referring to the tasks included in the tests, in contrast with Binet's tests which were based on the notion of mental age and associated test items to determine the IQ. This new IQ is called the "deviation IQ" because it indicates how much a person deviates from the mean performance in the same group (age and sex groups) a person belongs to.

It is interesting that in contemporary test theory (item response theory, or IRT), measurement is again task comparative, in that the individual's test score is expressed on the same scale as the difficulty levels of the items themselves. It took another educationally inspired scholar from a field outside psychology, Georg Rasch (1960), to develop such an approach. Binet studied law and physiology, whereas Rasch was a mathematician and statistician.

More recently, in a widely used system for interpreting the WAIS-IV (Lichtenberger & Kaufman, 2009), an overall IQ score is supported only if there are no large differences among ability indicators for four clusters of tests – verbal comprehension, working memory, perceptual reasoning, and processing speed – and cluster scores are computed only if there are no large differences between scores on constituent subtests. A similar system has been developed for the analysis of WISC-V scores in child intelligence testing (Kaufman, Raiford, & Coalson, 2016). Factor analyses of WAIS-IV subtest scores now support aggregate scores that map onto CHC theory (Lichtenberger & Kaufman, 2009).

The Measurement Quest for a Causa Efficiens

There have been and still are strong beliefs among some psychologists about the genetic origin of individual differences in measured intelligence. For an overview of the literature, see Deary, Penke, and Johnson (2010), and for a report on the most recent progress in explaining the genetic contribution to intelligence scores, see Plomin and von Stumm (2018). The genetic and environmental contributions to intelligence have been empirically approached in various ways, including through research with monozygotic twins and adopted children and through analyses of genes. We will not discuss these lines of research, but instead we review research involving elementary cognitive tasks, based on the hypothetical reasoning from the part of the researchers in question that the simpler the tasks are, the less influence there is from learning and experience and, therefore, the closer one would get to the ultimate basis of intelligence.

Measures of mental speed from simple experimental tasks, such as mean simple reaction time, its variance (time in response to a stimulus and how much it varies), choice reaction time (time to make a choice between response options), and inspection time (time needed to identify a stimulus) (Barrett, Petrides, & Eysenck, 1998; Eysenck, 1967; Jensen, 1982), were investigated as correlates of IQ, clearly with some success, as summarized by Sheppard and Vernon (2008). Eysenck also discussed biological underpinnings and neurochemical pathways (Eysenck, 1982, 1988).

An interesting new development with respect to elementary cognitive tasks and choice response times is that the effect of the speed-accuracy trade-off can be separated from efficiency of information accumulation in the drift diffusion model (DDM) and that the efficiency parameter (drift rate) appears to be a cognitive correlate of IQ (Ratcliff et al., 2010). The DDM stipulates that people extract relevant information from a stimulus in a two-alternative forced choice task (e.g., a numerosity task wherein a subject is asked to indicate whether a 10 x 10 grid holds less or more than fifty asterisks) and that the balance between the two options evolves with a trend (called drift rate) toward one of both possible answers but with random disturbances of the trend while the information cumulates, until one of the two decision boundaries is reached and a choice is made. The process takes more time the weaker the trend is and the farther the decision boundaries are separated. Wider boundaries imply that more evidence is needed before one can decide, and so reflect caution on the part of the examinee, but they also lead to a higher degree of accuracy. The separation width of the boundaries determines the speed-accuracy trade-off.

Is this a return to the early work of James McKeen Cattell? In a way it is, but in two other ways it is not. First, the simple cognitive tasks Cattell proposed and used were considered as *causa formalis*, whereas in the more recent work, what one intends to measure is the *causa efficiens* or an elementary *causa materialis*. Second, response times are confounded with the speed-accuracy trade-off, and only after controlling for the trade-off does one obtain the true measure (Ratcliff et al., 2010), in line with Spearman's idea of controlling for confounds. In fact, Spearman (1927) was already aware of the speed-accuracy trade off (p. 250).

Performance (mostly response times) in simple cognitive tasks forms one end – the simple one – of the cognitive-correlates range. At the other, more complex end is working memory (Engle et al., 1999; Kyllonen & Christal, 1990), with work by Hunt, Frost, and Lunneborg (1973) and Hunt, Lunneborg, and Lewis (1975) situated in between the two extremes. In the cognitive-correlates approach, tasks are presented that do not belong to intelligence tests but are correlated with a measurement of intelligence hoping to find correlates that can be interpreted as contributing factors. For example, Hunt and colleagues (1975) used letter comparison tasks (AA, Aa, ab, bA, etc.) with the request to indicate whether the letters were the same. It turned out that more time was needed if the letters were the same but physically different (a and A) and that the extra time necessary was negatively correlated with verbal ability. This result was seen as evidence that access to letter names is a correlate of verbal ability. The cognitive-correlates studies are all endeavors to uncover the *causa efficiens* as more or less distal factors (rather than decompositions of intelligence test tasks) and therefore possible causes. Some of these measures were also considered as measures of intelligence through its *causa efficiens*.

Robert J. Sternberg's Revolution

To discuss Robert J. Sternberg's contribution to the field of intelligence, a distinction must be made between his early work (e.g., Sternberg, 1977a, 1977b) and his later work on the triarchic theory of intelligence (e.g., Sternberg, 1985) and his theory on successful intelligence (Sternberg, 1997). Some would see his later work as revolutionary, and from a measurement point of view, his early work certainly is. Sternberg did not start from a definition of intelligence; the aim of his early work was to find out what a definition could be. In our interpretation, he wanted to find the *causa materialis* (what intelligence consists of), not primarily in terms of individual differences but in terms of mental processes. He considered his approach as an antidote to factor analysis and the "psychometric" approach (Sternberg, 1977a). Ironically, at about the same time, Susan Embretson, a psychometrician, set out a very similar research agenda, using precisely psychometric models (e.g., Whitely, 1976, 1977).

Why the work of these two scholars is very similar will be explained after Sternberg's approach is discussed.

The basic feature of Sternberg's early work is that he created differences between tasks (analogies) by manipulating process-related task attributes based on a process theory for how people solve the tasks (Sternberg, 1977a, 1977b). This is a cognitive-psychological experimental method of modeling cognitive processes. The revolutionary part of the approach was that he looked at individual differences in the effects of the task attributes on the total time used to solve the problems. In the 1970s, at the time of his work, mixed models were not yet popular, while he was basically formulating and using mixed models for response times, but in a premature way because the appropriate methods were not yet available. Mixed models are regression models with varying regression weights. In Sternberg's work the predictors in the regression are task attributes and the predictor weights indicate the effect of the task attributes on the response time. In a mixed model, the weights of the predictors are allowed to vary across persons.

Mixed models may sound like just another method, one of many from the statistical toolbox, but they are in fact a fundamental and revolutionary concept when applied to cognitive tasks. It is a common but misleading idea that statistical methods are just tools; instead, they also are theories about the data expressed mathematically rather than in words, as nicely illustrated in Sternberg's (1977a, 1977b) early work. The mixed models he used are the formal side of a theory saying that the problems are being solved through well-defined hypothesized processes and that people differ in the time each of these processes take. Instead of using the now easily available software for mixed models, he made use of one multiple regression per individual. The resulting regression weights are measurements of cognitive processes, a powerful method to reveal and measure covert processes. Each of the weights refers to a cognitive component (process) of the responses to analogy problems.

Next, these estimates of individual regression coefficients were used in turn as predictors for reasoning ability tests scores, and the multiple correlations were found to be very high, from about 0.70 to almost 0.90, depending on the type of analogy (Sternberg, 1977b). The most predictive components could then be interpreted as the nature of intelligence as far as reasoning is concerned. Problem solved. His 1977 *Psychological Review* article (Sternberg, 1977b) was a reason for optimism. Not only did the results give us the impression that the nature of intelligence (and a definition) could be found, but after more than fifty years, the experimental and correlational approach were again integrated in a cross-fertilizing way for the study and measurement of intelligence. Unfortunately, as was later found out, it was too optimistic an interpretation of the potential of the new approach. For the sake of completeness, Sternberg has used the same method for correctness of the response data (instead of response time) as the dependent variable, but the analogy tasks he used were so easy that the accuracy data were not really informative (Sternberg, 1977a). He has also applied the same methods to other types of problems in addition to just analogies (Sternberg, 1983).

The better approach for binary data such as accuracy of responses (instead of response times) was independently introduced by Embretson (who first published under the name of Whitely) (Whitely, 1976, 1977). Using psychometric models for correct versus incorrect responses, Embretson did the very same as Sternberg did for response times. Her models too were early versions of mixed models, but taking the binary nature of the data into account (correct versus incorrect responses). In particular they were IRT models adjusted to incorporate task attributes. Her approach was a joint analysis of cognitive processes and individual differences. Although the ideas of both scholars were excellent and revolutionary, they have not led to breakthroughs, either with respect to finding the *causa materialis* or with respect to methods of measurement that are broadly used.

It is an unfortunate outcome, first, because the logic of starting from the domain (the *causa formalis*) to find the substance (*causa materialis*) is in principle an excellent and open approach (rather than starting from a definition); second, it is an integration of the two major approaches in psychology (experimental psychology and individual differences); and third, the cognitive-components approach (individual differences in the effects of task differences) has a stronger reference basis in tasks because the components refer directly to effects of task attributes. The results of such an approach have more meaning and explanatory value than purely interpersonal comparison of performance levels. The individual differences now clearly refer to the effects of task attributes and the corresponding cognitive processes. In addition, when item response theory is used as in Embretson's work, it is also possible to compare persons with items and with the process-based features of tasks.

The major reason why the cognitive-component approach to measuring intelligence is no longer pursued is a change in the views of Sternberg himself. He found that the components do not have an incremental value beyond the traditional way of measuring intelligence and he switched from a *causa materialis* approach to a *causa finalis* approach, which has led to his triarchic theory of intelligence. In that theory, the cognitive components from his earlier work, applied to abstract problems, are interpreted as the basis for analytical intelligence, while the other two types of intelligence in the triarchic theory (Sternberg, 1983) are creative intelligence and practical intelligence (where the components are applied in more complex ways). In combination, they are the basis of meaningful accomplishments in life. His later theory of successful intelligence (Sternberg, 1997) is an extension of the triarchic theory with the idea that intelligence implies capitalization on one's strengths and compensation for and correction of one's weaknesses. Finally, in the augmented theory of successful intelligence, the notion of wisdom is also given a prominent place (Sternberg, 2003). Sternberg's triarchic theory also came with a test, the Sternberg Triarchic Abilities Test (Sternberg et al., 2001). The test is no longer used (see Sternberg, 2010). The identification criterion for intelligence in the later work of Sternberg and for the way he measures intelligence is the finality of intelligence: a successful life according to one's own norms.

It is easy to see that this approach has more appeal than the rather laborious cognitive-components experiments followed by a not so easy data analysis, with a result that has no incremental predictive value (Sternberg, 1981). Other reasons are the dominance and popularity of factor models, not only for intelligence. In the minds and practice of test developers and for the operational definition of constructs, factor models have become and still are considered by many as a necessary ingredient everyone is familiar with. The alternatives of mixed models proposed by Sternberg and Embretson are much less familiar and have the (incorrect) reputation for being difficult to apply, even though (ironically), formally speaking, they are simplified versions of the now so popular confirmatory factor models. Sociological factors sometimes drive what is considered the optimal approach. Cognitive components have lost their innovation potential. Neurobiological approaches have taken over as one among several innovative approaches in the search for the Holy Grail of intelligence (e.g., Santarnecchi et al., 2017).

The End of the History of Intelligence Measurement

Perhaps we are experiencing the end of the history of intelligence measurement, neither because of a general consensus on how to measure intelligence, nor because one method works better than other methods, but because we have seen everything: measurement approaches based on all four of Aristotle's *causae*. Another way of measuring intelligence does not seem possible; they can only be variants of the same. Or is this too easy as a conclusion? The future always surprises.

Although all measurement methods may seem meaningful, especially given the historical context from which they stem, apparently from none of the methods can one expect that intelligence is measured in its true nature, which is still unknown, if any true nature exists at all with respect to intelligence. However, for many approaches, it has also been shown that they are useful and have predictive value.

Our tentative conclusion is that intelligence is an undefinable concept, not clearly delineable. For some, it is rather narrow, for others, it is very broad, and it probably evolves depending on time and changes to some extent with the culture. It is a useful construction, but "nothing in particular, it is a construct humans have invented" (Sternberg, 2018, p. 308). Such a view can be interpreted in line with two other relativistic views, an old view on intelligence as an alternative to Spearman's general factor and a recent view on emotion away from emotions being invariant notions. The former is Thomson's (1916) sampling theory, which implies that intelligence is a loose population of elements (called "bonds" by Thomson) from which elements are sampled depending on the kind of task. The elements are not further specified, but one can think, for example, of cognitive processes or connections in the brain. The theory is experiencing a revival, among others through Kovacs and Conway's (2016) process overlap theory, which implies that positive correlations between tests stem from overlap of processes involved in test performance and that general intelligence and thus *g* is constituted by the total set of all these processes instead of being itself a constituting factor underlying test performance. The latter is Feldman Barrett's theory of constructed emotion: "Anger is a population of diverse instances, not a single automatic reaction in the true sense of the phrase. The same holds for every other category of emotion, cognition, perception, and other type of mental event" (Feldman Barrett, 2017, p. 223). This seemingly nihilist view does not imply that intelligence does not exist; it does exist, not as an essence or an invariance, but instead as a constructed and not clearly delineated set of elements (*per causa* or from all four *causae*). Intelligence does exist in the same sense as most of our everyday concepts "exist" and have meaning, like "sadness," "joy," "love," "altruism," etc., as rather loose somewhat overlapping clusters of exemplars. Just as saltimbocca and other dishes can be found on the menu of a restaurant, "intelligence" can be found on the menu of research in psychology. Experts agree on many examples of intelligence and on many counterexamples, even without a definition, and there also is a gray zone where experts have contradictory opinions. Intelligence also is an open concept in that ever more instances can be generated, just as for the concept of "sentence." Some sentences are accepted as well-formed sentences and others barely are, or they are not. All depends on changing criteria for what a sentence is. However, independent of the criteria, it is always possible to generate new sentences. A possible biological explanation for the open and distributed concept of intelligence can be found in Geary's (2018) conjecture that the efficiency of mitochondrial functioning is related to general intelligence.

Ironically, intelligence is at the same time the most successful construct in psychology with regard to measurement while it has become famous for its reputation of being impossible to define. Apparently, measurement does not imply clarity on the variable that is being measured. Intelligence seems a concept without an essence. Invariance in the sense Spearman was looking for may not exist, as already noted by Wechsler (1975), not even in milder versions such as types of intelligence. Not everyone may agree, especially because it goes against the way we like to think in our discipline about psychological constructs and the way we humans think about everything else. The human mind works with categories as representations. One may think that a dimension such as intelligence, or multiple intelligence dimensions, contradict the categorizing mind, but a dimension with an identity and with invariance is just another variant of how the need for definable categories is fulfilled. Intelligence is not a clear concept, and, based on the lack of consensus regarding its definition, it is possible that it never will be. However, apparently the lack of a definition of intelligence does not prevent its measurement from being useful and having predictive value, just as one who orders saltimbocca has some idea of what to expect. Though the dish has no precise culinary definition, diners who appreciate it develop an understanding of the commonalities across different exemplars which are loosely related to one another.

References

Agnoli, P., & D'Agostini, G. (2005). Why does the meter beat the second? https://arxiv.org/abs/physics/0412078

Aquinas, Thomas (1961). *Commentary on Aristotle's Metaphysics* (trans. John P. Rowan). Notre Dame, IN: Dumb Ox.

Barnes, B., & Sternberg, R. J. (1989). Social intelligence and decoding of nonverbal cues. *Intelligence*, *13*, 263–282.

Barrett, P. T., Petrides, K. V., & Eysenck, H. J. (1998). Estimating inspection time: Response probabilities, the BRAT IT algorithm, and IQ correlations. *Personality and Individual Differences*, *24*, 405–419.

Binet, A., & Simon, T. (1905). Méthodes nouvelles pour le diagnostic du niveaux intellectuel des anormaux. *L'Année Psychologique*, *11*, 191–244.

Binet, A., & Simon, T. (1908). Le développement de l'intelligence chez les enfants. *L'Année Psychologique*, *14*, 1–95.

Black, E. (2003). *War against the weak: Eugenics and America's campaign to create a master race*. New York: Four Walls Eight Windows.

Burt, C. (1949). The structure of the mind: A review of the results of factor analysis. *British Journal of Educational Psychology*, *19*, 100–111, 176–199.

Carroll, J. B. (1993). *Human cognitive abilities*. New York: Cambridge University Press.

Cattell, J. M. (1890). Mental tests and measurements. *Mind*, *15*, 373–380.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1–22.

Cianciolo, A. T., & Sternberg, R. J. (2004). *Intelligence: A brief history*. Malden, MA: Blackwell.

De Boeck, P. (2013). Intelligence, where to look, where to go? *Journal of Intelligence*, *1*, 5–24.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Review Neuroscience*, *11*, 201–211.

Doehring, D. G., Reitan, R. M., & Kløve, H. (1961). Changes in patterns of intelligence test performance associated with homonymous visual field defects. *Journal of Nervous and Mental Disease*, *123*, 227–233.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Eysenck, H. (1967). Intelligence assessment: A theoretical and experimental approach. *British Journal of Educational Psychology*, *37*, 81–98.

Eysenck, H. J. (1982). *A model for intelligence*. Berlin: Springer-Verlag.

Eysenck, H. J. (1988). The biological basis of intelligence. In S. H. Irvine & J. W. Berry (Eds.), *Human abilities in cultural context*. New York: Cambridge University Press.

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Härtel.

Feldman Barrett, L. (2017). *How emotions are made. The secret life of the brain*. Boston: Houghton Mifflin Harcourt.

Galton, F. (1883). *Inquiry into the human faculty and its development*. London: MacMillan.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligence*. New York: Basic Books.

Geary, D. C. (2018). Efficiency of mitochondrial functioning as the fundamental biological mechanism of general intelligence. *Psychological Review*, 125, 1028–1050.

Giacomo, P. (1984). News from the BIPM. *Metrologia*, 20, 25–30.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.

Goleman, D. (2006). *Social intelligence: The new science of human relationships*. New York: Bantam Books.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24, 13–23.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, *53*, 267–293.

Gustaffson, J. (1984). A unifying model for the structure of human intellectual abilities. *Intelligence*, *8*, 179–203.

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, *2*, 41–52.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, *75*, 242–259.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystalized generalized intelligences. *Journal of Educational Psychology*, *57*, 253–270.

Horn, J. L., & Knapp, J. R. (1973). On the subjective character of the empirical base of Guilford's structure-of-intellect model. *Psychological Bulletin*, *80*, 33–43.

Hunt, E., Frost, N., & Lunneborg, C. (1973). Individual differences in cognition: A new approach to cognition. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (vol.7, pp. 87–122). New York: Academic Press.

Hunt, E., & Jaeggi, S. M. (2013). Challenges for research on intelligence. *Journal of Intelligence*, *1*, 36–54.

Hunt, E., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology*, *7*, 194–227.

Intelligence and Its Measurement: A Symposium. (1921). *Journal of Educational Psychology*, *12*, 123–147.

Jensen, A. R. (1982). The chronometry of intelligence. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol.1, pp. 255–310). Hillsdale, NJ: Lawrence Erlbaum.

Johnson, W., & Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33, 393–416.

Kaufman, A. S., Raiford, S. E., & Coalson, D. L. (2016). *Intelligent testing with the WISC-V.* Hoboken, NJ: Wiley.

Kell, H. J., & Lang, J. W. B. (2017). Specific abilities in the workplace: More important than *g*? *Journal of Intelligence*, *5*, 13. https://doi.org/10.3390/jintelligence5020013

Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, *27*, 151–177.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*, 389–433.

Legg, S., & Hutter, M. (2007). *A collection of definitions of intelligence*. Technical report. www.researchgate.net/publication/1895883

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Lichtenberger, O., & Kaufman, A. S. (2009). *Essentials of WAIS-IV assessment*. New York: Wiley.

Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, *17*, 433–442.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10.

Melnick, M. D., Harrison, B. R., Park, S., Benneto, L., & Tadin, D. (2013). A strong interactive link between sensory discriminations and intelligence. *Current Biology*, *23*, 1013–1017.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature Review Genetics*, *19*, 148–159.

Rapaport, D., Gill, M., & Schafer, R. (1945). *Diagnostic psychological testing*, vol. 1. Chicago: The Year Book Publishers.

Rapaport, D., Gill, M., & Schafer, R. (1946). *Diagnostic psychological testing*, vol. 2. Chicago: The Year Book Publishers.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danish Institute for Educational Research.

Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, ageing, and IQ in two-choice tasks. *Cognitive Psychology*, *60*, 127–157.

Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence*. London: Lewis.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales, Fifth Edition*. Itasca, IL: Riverside Publishing.

Roid, G. H., & Tippin, S. M. (2009). Assessment of intellectual strengths and weaknesses with the Stanford-Binet Intelligence Scales – Fifth Edition (SB5). In J. A. Naglieri & S. Goldstein (Eds.), *Practitioner's guide to assessing intelligence and achievement*. Hoboken, NJ: John Wiley & Sons.

Santarnecchi, E., Emmendorfer, A., Tadayon, S., Rossi, S., Rossi, A., & Pascual-Leone, A. (2017). Network connectivity correlates of variability in fluid intelligence performance. *Intelligence*, *65*, 35–47.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*, 414–429.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Intelligence*, *44*, 535–551.

Spearman, C. (1904a). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–295.

Spearman, C. (1904b). The proof and measurement of association between two things. *American Journal of Psychology*, *15*, 72–101.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition*. London: Macmillan.

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Stern, W. (1912). *Die psychologischen Methoden der Intelligenzprüfung und deren Anwendung an Schulkindern*. Leipzig: Barth.

Sternberg, R. J. (1977a). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1977b). Component processes in analogical reasoning. *Psychological Review*, *84*, 353–378.

Sternberg, R. J. (1981). Intelligence and non-entrenchment. *Journal of Educational Psychology*, *73*, 1–16.

Sternberg, R. J. (1983). Components of human intelligence. *Cognition*, *15*, 1–48.

Sternberg, R. J. (1985). *Beyond IQ*. New York: Cambridge University Press.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. New York: Cambridge University Press.

Sternberg, R. J. (2010). *College admissions for the 21st century*. Cambridge, MA: Harvard University Press.

Sternberg, R. J. (2018). Successful intelligence in theory, research, and practice. In R. J. Sternberg (Ed.), *The nature of intelligence* (pp. 308–322). New York: Cambridge University Press.

Sternberg, R. J., Castejón, J. L., Prieto, M. D., Hautamäki, J., & Grigorenko, E. (2001). Confirmatory factor analysis of the Sternberg Triarchic Abilities Test (multiple-choice items) in three international samples: An empirical test of the triarchic theory of intelligence. *European Journal of Psychological Assessment*, *17*, 1−16.

Sternberg, R. J., & Detterman, D. K. (Eds.) (1986). *What is intelligence?* Norwood, NJ: Ablex.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., et al. (2000). *Practical intelligence in everyday life*. Cambridge, MA: Cambridge University Press.

Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.

Thomson, G. H. (1916). A hierarchy without a general factor. *British Journal of Psychology*, *8*, 271–281.

Thorndike, E. L. (1921). Intelligence and its measurement: A symposium – I. *Journal of Educational Psychology*, 12, 124–127.

Thurstone, L. L. (1924). *The nature of intelligence*. London: Kegan Paul, Trench Trubner & Co.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Todd, J., Coolidge, F., & Satz, P. (1977). The Wechsler Adult Intelligence Scale Discrepancy Index: A neuropsychological evaluation. *Journal of Consulting and Clinical Psychology*, *45*, 450–454.

Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen.

Wagner, R. K., & Sternberg, R. J. (1986). Tacit knowledge and intelligence in the everyday world. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 51–83). Cambridge, MA: Cambridge University Press.

Wechsler, D. (1944). *The measurement of adult intelligence* (3rd ed.). Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1975). Intelligence defined and undefined: A relativistic appraisal. *American Psychologist*, *30*, 135–139.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale – Fourth Edition*. San Antonio, TX: Pearson.

Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence™ – Fourth Edition*. San Antonio, TX: Pearson.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children® – Fifth Edition*. San Antonio, TX: Pearson.

Whitely, S. E. (1976). Solving verbal analogies: Some cognitive components of intelligence test items. *Journal of Educational Psychology*, *68*, 234–242.

Whitely, S. E. (1977). Information processing on intelligence test items: Some response components. *Applied Psychological Measurement*, *1*, 465–476.

Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review, Monograph Supplement*, *3* (6).

Wundt, W. (1902). *Grundriss der Psychologie*. Leipzig: Engelmann.

Yerkes, R. M. (1921). *Psychological examining in the US Army*. Memoirs of the National Academy of Sciences, vol. 15. Washington: US Government Printing Office.

5 Factor-Analytic Models of Intelligence

Robert Walrath, John O. Willis, Ron Dumont, and Alan S. Kaufman

> The great tragedy of Science – the slaying of a beautiful hypothesis by an ugly fact.
>
> Thomas Huxley (1870)

> First get your facts, then you can distort them at your leisure.
>
> Mark Twain (in Kipling, 1899)

Clearly, there are many ways to define intelligence. Wasserman (2019, pp. 37–38) lists thirty-one definitions of the concept, provided by psychologists from Herbert Spencer (1855) to Robert J. Sternberg (1986), based in part on those psychologists who responded in 1921 to a survey regarding their opinions about the definition of the term *intelligence*. Sternberg and Detterman (1986) provided an updated symposium with more definitions and some overlap of components. Sattler (2008, p. 223) provided an additional list of nineteen different definitions that have been suggested over the years by several of the major experts in the field of psychology. Although intelligence, like Freud's *ego*, is probably best thought of as a process, it is treated in much of the literature and often in professional practice as a thing. The lack of a single, accepted definition of intelligence contributes, of course, to disagreements about how to assess intelligence. Without agreement on the definition of intelligence – and even on whether IQ exists – it is difficult to reach agreement on how to measure intelligence.[1]

Global Intellectual Ability versus Separate Abilities

A persistent and unresolved question in both professional theories and lay conceptualizations of intelligence has been whether an individual has one overall level of "intelligence" or, instead, whether what we call "intelligence" is actually

1. For information about the major theories of intelligence that have influenced testing, see Carroll (1993, pp. 30–72); Chen and Gardner (2019); Daniel (1997); Horn and Blankson (2012); A. S. Kaufman (2009, 2018); A. S. Kaufman, Raiford, and Coalson (2016, pp. 6–20); Kornhaber (Chapter 27 in this volume); Naglieri, Das, and Goldstein (2012); Sattler (2018, pp. 223–255); Schneider and McGrew (2019); and Sternberg (2019; see also Sternberg (Chapter 3 in this volume; Chapter 28 in this volume). For information regarding some of the many disputes about the construct and measurement of intelligence, see, for example, Eysenck versus Kamin (1981); Gould (1981); Herrnstein and Murray (1994); Jacoby and Glauberman (1995); Jensen (1998); A. S. Kaufman, Schneider, and Kaufman (in press); and Sternberg(Chapter 1 in this volume) among a great many other sources (it is a contentious field).

a set of several separate abilities. These theorists could be characterized respectively as "lumpers" and "splitters" (McKusick, 1969). Although apparently dichotomous, this fundamental question has spawned continua of hotly debated theories.

At one end, there is the extreme lumper position that each person has a single level of cognitive ability (often referred to as *g*, as discussed below). The *expression* of this intelligence may vary with different tasks, and as a function of education, sensory and motor abilities, and other influences, but the individual has one, single level of reasoning ability that will be seen on a wide variety of intelligence tests. This theoretical perspective matches the common observation that, among our friends and acquaintances, some individuals are consistently pretty smart about almost everything and some are consistently incompetent and clueless. Most of us can categorize the people we know as "smart," "dumb," or something in between. Theorists and practitioners who adhere to this position tend to consider the total score on an intelligence test to be an approximation of the individual's overall level of intelligence, although scores will vary somewhat across different tests.

The opposite extreme, the splitter end of this continuum, is the position that there is a set of several higher-order cognitive abilities that are more or less independent of each other. A person might demonstrate, for example, a high level of verbal knowledge, vocabulary, and verbal reasoning ability, but at the same time might be weak in visual-spatial thinking and unable to read a map or to "see" how a decorator's floor plan would translate into the actual layout of furniture in the real room. Most of us can think of acquaintances who may be terribly clever in some ways and notably incompetent in others. Theorists and practitioners who adhere to this extreme splitter position tend to ignore or de-emphasize total scores on intelligence tests and to focus on patterns of strengths and weaknesses.

Other splitter theorists focus their attention on different mental *processes* (rather than a set of discrete *abilities*), such as planning; attention; and dealing with information in a step-by-step, sequential process or in an all-at-once holistic approach. Again, this theoretical perspective is mirrored in popular psychology. People, even if not professionally trained in intelligence theory, may nonetheless characterize themselves and others as, for example, either sequential or holistic thinkers.

Still other splitter theorists object to the narrow scope of intelligence as it is measured by most existing intelligence tests. They note that the oral question-and-answer, paper-and-pencil, and picture-and-puzzle intelligence tests de-emphasize or entirely omit such essential capacities as practical intelligence, creativity, artistic and musical abilities, and rational thinking. In an early episode of the popular comedy show *Everybody Loves Raymond* (1996), Ray and his wife Debra allow themselves to be given an IQ test by Ray's brother, and the results show Debra to have scored higher then Ray. Ray responds to this finding by claiming, "I didn't know he was going to learn that you're smart and I'm an idiot," to which Debra responds, "There's more than one kind of intelligence, Ray." Not to be outdone, Ray's father chimes in, "That's right. There's street smarts."

General Intelligence – Spearman's g

British psychologist Charles Spearman (1904) proposed a conception of intelligence perhaps most widely (though by no means universally) accepted by authors and users of intelligence tests. His idea was that each person has a certain general level of intellectual ability, which the person can demonstrate in most areas of endeavor, although it will be expressed differently under different circumstances. This general intelligence is commonly referred to by the single italicized letter *g*.

As noted above, Spearman's general ability theory is appealing on a commonsense level. One often notices, for example, that some colleagues are generally pretty smart at most things while others have a lack of ability that seems to extend with equally broad application to many endeavors. There is also, as Spearman showed, statistical support for the general ability theory. Using the statistical techniques of factor analysis (a group of various mathematical methods using statistical correlations to investigate groups of tests or subtests that have higher correlations with each other and lower correlations with other tests or subtests; see, for example, Carroll, 1985; Keith & Reynolds, 2019; and for a brief, very clear, basic explanation, A. S. Kaufman, Schneider, & Kaufman, in press) to examine a number of mental aptitude tests, he observed that people who performed well on one cognitive test tended to perform well on other tests, while those who scored poorly on one test tended to score poorly on others. Spearman demonstrated that measures of different mental abilities correlated substantially with each other. People with high verbal abilities are likely also to have high spatial and quantitative abilities, and so on. Persons with higher IQs apparently are also likely to be taller and have more body symmetry than persons with lower ability scores (Prokosch, Yeo, & Miller, 2005; Silventoinen et al., 2006; see also Haier, Chapter 19 in this volume). Spearman postulated that those positive correlations across different tests indicated that there must be a general function or "pool" of mental energy, which he named the general factor, or *g* (Spearman, 1904, 1927). Spearman also acknowledged specific factors (*s*) representing particular tests or subtests, but he was slow to accept the concept of separate abilities such as fluid reasoning or processing speed that generalized across several tests.

Most intelligence tests in use today are based, at least in part, on general ability theory. Critics (e.g., Gould, 1981; Horn & Blankson, 2012) assert that correlations with older tests based on the *g* theory are used to justify new tests based on the same theory, which, they claim, simply adds more circular and artificial support to the construct of *g*.

Many immediate or enduring nonintellectual influences can affect the expression of *g* (e.g., Wechsler, 1926). For instance, a math "phobia," lack of training in higher math, or an interacting combination of the two forces could prevent the successful expression of a person's full *g* in the area of mathematics.

Some problems require more than *g* for their solution. For instance, solving problems in engineering, housekeeping, teaching, farming, mechanics, and medicine usually requires specialized knowledge, skills, and ways of thinking (See Ackerman, Chapter 48 in this volume; Flanagan, Ortiz, and Alfonso, 2013, pp. 13, 17, 390–391;

Flynn, 2007; Hedlund, Chapter 30 in this volume; Horn & Blankson, 2012, p. 78). Furthermore, emotions and intellect often interact, sometimes aiding and sometimes interfering with one another in solving problems (e.g., Wechsler, 1940, 1943, 1950). For example, frustration tolerance, impulsiveness, and persistence are important components of test performance.

The *g* theory of intelligence is not necessarily linked to theories of either hereditary or environmental influences on intelligence (e.g., Eysenck versus Kamin, 1981), although proponents of the *g* theory often emphasize hereditary influences (e.g., Bouchard, 2018; Gottfredson, 2008). The idea necessary for acceptance of the *g* theory is that intelligence operates primarily as a single capacity.

Aging, brain damage, disease, deprivation, and disturbance are known to affect some expressions of intelligence differentially (Horn & Blankson, 2012, pp. 78–82; McGrew, LaForte, & Schrank, 2014, pp. 146–143). For example, a stroke may impair one function, such as speech, while sparing others, such as drawing. Sacks (1970) offers many instructive examples of differential effects of diseases and injuries. Springer and Deutsch (1993) discuss early split-brain studies. General ability theorists might hold that it is the expression of intelligence that is affected, and that intelligence itself is still mostly unitary, even though its application is unevenly handicapped.

For more than three-quarters of a century, Spearman's *g* theory was the only one that mattered for practical assessment of intelligence (see Sternberg, Chapter 3 in this volume). Indeed, Spearman's *g* was at the root of Terman's (1916) Stanford-Binet adaptation of Binet's test (Binet & Simon, 1916/1980) in the United States, forming the foundation for offering only a single score, the global IQ (A. S. Kaufman, 2009). However, as explained by A. S. Kaufman and colleagues (in press; section "The Glorious Death of Spearman's Two-Factor (*g*) Theory," para. 3):

> Shortly after Spearman's theory was proposed, several scholars (e.g., Cattell, 1941; Thurstone, 1931) obtained data showing that there were many correlations that could not be fully explained by a general factor. Two decades after the strict version of Spearman's theory was proposed, he was still calling his theory the "Two-Factor Theory" and he focused on general ability as the essence of intelligence, but even his own data obliged him to acknowledge the existence of a third kind of factor, the group factor, by which similar tests could be grouped together (Spearman, 1927, p. 222).

Nonetheless, until 1939, intelligence tests generally offered only a total score to be taken as an approximation of *g*. David Wechsler's (1939) Wechsler-Bellevue Intelligence Scale offered, in addition to the Full Scale IQ (a proxy for *g*), two component IQs (Verbal and the nonverbal Performance), which inspired an industry of profile analysis as clinicians and researchers interpreted various patterns of subtest scores from diverse perspectives (e.g., A. S. Kaufman, 1979, 1994; Rapaport, Gill, & Schafer, 1945–1946; Zimmerman & Woo-Sam, 1973). Ultimately, another industry was formed that was dedicated to condemnation of the practice of profile interpretation – for example, McDermott, Fantuzzo, and Glutting (1990), who proclaimed, "Just say no to subtest analysis: A critique on Wechsler theory and practice." That debate continues to the present day (e.g., Canivez & Watkins, 2016; Flanagan & Kaufman, 2009; A. S. Kaufman, 2009; A. S. Kaufman et al., 2016; Lichtenberger &

Kaufman, 2013; Watkins, Glutting, & Youngstrom, 2005). Ironically, Wechsler provided clinicians with a profile of IQs and subtest scaled scores to interpret – and he championed the interpretation of subtest profiles for diagnosis of brain damage and psychopathology (Benisz, Dumont, & Willis, 2015; A. S. Kaufman, 2009; Wasserman, 2019; Wechsler, 1958) – but he always considered the Wechsler-Bellevue and all his subsequent intelligence scales to be measures of global intellectual ability, measures of *g*. In fact, Wechsler's personal definition was "Intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment" (Wechsler, 1940. pp. 444–445).

Thurstone's Primary Mental Abilities

Other theorists (e.g., Edward L. Thorndike, 1927) have historically placed more importance on separate areas of intelligence and argued that *g* and specific factors (referred to as "*s*" by Spearman) interact to determine the expression of intelligence in different situations. There are many different conceptions of the specific mental factors. In 1938, Louis L. Thurstone, an outspoken opponent of Spearman's *g*, offered a differing theory of intelligence. Thurstone, who had developed methods for scaling psychological measures, assessing attitudes, and testing theory, developed new factor-analytic techniques (statistical procedures to reduce large numbers of *variables* into fewer numbers of *factors*) to determine the number and nature of latent constructs within a set of observed variables. Using his new methods, Thurstone argued that Spearman's *g* resulted from a statistical artifact based upon the mathematical procedures that Spearman had used. Thurstone believed that human intelligence should not be regarded as a single unitary trait, and in its place, he proposed the theory of primary mental abilities (1938), a model of human intelligence that challenged Spearman's unitary conception of intelligence.

Thurstone's early (1931) theory, based upon an analysis of mental test data from samples composed of people with similar overall IQs, suggested that intelligent behavior does not arise from a general factor, but instead emerges from different "primary mental abilities" (Thurstone, 1938). The abilities that he described were verbal comprehension, inductive reasoning, perceptual speed, numerical ability, verbal fluency, associative memory, and spatial visualization.

British psychologist P. E. Vernon (1950, 1961) proposed a hierarchical group factor theory of the structure of human intellectual abilities, based upon factor analysis. His proposed intellectual structure had at the highest level general ability (*g*) with major, minor, and specific factors tiered below *g*. Major factors were verbal-educational and spatial-mechanical, while the minor group included such factors as verbal fluency, numerical, and psychomotor abilities. Specific factors (lowest in the hierarchy) referred to narrow ranges of behavior. Because Vernon's theory included both a general factor and group factors, it may be viewed as something of a compromise between Spearman's two-factor theory (which was composed of *g* and *s*, but did not include group factors) and Thurstone's multiple-factor theory (which did not have a general factor).

Guilford's Structure of Intellect Model

One prominent multifactor theorist was J. P. Guilford (1967), who devised the Structure of Intellect (SOI) model. Guilford's theory laid out, in a three-dimensional model, five different mental operations needed to solve problems (such as *convergent production* or *divergent production*) on four different contents (such as *symbolic* or *figural*), yielding six kinds of products (such as *classes* or *relations*) for a total of 120 (5 x 4 x 6 = 120) possible intellectual factors. Guilford's model, because of the huge number of intellectual abilities it posited, was the most dramatic contrast to Spearman's unitary *g* theory.

Despite the clear distinction between Spearman's single-factor model and Guilford's multidimensional model, both suffered from a similar problem. As A. S. Kaufman (2009, p. 52) notes:

> If one ability was too few to build a theory on, then 120 was just as clearly too many. And Guilford did not stop at 120. He kept refining the theory, adding to its complexity. He decided that one Figural content was not enough, so he split it into figural-auditory and figural-visual (Guilford, 1975). Nor was a single memory operation adequate, so he subdivided it into memory recording (long-term) and memory retention (short-term) (Guilford, 1988). The revised and expanded SOI model now included 180 types of intelligence!

Verbal-Perceptual-Image Rotation Model

Thomas Bouchard, Wendy Johnson, and colleagues (e.g., Bouchard, 2018; Johnson & Bouchard, 2005; Johnson, te Nijenhuis, & Bouchard, 2007b; Major, Johnson, & Deary, 2012) developed the four-stratum verbal-perceptual-image rotation model, which has an overarching *g*-factor at the apex with verbal, perceptual, and image rotation abilities in the third stratum. In the second stratum, the verbal factor includes verbal, scholastic, fluency, and numerical abilities; the perceptual factor includes content memory, perceptual speed, spatial and (shared with the verbal factor) numerical abilities; the image rotation factor involves mental rotation of three-dimensional images. The first stratum consists of specific tests of the abilities.

> In this implementation, the third-stratum verbal and perceptual abilities, though separable, were highly correlated (.80), as were the perceptual and image rotation abilities (.85). The verbal and image rotation abilities, however, were much less correlated (.41), though g contributed similarly to all of them. (Johnson et al., 2007a, p. 543)

Although Bouchard, Johnson, and colleagues present strong experimental evidence for the preeminence of *g* and strongly dispute models without *g* at the top, Bouchard (2018, pp. 20, 22) stresses that:

> It is important to realize that *g* is not the only mental ability. There are important special abilities. One of the goals of the Minnesota Study of Twins Reared Apart (MISTRA) was to formally test competing models regarding the structure of mental abilities. Advances in confirmatory factor analysis had made clear that it would be possible to pit models against each other in meaningful ways if the appropriate test

> batteries were assembled (Jöreskog, 1969). The MISTRA battery was assembled with this goal in mind and that body of data, gathered over 20 years, has been made publicly available (Johnson & Bouchard, 2011). The process of pitting theories against each other is called "Strong Inference" (Platt, 1964), a practice in short supply in the social sciences (Bouchard, 2009).

Bouchard (e.g., 2018) argues that data, such as studies of identical and nonidentical twins raised together or apart and unrelated children raised together, strongly support the importance of hereditary influences on *g*, and argues that the mechanism is a genetically determined drive toward exploration creating experiences that enhance cognitive ability.

One Influential Synthesis: Cattell, Horn, and Carroll

Spearman (1904) had originally insisted that the separate *s* factors were limited to their particular tests or subtests. Eventually, though, he recognized that some factors were common to multiple measures, making them "group factors," but, unlike *g*, the group factors were not common to all measures (Spearman, 1927). The final version of Spearman's theory with the two factors, one *g* and various *s* and group factors, was closer to Thurstone's formulation than his original theory had been.

At the other end of the continuum, when Thurstone administered his tests to an intellectually heterogeneous group of children, he found that his seven primary abilities were *not* entirely separate; instead he found evidence of a second-order factor that he theorized might be related to *g* (Sattler, 2008, 2018). According to Ruzgis (1994), the final version of Thurstone's theory, which accounted for the presence of both a general factor and the seven specific abilities, helped lay the groundwork for future researchers who proposed hierarchical theories and theories of multiple intelligences. Thurstone's final formulation was closer than his original theoretical framework to Spearman's model. In the end, the two extremes of the lumper-splitter continuum (Spearman and Thurstone) each gravitated a bit toward the center.

Cattell and Horn's Gf-Gc Model

Probably the best-known and most widely accepted theories of intellectual factors derive from the model of Raymond B. Cattell (1941) and his student, John L. Horn (1965, 1994). Cattell first proposed two types of intelligence: G*f* and G*c*, which refer, respectively, to "fluid intelligence" and "crystallized intelligence" (Cattell, 1941, 1963). Cattell and Horn and colleagues (e.g., Cattell & Horn, 1978; Horn, 1965, 1985, 1988; Horn & Blankson, 2012; Horn & Cattell, 1966), drawing on factor-analytic studies and evidence from "neurological damage and aging" and "genetic, environmental, biological, and developmental variables" (Horn & Blankson, 2012, pp. 75–77), gradually expanded this initial bifurcation of *g* into eight or nine primary abilities. Horn (1985, 1994) argued unyieldingly against the reality of a single general ability factor (*g*), because he did not believe that research supported a unitary theory. For another perspective on G*f-Gc* theory, see Woodcock (1993).

*G*f*, fluid intelligence*, refers to inductive, general sequential, and quantitative reasoning with materials and processes that are new to the person doing the reasoning. Fluid abilities allow an individual to think and act quickly, solve novel problems, and encode short-term memories. The vast majority of fluid reasoning tasks on intelligence tests use nonverbal, relatively culture-free stimuli, but require an integration of verbal and nonverbal thinking.

*G*c*, crystallized intelligence*, refers to the application of acquired knowledge and learned skills to answering questions and solving problems presenting at least broadly familiar materials and processes. It is reflected in tests of knowledge, general information, use of language (vocabulary), and a wide variety of acquired skills (Horn & Cattell, 1966). Most verbal subtests of intelligence scales primarily involve crystallized intelligence.

Carroll's Three-Stratum Hierarchy

John B. Carroll (1993) undertook a truly staggering reanalysis of all of the usable correlational studies of mental test data that he could find. He winnowed a collection of about 1,500 studies down to a set of 461 data sets that met four technical criteria (Carroll, 1993, pp. 78–80, 116; 1997/2012, pp. 883–884) and then subjected the data from those studies to a uniform process of reanalysis by exploratory factor analysis (Carroll, 1993, pp. 80–91; 1997/2012, p. 884). Carroll noted that this massive project was "in a sense an outcome of work I started in 1939, when . . . I became aware of L. L. Thurstone's research on what he called 'primary mental abilities' and undertook, in my doctoral dissertation, to apply his factor-analytic techniques to the study of abilities in the domain of language" (1993, p. vii). As a result of his reanalysis of the 461 data sets, Carroll presented extensive data in the domains of language, reasoning, memory and learning, visual perception, auditory reception, idea production, cognitive speed, knowledge and achievement, psychomotor abilities, miscellaneous domains of ability and personal characteristics, and higher-order factors of cognitive ability (1993, p. 5). Based on his data, Carroll (1993, pp. 631–655) presented "A Theory of Cognitive Abilities: The Three-Stratum Theory" with "*narrow* (stratum I), *broad* (stratum II), and *general* (stratum III)" (p. 633) abilities. See Carroll (1997/2012) for further discussion.

Integration of Horn-Cattell and Carroll Models to Form CHC Theory

The remarkable similarity between Carroll's *broad* stratum II abilities and Cattell and Horn's expanded G*f*-G*c* abilities suddenly became apparent at a meeting in March 1996 convened by the publisher of the Woodcock-Johnson Psycho-Educational Battery (Woodcock & Johnson, 1977) to begin the process of developing the Woodcock-Johnson – Revised (Woodcock & Johnson, 1989). Kevin McGrew (2005) describes this "fortuitous" meeting, which included Richard Woodcock, John L. Horn, and John B. Carroll, among other important figures in test theory and development, including McGrew. McGrew considers that meeting the "flash point that resulted in *all* subsequent theory-to-practice bridging events leading to today's CHC theory and related assessment developments" (p. 144).

"CHC" stands for "Cattell-Horn-Carroll," a synthesis of the work of Cattell and Horn with that of Carroll. McGrew (2005, p. 148) believes that the term and abbreviation "Cattell-Horn-Carroll theory" and "CHC" were first published in Flanagan, McGrew, and Ortiz (2000) and first formally defined in print in his and Woodcock's technical manual for the third edition of the Woodcock-Johnson battery (McGrew & Woodcock, 2001). CHC theory synthesizes two of the most widely recognized theories of intellectual abilities (McGrew, 2005; Sternberg & Kaufman, 1998).

Although Horn and Carroll agreed to the use of the term Cattell-Horn-Carroll (McGrew, 2005, p. 149), Horn and Carroll always disagreed sharply about *g* or the general stratum III (McGrew, 2005, p. 174). Horn, like Thurstone in his earlier formulations, consistently and adamantly maintained that there was no single *g*. Carroll always considered *g* or stratum III essential to his hierarchical, three-stratum theory.

Carroll (1993, 1997/2012) wrote, "There are a fairly large number of distinct individual differences in cognitive ability, and . . . the relationships among them can be derived by classifying them into three different strata: stratum I, 'narrow' abilities; stratum II, 'broad' abilities; and stratum III, consisting of a single 'general' ability" (p. 883). Carroll's model, although strikingly similar to that proposed by Cattell and Horn, differs in several substantial ways. First, as noted above, Carroll included at stratum III the general intelligence factor (*g*) because he believed that the evidence for such a factor is overwhelming. Second, where Cattell and Horn differentiate quantitative knowledge as a separate Gf-G*c* factor, labeled G*q*, Carroll believed quantitative ability is best subsumed as a narrow G*f* ability. Third, while the Cattell-Horn model included measures of reading and writing as a combined, separate factor (G*rw*), Carroll believed these to be narrow abilities subsumed in the G*c* factor.

Applications of CHC Theory: Cross-Battery Assessment and Test Development

CHC theory provided the basis for the McGrew, Flanagan, and Ortiz integrated cross-battery approach to assessment (see, for example, Flanagan & McGrew, 1997; Flanagan et al., 2013). Although the CHC cross-battery approach quickly gained many adherents among evaluators, it does not meet with universal approval (see, for example, Floyd, 2002; Naglieri et al., 2012; Ortiz & Flanagan, 2002a, 2002b; Watkins, Youngstrom, & Glutting, 2002).

CHC theory also, to varying degrees, contributed to the structure of many recent tests of cognitive ability. The Woodcock-Johnson (WJ) IV (Schrank, McGrew, & Mather, 2014; see also Schrank, Decker, & Garruto, 2016) and its two previous editions are explicitly based on CHC theory, and the WJ IV provides scores for nine of the most commonly agreed-upon CHC broad (stratum II) abilities and several narrow (stratum II) abilities. Some other cognitive ability tests with explicit, though not exclusive, CHC foundations include the Kaufman Assessment Battery for Children – Second Edition (KABC-II; A. S. Kaufman & Kaufman, 2004; KABC-II/NU; A. S. Kaufman et al., 2018), Differential Ability Scales – Second Edition (DAS-II; Elliott, 2007), and Stanford-Binet Intelligence Scale – Fifth Edition (SB 5; Roid, 2003). CHC or G*f*-G*c* abilities are cited in the test manuals to help explain and describe scales and subtests for many tests, including the Leiter International Performance Scale – Third Edition (Leiter-3; Roid et al., 2013), the Reynolds Intellectual Assessment Scales – Second Edition (RIAS-2; Reynolds & Kamphaus, 2015), and recent editions of the Wechsler intelligence scales, such as the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV; Wechsler, 2008), Wechsler Intelligence Scale for Children – Fifth Edition (WISC-V; Wechsler, 2014), and Wechsler Preschool and Primary Scale of Intelligence – Second Edition (WPPSI-IV; Wechsler, 2012). There is a growing body of research showing relationships between various CHC factors and different aspects of school achievement (e.g., Caemmerer et al., 2018; Cormier et al., 2017; Evans et al., 2002; Flanagan et al., 2013; Floyd, Evans, & McGrew, 2003; Hale et al., 2008; Hale et al., 2001).

Cognitive Abilities: What's in a Name?

CHC theory continues to evolve (Schneider & McGrew, 2019). Complete agreement still has not quite been reached on the broad (stratum II) abilities, and the narrow (stratum I) abilities within each broad ability are occasionally redefined. Current formulations can be found in Flanagan and colleagues (2013), McGrew and colleagues (2014, pp. 2–8, 243–252), and Schneider and McGrew (2019). Those books, and others cited earlier, classify a great many intelligence and achievement test subtests by broad (stratum II) and narrow (stratum I) CHC abilities on the basis of factor-analytic research and surveys of expert opinion. The names and the abbreviations or symbols for the abilities are taken, with alterations, from Carroll, 1993, who observed (p. 644), "The naming of a factor in terms of a process, or the assertion that a given process or component of mental architecture is involved in a factor, can be based only on inferences and makes little if any contribution to explaining or accounting for that process unless clear criteria exist for defining and identifying processes."

Even more broadly, we need to be careful not to confuse verbal names for factors with the factor-analytic bases for them. For example, G*v* has been referred to as, among other things, "visual-spatial thinking," which sounds like a high-level cognitive process, and "visual perception," which sounds more physiological than intellectual. By either name, it is the same G*v*, defined by loadings of various tests or subtests on the same factor, and we should not be distracted, biased, or misled by the verbal name assigned by an author. For example, when Cohen (1959) made a tremendous contribution to the field by publishing his factor analysis of the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1949), he also, we believe, inadvertently caused decades of misunderstanding by assigning the name "freedom from distractibility" to a factor consisting of the Arithmetic, Digit Span, and Coding subtests. Generations of psychologists and educators consequently persisted in the misguided belief that those subtests were definitively diagnostic of attention deficit hyperactivity disorder. A. S. Kaufman (1979) tried to resolve this confusion by neutrally calling his derived score for those three subtests simply "the third factor," but in our personal experience, the misunderstanding remained robust. This cautionary tale might encourage us to take advantage of the more or less implication-free abbreviations and symbols offered by current formulations of CHC theory. The following discussion draws heavily on presentations in Carroll (1993); Flanagan and Dixon (2014); Flanagan and McGrew (1997); Flanagan and colleagues (2013); A. S. Kaufman and colleagues (2018); McGrew (1997, 2018); McGrew and Schneider (2018); Schneider and McGrew (2019).

Definitions of CHC Abilities

Fluid and crystallized intelligence, described earlier, were the original Cattell-Horn G*f*-G*c* factors. As noted, over the years, the original dichotomous G*f*-G*c* theory was expanded to include additional abilities. These additional broad (stratum II) abilities are defined here, together with some of the more recent modifications to CHC theory. "The precise number of broad and narrow abilities depends on each researcher's perspective and preferences" (A. S. Kaufman et al., 2018, p. 5).

*G*kn *or general (domain-specific knowledge)* is the depth, breadth, and mastery of specialized declarative and procedural knowledge typically acquired through one's professional education, career, hobby, or other passionate interest. Examples of narrow abilities might include: foreign language proficiency (KL), which is similar to language development (G*c*-LD) except that it is proficiency in another language. Knowledge of signing (KF) is the knowledge of finger spelling and signing (e.g., American Sign Language). Skill in lip reading (LP) is competence in the ability to understand communication from others by watching the movement of their mouths and their facial expressions. Geography achievement (A5) is the range of geography knowledge (e.g., capitals of countries). General science information (K1) is the range of scientific knowledge (e.g., biology, physics, engineering, mechanics, electronics). Knowledge of culture (K2) is the range of knowledge about the humanities (e.g., philosophy, religion, history, literature, music, and art). Mechanical knowledge (MK) is the range of knowledge about the function, terminology, and operation of ordinary tools, machines, and equipment. Knowledge of behavioral content (BC) is the knowledge of or sensitivity to nonverbal human communication/interaction systems (e.g., facial expressions and gestures).

*G*v*, or visual-spatial thinking*, involves a range of visual processes, ranging from fairly simple visual perceptual tasks to higher-level, visual, cognitive processes. Woodcock and Mather (1989, p. 15) define G*v* in part: "In Horn-Cattell theory, 'broad visualization' requires fluent thinking with stimuli that are visual in the mind's eye." Although G*f* tasks are also often nonverbal (e.g., matrix tests), G*v* does not include the aspect of dealing with novel stimuli or applying novel mental processes that characterizes G*f* tasks. Many writers seem to consider G*v* a relatively low-level cognitive ability, more perceptual than intellectual. However, the "fluent thinking with stimuli that are visual in the mind's eye" may well be a higher-level intellectual process on a par with G*c* and G*f* (see, for example, Johnson & Bouchard, 2005, and Johnson et al., 2007b, who differentiate perceptual from image rotation abilities). Engineers, auto mechanics, architects, nuclear physicists, sculptors, carpenters, and parts department managers all use G*v* to deal with the demands of their jobs. Elliott (2007), for example, made two subtests each of G*f*, G*c*, and G*v* abilities the core subtests for the general conceptual ability summary score for the school-age and upper early years levels of the Differential Ability Scales – Second Edition. Other CHC abilities are included among the diagnostic subtests, but are not counted in the general conceptual ability score. Schneider and McGrew (2019) have proposed or maintained the additional narrow abilities of visualization (Vz), imagery (Im), visual memory (Mv), and spatial scanning (SS).

*G*a*, auditory processing*, involves tasks such as recognizing similarities and differences between sounds; recognizing degraded spoken words, such as words with sounds omitted or separated (e.g., "tel – own" and /t/ /ĕ/ /l/ /ĕ/ /f/ /ō/ /n/ both as "telephone"), and mentally manipulating sounds in spoken words (e.g., "say *blend* without the /l/ sound" or "change the /ĕ/ in *blend* to /ī/"). Phonemic awareness skills, terribly important for acquisition of reading skills (e.g., Carroll, 1997/2012, p. 888; Farrall, 2012, pp. 163–192; Kilpatrick, 2015), are G*a* tasks. Schneider and McGrew (2019) include the narrow abilities of phonetic coding (PC) to G*a,* with subcategories of speech sound discrimination (US), resistance to auditory distraction (UR), maintaining judgment and rhythm (U8), and memory for sound patterns (UM).

*G*s*, processing speed* or attentional speediness, refers to measures of clerical speed and accuracy, especially when there is pressure to maintain focused attention and concentration. Most recently, Schneider and McGrew (2019) have added or maintained the narrow abilities of perceptual speed (P), with subcategories of perceptual speed-search (Ps) and perceptual speed-compare (Pc), together with the narrow abilities of number facility (N), reading speed (fluency) (RS), and writing speed (fluency) (WS).

*G*t*, decision/reaction time or speed*, reflects the immediacy (quickness) with which an individual can react and make a decision (decision speed) to typically simple stimuli. It can be difficult to distinguish between G*s* tasks, which are relatively common on intelligence tests, and G*t* tasks, which are more often found on computerized neuropsychological measures of vigilance and reaction time. G*s* tasks generally require a sustained effort over at least two or three minutes and simply measure the number of simple items completed (or number right minus number wrong) for the entire span of time. G*t* tasks are more likely to measure response speed to each item or a few items. Schneider and McGrew (2019) note additional narrow abilities here that include simple reaction time (R1), choice reaction time (R2), inspection time (IT), semantic processing speed (R4), and mental comparison speed (R7).

*G*sm*, short-term or immediate memory*, or G*wm* (short-term working memory), refers to the ability to take in and hold information in immediate memory and then to use it within a few seconds. Given the relatively small amount of information that can be held in short-term memory, information is typically retained for only a short period of time before it is lost. When additional tasks are required that tax an individual's short-term memory abilities, information in short-term memory is either lost or transferred and stored as acquired knowledge through the use of long-term storage and retrieval (G*lr*). G*sm* is divided in current CHC formulations into memory span (MS) and working memory capacity (MW) with a distinction between simple recall (MS) (e.g., repeating increasingly long series of dictated digits) and mental manipulation of material held in short-term memory (MW) (e.g., repeating the dictated series in reversed sequence). This is another example of the difficulty with verbal labels for abilities, since "working memory" is used by many authors to mean not MW but MS, particularly with reference to brief retention on the way to long-term storage. The different meanings of the terms can cause considerable confusion. Factor analyses suggested that short-term visual memory (such as recognizing in a group of pictures the one picture that had been seen earlier) is a narrow ability within G*v* rather than G*sm*. However, more recently Schneider and McGrew (2019) have added Wa (auditory short-term storage), Wv (visual-spatial short-term storage), and AC (attentional control) to G*wm*.

*G*lr*, long-term storage and retrieval*, involves memory storage and retrieval over longer periods of time than G*sm*. How much longer varies from task to task. It is important to note that G*lr* refers to the *efficiency* of storage and retrieval, not *what* is stored (that is Gc). G*lr* is usually measured with controlled learning tasks in which the efficiency of learning and recalling, for example, rebus symbols for words, is assessed during the learning, and then, on some tests, retention is assessed with a delayed recall measure. Schneider and McGrew (2019) have divided G*lr* into learning efficiency (G*l*) and retrieval fluency (G*r*). Learning efficiency includes MA (associative memory), MM (meaningful memory), and free-recall memory (M6), while retrieval fluency (G*r*) includes several narrow abilities related to speed and ease of retrieval. Although the distinction between ease of learning (storage) and fluency of retrieval fits clinical and research findings, many existing tests assess both broad abilities in a single score.

*G*rw *includes reading and writing abilities*, which were part of G*c* in Carroll's (1993) formulation. Some authorities (e.g., Flanagan et al., 2013) divide G*rw* into G*rw*-R (reading) and G*rw*-W (writing) with narrow abilities within each, but even those several narrow, stratum I abilities may not be sufficiently detailed to satisfy educators specializing in literacy (Farrall, 2012; Kilpatrick, 2015).

*G*q*, quantitative knowledge*, is distinct from quantitative *reasoning*, which is a narrow ability (RQ) within G*f*. G*q* refers to the depth and breadth of declarative and procedural knowledge related to mathematics. According to Schneider and McGrew (2019), G*q* also contains the narrow abilities mathematical knowledge (KM) and mathematical achievement (A3).

The last two broad abilities (G*rw* and G*q*) raise the question of the distinction between "ability" and "achievement." Carroll (1993, p. 510, emphasis in original) discusses this problem: "It is hard to draw the line between factors of cognitive abilities and factors of achievement. Some will argue that *all* cognitive abilities are in reality learned achievements of one kind or another." Carroll suggests we "conceptualize a continuum that extends from the most general abilities to the most specialized types of knowledges." Flanagan and colleagues (2002, p. 21) quote those comments from Carroll (1993, p. 510) and then also quote Horn: "Cognitive abilities are measures of achievements, and measures of achievements are just as surely measures of cognitive ability" (Horn, 1988, p. 665). They reach the same conclusion as Carroll: "Thus, rather than conceiving of cognitive abilities and academic achievements as mutually exclusive, they may be better thought of as lying on an *ability continuum* that has the most general types of abilities at one end and the most specialized types of knowledge at the other (Carroll, 1993)" (Flanagan et al., 2002, p. 21). See Niileksa and colleagues (2016) and Cormier and colleagues (2017) for a discussion of ability and achievement with the WJ-IV and Caemmerer and colleagues (2018) with the WISC-V and WIAT-III.

Additional Abilities

Space does not permit detailed discussions of several additional abilities, which are listed below. See Flanagan and Dixon (2014), Flanagan and colleagues (2013), and Schneider and McGrew (2019) for recent discussions of these broad abilities, which are usually assessed in neuropsychological rather than cognitive evaluations, although some of these abilities may affect performance on cognitive tests.

Olfactory abilities (G*o*): The ability to detect and process meaningful information in odors. Perceiving, discriminating, and manipulating odors. The G*o* domain is likely to contain more narrow abilities than are currently listed in the CHC model. Olfactory memory (OM) is the ability to recognize previously encountered distinctive odors.

Tactile (haptic) abilities (G*h*): The ability to detect and process meaningful information in haptic (touch) sensations, perceiving, discriminating, and manipulating touch stimuli. Currently there are no well-supported narrow G*h* cognitive ability factors, although tests of these abilities are used by neuropsychologists and occupational therapists.

Kinesthetic abilities (G*k*): The ability to detect and process meaningful information in proprioceptive sensations, perceiving, discriminating, and manipulating sensations of body movement. Currently there are no well-supported narrow G*k* cognitive ability factors.

Psychomotor abilities (G*p*): The ability to perform skilled physical body motor movements (e.g., movement of fingers, hands, legs) with precision, coordination, or strength. The G*p* domain is likely to contain more narrow abilities than are currently listed in the CHC model. Aiming (AI) is the ability to precisely and fluently execute a sequence of eye-hand coordination movements for positioning purposes. Manual dexterity (P1) is the ability to make precisely coordinated movements of a hand or a hand and attached arm. Finger dexterity (P2) is the ability to make precisely coordinated movements of the fingers (with or without the manipulation of objects). Static strength (P3) is the ability to exert muscular force to move (push, lift, pull) a relatively heavy or immobile object. Gross body equilibrium (P4) is the ability to maintain the body in an upright position in space or to regain balance after balance has been disturbed. Multi-limb coordination (P6) is the ability to make quick, specific, or discrete motor movements of the arms or legs. Arm-hand steadiness (P7) is the ability to precisely and skillfully coordinate arm-hand positioning in space. Control precision (P8) is the ability to exert precise control over muscle movements, typically in response to environmental feedback.

Other Formulations

Although they are slightly or substantially outside the factor-analytic focus of this chapter, there are other important and influential theories and models that bear mention and that share the concept of dividing cognitive abilities into categories.

Planning, Attention, Simultaneous, Successive (PASS)

Building on the work of Russian psychologist A. R. Luria (1966, 1973, 1980), J. P. Das, Jack Naglieri, and colleagues (e.g., Das, Kirby, & Jarman, 1979; Das, Naglieri, & Kirby, 1994; Naglieri & Das, 2002; Naglieri et al., 2012; Naglieri & Otero, 2019a) have developed the Planning, Attention, Simultaneous, Successive (PASS) theory of intelligence. Luria posited three functional units or "blocks": arousal and attention (the attention in PASS), representing Luria's Block 1; taking in, processing, and storing information (the simultaneous and successive processes in PASS), or Block 2 coding processes; and synthesizing information and regulating behavior (the planning in PASS), which are the executive functions associated with Block 3.

The Kaufman Assessment Battery for Children (K-ABC; A. S. Kaufman & Kaufman, 1983; A. S. Kaufman, Kaufman, & Goldsmith, 1984) was a pioneering test based on simultaneous versus sequential (successive) processing, the components of Luria's second processing unit (Block 2). The second edition of the Kaufman Assessment Battery for Children (KABC-II; A. S. Kaufman & Kaufman, 2004; KABC-II/NU; A. S. Kaufman et al., 2018; see also Drozdick et al., 2019; A. S. Kaufman, 2018) is uniquely designed to permit interpretation on the basis of four Luria-based processes or on the basis of five CHC factors: sequential processing or G*sm*, simultaneous processing or G*v*, learning or G*lr*, planning or G*f*, and G*c*.

Naglieri, Das, and Goldstein's (2014) Cognitive Assessment System – Second Edition (CAS2) is built on the planning, attention, simultaneous, and successive (PASS) theory; there are three planning, three attention, three simultaneous, and four successive subtests (Naglieri & Otero, 2019a, 2019b).

As with CHC theory, there is evidence of correlations of PASS measures with different aspects of educational achievement. There is also evidence of the utility of PASS profiles for planning instruction (e.g., Naglieri & Johnson, 2000; Naglieri & Otero, 2019a). Differences between scores of African-American and Euro-American students are notably smaller on the PASS-based CAS2 and KABC-II than on other comprehensive cognitive ability tests in current use (Drozdick et al., 2019; A. S. Kaufman, 2018; Naglieri et al., 2014; Naglieri & Otero, 2019b).

Triarchic Theory

Many experts argue, for various reasons, that none of the theories discussed above goes far enough (e.g., Chen & Gardner, 2019; Gardner, 2003; J. C. Kaufman, 2015; Plucker, Karwowski, & Kaufman, Chapter 45 in this volume; Stanovich, 2009; Stanovich, Toplak, & West, Chapter 46 in this volume).

Robert J. Sternberg (1982, 1985, 1997, 1999, 2002, 2003, 2005, 2019, Chapter 28 in this volume) has developed the theory of:

> successful intelligence [which] is (1) the use of an integrated set of abilities needed to attain success in life, however an individual defines it, within his or her sociocultural context. People are successfully intelligent by virtue of (2) recognizing their strengths and making the most of them, at the same time that they recognize their weaknesses and find ways to correct or compensate for them. Successfully intelligent people (3) adapt to, shape, and select environments through (4) finding a balance in their use of analytical, creative, and practical abilities. (Sternberg, 2005, p. 104)

Although not strictly speaking a factor-analytic theory of intelligence, Sternberg's theory is supported by studies showing the "factorial separability of analytic, creative, and practical abilities" (Sternberg, 2005, pp. 104–105). Sternberg also points to evidence of effective instructional interventions based on the theory. The triarchic theory of successful human intelligence expands considerably the domain of intelligence beyond what is measured by most current tests. We believe that Sternberg's theory comes much closer to Wechsler's famous definition of intelligence ("the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment"; Wechsler, 1958, p. 7) than do any of Wechsler's own intelligence tests.

Multiple Intelligences

Gardner (1983, 1994, 1999, 2003; Chen & Gardner, 2005, 2019) argues for the existence of at least eight "intelligences," including linguistic, logical-mathematical, musical, spatial, bodily-kinesthetic, naturalistic, interpersonal, and intrapersonal. "The identification of intelligences is based on empirical evidence and can be revised on the basis of new empirical findings" (Chen & Gardner, 2005, p. 79).

Rationality

Stanovich (2009) agrees with Sternberg and Gardner that the aspects of intelligence measured by traditional tests, which he terms "MAMBIT (to stand for: the mental abilities measured by intelligence tests)" (p. 13), are too narrow. He focuses particularly on the absence of measures of rational thinking (e.g., Stanovich, Toplak, & West, Chapter 46 in this volume; Sternberg, 2002). However, rather than including rational thinking and other abilities in a definition of "intelligence," Stanovich argues for separating MAMBIT from other abilities, such as rational decision-making.

A Parting Thought

Factor-based theories of intelligence have proliferated since Spearman (1904) started the ball rolling more than a century ago. The once-extreme "lumper-splitter" dichotomy has become less extreme and the pendulum has rested somewhere in between the two ends, though decidedly closer to the Thurstone than the Spearman end. The uneasy balance between *g* and multiple abilities is probably best reflected by CHC theory, which reflects an integration of the life's work of John B. Carroll (a believer in *g*) and John L. Horn (a devout nonbeliever), and forms the foundation of most contemporary IQ tests. We believe that CHC theory has important positive features and merits a key role in the assessment of intelligence. But, however well researched CHC theory may be, it reflects only one-third of Sternberg's theory, and perhaps a similar portion of Gardner's theory – but, as Stanovich aptly points out, MAMBIT is too narrow. At present, CHC theory and, to a lesser extent, Luria's neuropsychological theory provide the theoretical basis of virtually all major tests of cognitive abilities. It is time for that status quo to change. The time has come for developers of individual clinical tests of intelligence to broaden their basis of test construction beyond the analytic dimension of Sternberg's triarchic theory and to begin to embrace the assessment of both practical intelligence and creativity.

References

Benisz, M., Dumont, R., & Willis, J. O. (2015). From psychometric testing to clinical assessment: Personalities, ideas and events that shaped David Wechsler's views of intelligence and its assessment. In S. Goldstein, D. Princiotta, & J. A. Naglieri (Eds.), *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts* (pp. 163–179). New York: Springer.

Bouchard, T., Jr. (2018). Hereditary ability: *g* is driven by experience-producing drives. In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 15–29). New York: Cambridge University Press.

Binet, A., & Simon, T. (1916/1980). *The development of intelligence in children* (trans. E. S. Kite). Facsimile limited edition issued by Lloyd M. Dunn. Nashville, TN: Williams Printing Co.

Caemmerer, J. J., Maddocks, D. L., Keith, T. Z., & Reynolds, M. R. (2018). Effects of cognitive abilities on child and youth academic achievement: Evidence from the WISC-V and WIAT-III. *Intelligence*, *68*, 6–20. https://doi.org/10.1016/j.intell.2018.02.005

Canivez, G. L., & Watkins, M. W. (2016). Review of the Wechsler Intelligence Scale for Children – Fifth Edition: Critique, commentary, and independent analyses. In A. S. Kaufman, S. E. Raiford, & D. L. Coalson (Eds.), *Intelligent testing with the WISC-V* (pp. 684–702). Hoboken, NJ: John Wiley & Sons.

Carroll, J. B. (1985). Exploratory factor analysis: A tutorial. In D. K. Detterman (Ed.), C*urrent topics in human intelligence* (vol. 1, pp. 25–58). Norwood, NJ: Ablex.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Carroll, J. B. (1997/2012). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 883–890). New York: Guilford Press.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin*, *38*, 592.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1–22. http://dx.doi.org/10.1037/h0046743

Cattell, R. B., & Horn, J. L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, *15*, 139–164. https://doi.org/10.1111/j.1745-3984.1978.tb00065.x

Chen, J-Q., & Gardner, H. (2005). Assessment based on multiple-intelligence theories. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 77–102). New York: Guilford Press.

Chen, J-Q., & Gardner, H. (2019). Assessment from the perspective of multiple-intelligence theory. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 164–173). New York: Guilford Press.

Cohen, J. (1959). The factorial structure of the WISC at ages 7–6, 10–6, and 13–6. *Journal of Consulting Psychology*, *23*, 285–299. http://dx.doi.org/10.1037/h0043898

Cormier, D. C., McGrew, K. S., Bulut, O., & Funamoto, A. (2017). Revisiting the relations between the WJ-IV measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during the school-age years. *Journal of Psychoeducational Assessment*, *35*(8), 731–754. https://doi.org/10.1177/0734282916659208

Daniel, M. H. (1997). Intelligence testing: Status and trends. *American Psychologist*, *52*(10), 1038–1045.

Das, J. P., Kirby, J. R., & Jarman, R. F. (1979). *Simultaneous and successive cognitive processes*. New York: Academic Press.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Needham Heights, MA: Allyn & Bacon.

Drozdick, L. W., Singer, J. K., Lichtenberger, E. O., Kaufman, J. C., Kaufman, A. S., & Kaufman, N. L. (2019). The Kaufman Assessment Battery for Children – Second Edition and KABC-II Normative Update. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 333–359). New York: Guilford Press.

Elliott, C. D. (2007). *Differential ability scales – Second edition*. San Antonio, TX: Pearson.

Evans, J. J., Floyd, R. G., McGrew, K. S., & Leforgee, M. H. (2002). The relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review*, *31*(2), 246–262. https://doi.org/10.1002/pits.10083

Everybody Loves Raymond. Standard Deviation. Directed by J. Meyer. Written by P. Rosenthal & S. Skrovan. Where's Lunch, Worldwide Pants, & HBO Independent Productions (October 4, 1996). www.imdb.com/title/tt0574226

Eysenck, H. J., versus Kamin, L. J. (1981). *The intelligence controversy.* Hoboken, NJ: Wiley-Interscience.

Farrall, M. L. (2012). *Reading assessment: Linking language, literacy, and cognition*. Hoboken, NJ: John Wiley & Sons.

Flanagan, D. P., & Dixon, S. G. (2014). The Cattell-Horn-Carroll theory of cognitive abilities. In Reynolds, C. R., Vannest, K. J., & Fletcher-Janzen, E. (Eds.), *Encyclopedia of special education* (4th ed.). Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781118660584.ese0431

Flanagan, D. P., & Kaufman, A. S. (2009). *Essentials of WISC-IV assessment* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Flanagan, D. P., & McGrew, K. S. (1997). A cross-battery approach to assessing and interpreting cognitive abilities: Narrowing the gap between practice and cognitive science. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (pp. 314–325). New York: Guilford Press.

Flanagan, D. P., McGrew, K. S., & Ortiz, S. O. (2000). *The Wechsler Intelligence Scales and Gf-Gc theory: A contemporary approach to interpretation*. Boston: Allyn & Bacon.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. (2013). *Essentials of cross-battery assessment* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Flanagan, D. P., Ortiz, S. O., Alfonso, V., & Mascolo, J. T. (2002). *The achievement test desk reference: Comprehensive assessment of learning disabilities*. Boston: Allyn & Bacon.

Floyd, R. (2002). The Cattell-Horn-Carroll (CHC) Cross-Battery Approach: Recommendations for school psychologists. *Communiqué*, *30*(5), 10–14.

Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and mathematics achievement across the school-age years. *Psychology in the Schools*, *60*(2), 155–171. https://doi.org/10.1002/pits.10083

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. New York: Cambridge University Press.

Gardner, H. (1983). *Frames of mind*. New York: Basic Books.

Gardner, H. (1994). Multiple intelligences theory. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 740–742). New York: Macmillan.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.

Gardner, H. (2003). Multiple intelligences after twenty years. Paper presented at the annual meeting of the American Education Research Association, Chicago, April.

Gottfredson, L. S. (2008). Of what value is intelligence? In A. Prifitera, D. Saklofske, & L. G. Weiss (Eds.), *WISC-IV applications for clinical assessment and intervention* (2nd ed., pp. 545–563). Amsterdam, Netherlands: Elsevier.

Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Hale, J. B., Fiorello, C. A., Dumont, R., Willis, J. O., Rackley, C., & Elliott, C. D. (2008). *Differential Ability Scales – Second Edition* (neuro)psychological predictors of math performance for typical children and children with math disabilities. *Psychology in the Schools*, *45*(9). https://doi.org/10.1002/pits.20330

Hale, J. B., Fiorello, C. A., Kavanagh, J. A., Hoeppner, J. B., & Gaitherer, R. A. (2001). WISC-III predictors of academic achievement for children with learning disabilities: Are global and factor scores comparable? *School Psychology Quarterly*, *16*(1), 31–35. http://dx.doi.org/10.1521/scpq.16.1.31.19158

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Simon & Schuster (Free Press Paperbacks).

Horn, J. L. (1965). Fluid and crystallized intelligence: A factor analytic study of the structure among primary mental abilities, PhD thesis, University of Illinois.

Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 267–300). Hoboken, NJ: John Wiley & Sons.

Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate psychology* (rev. ed., pp. 645–685). New York: Academic Press.

Horn, J. L. (1994). The theory of fluid and crystallized intelligence. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 433–451). New York: Macmillan.

Horn, J. L., & Blankson, A. N. (2012). Foundations for a better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 73–98). New York: Guilford Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*, 253–270.

Huxley, T. 1870. Biogenesis and abiogenesis. *Nature*, *2*, 400–406.

Jacoby, R., & Glauberman, N. (Eds.) (1995). *The bell curve debate*. New York: Times Books.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

Johnson, W., & Bouchard, T., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416. https://doi.org/10.1016/j.intell.2004.12.002

Johnson, W., Bouchard, T., Jr., McGue, M., Segal, N. L.,Tellegen, A., Keyes, M., & Gottesman, I. I. (2007a). Genetic and environmental influences on the Verbal-Perceptual-Image Rotation (VPR) model of the structure of mental abilities in the Minnesota study of twins reared apart. *Intelligence*, *35*, 542–562. https://doi.org/10.1016/j.intell.2006.10.003

Johnson, W., te Nijenhuis, J., & Bouchard, T., Jr. (2007b). Replication of the hierarchical visual-perceptual-image rotation model in de Wolff and Buiten's (1963) battery of 46 tests of mental ability. *Intelligence*, *35*, 69–81. https://doi.org/10.1016/j.intell.2006.05.002

Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. Hoboken, NJ: John Wiley & Sons.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. Hoboken, NJ: John Wiley & Sons.

Kaufman, A. S. (2009). *IQ testing 101*. New York: Springer Publishing.

Kaufman, A. S. (2018). Many pathways, one destination: IQ tests, intelligent testing, and the continual push for more equitable assessments. In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 197–214). New York: Cambridge University Press.

Kaufman, A. S., & Kaufman, N. L. (1983). *The Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004). *The Kaufman Assessment Battery for Children – Second Edition*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., Kaufman, N. L., Drozdick, L. W., & Morrison, J. (2018). *Kaufman Assessment Battery for Children – Second Edition, normative update, manual supplement*. Bloomington, MN: NCS Pearson.

Kaufman, A. S., Kaufman, N. L., & Goldsmith, B. Z. (1984). *Kaufman Sequential or Simultaneous (K-SOS)?* Circle Pines, MN: American Guidance Service.

Kaufman, A. S., Raiford, S. E., & Coalson, D. L. (2016). *Intelligent testing with the WISC-V.* Hoboken, NJ: John Wiley & Sons.

Kaufman, A. S., Schneider, W. J., & Kaufman, J. C. (in press). Psychometric approaches to intelligence. In R. J. Sternberg (Ed.), *Human intelligence: An introduction*. New York: Cambridge University Press.

Kaufman, J. C. (2015). Why creativity isn't in IQ tests, why it matters, and why it won't change anytime soon . . . Probably. *Journal of Intelligence*, *3*, 59–72. https://doi.org/10.3390/jintelligence3030059

Keith, T. Z., & Reynolds, M. R. (2019). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 853–900). New York: Guilford Press.

Kilpatrick, D. A. (2015). *Essentials of assessing, preventing, and overcoming reading difficulties*. Hoboken, NJ: John Wiley & Sons.

Kipling, R. 1899. An interview with Mark Twain. In *Sea to sea: Letters of travel*. New York: Doubleday & McClure.

Lichtenberger, E. O., & Kaufman, A. S. (2013). *Essentials of WAIS-IV assessment* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Luria, A. R. (1966). *Human brain and psychological processes*. New York: Harper & Row.

Luria, A. R. (1973). *The working brain*. New York: Basic Books.

Luria, A. R. (1980). *Higher cortical functions in man* (2nd ed.). New York: Basic Books.

Major, J. T., Johnson, W., & Deary, I. J. (2012). Comparing models of intelligence in Project TALENT: The VPR model fits better than the CHC and extended Gf–Gc models. *Intelligence*, *40*, 543–559. https://doi.org/10.1016/j.intell.2012.07.006

McDermott, P. A., Fantuzzo, J. W., & Glutting, J. J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. *Journal of Psychoeducational Assessment*, *8*, 290–302.

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 92–104). New York: Guilford.

McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (2nd ed., pp. 136–181). New York: Guilford Press.

McGrew, K. S. (2018). *WJ IV norm-based and supplemental clinical test groupings for "intelligent" intelligence testing with the WJ IV* (MindHub Pub #3) [PowerPoint presentation]. www.iqscorner.com/2018/05/mindhub-pub-3-wj-iv-norm-based-and_16.html

McGrew, K. S., LaForte, E. M., & Schrank, F. A. (2014). *Technical manual. Woodcock-Johnson IV.* Rolling Meadows, IL: Riverside.

McGrew, K. S., & Schneider, W. J. (2018). *CHC theory revised: A visual-graphic summary of Schneider and McGrew's 2019 CHC update chapter* (MindHub Pub #4) [PowerPoint presentation]. www.iapsych.com/mindhubpub4.pdf

McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual. Woodcock-Johnson III*. Itasca, IL: Riverside Publishing.

McKusick, V. A. (1969). On lumpers and splitters, or the nosology of genetic disease. *Perspectives in Biology and Medicine*, *12*(2), 298–312.

Naglieri, J. A., & Das, J. P. (2002). Practical implications of general intelligence and PASS cognitive processes. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 855–884). New York: Erlbaum.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2012). Planning, attention, simultaneous, successive: A cognitive-processing-based theory of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 178–194). New York: Guilford Press.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2014). *Cognitive Assessment System – Second Edition*. Austin, TX: Pro-Ed.

Naglieri, J. A., & Johnson, D. (2000). Effectiveness of a cognitive strategy intervention to improve math calculation based on the PASS theory. *Journal of Learning Disabilities*, *33*, 591–597. https://doi.org/10.1177/002221940003300607

Naglieri, J. A., & Otero, T. M. (2019a). Redefining intelligence with the planning, attention, simultaneous, and successive theory of neurocognitive processes. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 195–218). New York: Guilford Press.

Naglieri, J. A., & Otero, T. M. (2019b). The Cognitive Assessment System – Second Edition. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 452–485). New York: Guilford Press.

Niileksala, C. R., Reynolds, M. R., Keith, T. Z., & McGrew, K. S. (2016). A special validity study of the Woodcock-Johnson IV: Acting on evidence for specific abilities. In D. P. Flanagan & V. C. Alfonso (Eds.), *WJ-V clinical use and interpretation: Scientist-practitioner perspectives*. Boston: Elsevier Press.

Ortiz, S. O., & Flanagan, D. P. (2002a). Cross-Battery Assessment revisited: Some cautions concerning "Some Cautions" (Part I). *Communiqué*, *30*(7), 32–34.

Ortiz, S. O., & Flanagan, D. P. (2002b). Cross-Battery Assessment revisited: Some cautions concerning "Some Cautions" (Part II). *Communiqué*, *30*(8), 36–38.

Prokosch, M. D., Yeo, R. A., & Miller, G. F. (2005). Intelligence tests with higher *g*-loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence*, *33*, 203–213. https://doi.org/10.1016/j.intell.2004.07.007

Rapaport, D., Gill, M., & Schafer, R. (1945–1946). *Diagnostic psychological testing* (2 vols.). Chicago: Year Book Medical.

Reynolds, C. R., & Kamphaus, R. W. (2015). *Reynolds Intellectual Assessment Scales, Second Edition*. Lutz, FL: Psychological Assessment Resources.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales, Fifth Edition: Technical manual*. Itasca, IL: Riverside Publishing.

Roid, G. H., Miller, L. J., Pomplun, M., & Koch, C. (2013). *Leiter International Performance Scale – Third Edition*. Wood Dale, IL: Stoelting.

Ruzgis, P. (1994). Thurstone, L. L. (1887–1955). In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 1081–1084). New York: Macmillan.

Sacks, O. (1970). *The man who mistook his wife for a hat and other clinical tales*. New York: Simon & Schuster.

Sattler, J. M. (2008). *Assessment of children: Cognitive foundations* (5th ed.). San Diego, CA: Jerome M. Sattler.

Sattler, J. M. (2018). *Assessment of children: Cognitive foundations* (6th ed.). San Diego, CA: Jerome M. Sattler.

Schneider, W. J., & McGrew, K. S. (2019). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 73–163). New York: Guilford Press.

Schrank, F. A., Decker, S. L., & Garruto, J. M. (2016). *Essentials of WJ IV cognitive abilities assessment*. Hoboken, NJ: Wiley.

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock-Johnson IV*. Rolling Meadows, IL: Riverside.

Silventoinen, K., Posthuma, D., van Beijsterveldt, T., Bartels, M., & Boomsma, D. I. (2006). Genetic contributions to the association between height and intelligence: Evidence from Dutch twin data from childhood to middle age. *Genes, Brain and Behavior*, *5*(8), 585–595. https://doi.org/10.1111/j.1601-183X.2006.00208.x

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*, 201–293. www.jstor.org/stable/1412107

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.

Spencer, H. (1855). *The principles of psychology*. London: Longman, Brown, Green, & Longmans.

Springer, S. P., & Deutsch, G. (1993). *Left brain, right brain* (4th ed.). San Francisco: Freeman.

Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT: Yale University Press.

Sternberg, R. J. (1982). Reasoning, problem solving, and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 225–307). New York: Cambridge University Press.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (1986). Intelligence is mental self-government. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence?: Contemporary viewpoints on its nature and definition* (pp. 141–148). Norwood, NJ: Ablex.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.

Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology*, *3*, 292–316. http://doi.org/10.1037/1089-2680.3.4.292

Sternberg, R. J. (2002). *Why smart people can be so stupid*. New Haven, CT: Yale University Press.

Sternberg, R. J. (2003). Construct validity of the theory of successful intelligence. In R. J. Sternberg, J. Lautrey, & T. I. Lubart (Eds.), *Models of intelligence: International perspectives* (pp. 55–80). Washington: American Psychological Association.

Sternberg, R. J. (2005). *The triarchic theory of successful intelligence*. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (2nd ed., pp. 103–119). New York: Guilford Press.

Sternberg, R. J. (2019). The triarchic theory of successful intelligence. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 174–194). New York: Guilford Press.

Sternberg, R. J., & Detterman, D. K. (1986). *What is intelligence?: Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex.

Sternberg, R. J., & Kaufman, J. C. (1998). Human abilities. *Annual Review of Psychology*, *49*, 1134–1139. www.scribd.com/doc/185741878/Sternberg-Kaufman-1998-AnnuRevPsych-Human-Abilities

Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.

Thorndike, E. L. (1927). *The measurement of intelligence*. New York: Bureau of Publications, Teachers College, Columbia University.

Thurstone, L. L. (1931). Multiple factor analysis. *Psychological Review*, *38*(5), 406–427.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen.

Vernon, P. E. (1961). *The structure of human abilities* (2nd ed.). London: Methuen.

Wasserman, J. D. (2019). A history of intelligence assessment: The unfinished tapestry. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (4th ed., pp. 3–55). New York: Guilford Press.

Watkins, M. W., Glutting, J. J., & Youngstrom, E. A. (2005). Issues in subtest profile analysis. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (2nd ed., pp. 251–268). New York: Guilford Press.

Watkins, M. W., Youngstrom, E. A., & Glutting, J. J. (2002). Some cautions regarding Cross-Battery Assessment. *Communiqué*, *30*(6), 16–20.

Wechsler, D. (1926). On the influence of education on intelligence as measured by the Binet-Simon tests. *Journal of Educational Psychology*, *17*, 248–257.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1940). Non-intellective factors in general intelligence. *Psychological Bulletin*, *37*, 444–445.

Wechsler, D. (1943). Non-intellective factors in general intelligence. *Journal of Abnormal and Social Psychology*, *38*, 101–103. http://dx.doi.org/10.1037/h0060613

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.

Wechsler, D. (1950). Cognitive, conative, and non-intellective intelligence. *American Psychologist*, *5*, 78–83. http://dx.doi.org/10.1037/h0063112

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale – Fourth Edition*. Bloomington, MN: Pearson.

Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence – Fourth Edition*. Bloomington, MN: Pearson.

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children – Fifth Edition*. Bloomington, MN: Pearson.

Woodcock, R. W. (1993). An information processing view of Gf-Gc theory. *Journal of Psychoeducational Assessment Monograph Series, Advances in Psychoeducational Assessment: Woodcock-Johnson Psychoeducational Battery – Revised*, 80–102.

Woodcock, R. W., & Johnson, M. B. (1977). *Woodcock-Johnson Psycho-Educational Battery.* Hingham, MA: DLM Teaching Resources.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psycho-Educational Battery – Revised*. Allen, TX: DLM Teaching Resources.

Woodcock, R. W., & Mather, N. (1989). WJ-R tests of cognitive ability – standard and supplemental batteries: Examiner's manual. R. W. Woodcock & M. B. Johnson, *Woodcock-Johnson Psychoeducational Battery – Revised*. Chicago: Riverside Publishing.

Zimmerman, I. L., & Woo-Sam, J. M. (1973). *Clinical interpretation of the Wechsler Adult Intelligence Scale*. New York: Grune & Stratton.

PART II

Development of Intelligence

6 Genetic Bases of Intelligence

Elena L. Grigorenko and Olga Burenkova

As many chapters do, this one started with a literature search. In a combined search of Ovid MEDLINE® All and PsychINFO, the term "intelligence" generated 176,700 entries. In comparison, the word "personality" generated 372,632 entries, more than twice the number of the entries on intelligence. Cross-referencing the terms "intelligence" and "genetic" returned 8,926 entries (i.e., 5% of all intelligence entries); cross-references of the terms "personality" and "genetic" generated 11,338 entries (i.e., 3% of all personality entries). Limiting both subsets to the period from 2010 to *current* returned 3,584 (i.e., 2% of all entries) for intelligence and 3,512 (i.e., 1% of all entries) for personality. Thus, the search started with twice the number of entries on personality compared with that on intelligence and ended with twice the number of entries on intelligence and genetics, compared with that on personality and genetics. For better or for worse, the studies on the genetics of intelligence are overrepresented in the field, given the base rates of studies of intelligence and personality in these two databases.

Clearly, the literature captured by cross-referencing the terms "intelligence" and "genetic" is vast; this topic has captured (and still does!) the attention of many researchers. It is simply impossible to provide a comprehensive overview of this literature in a single chapter. What is possible is to highlight the major threads of this literature. In this chapter, we shall strive to offer a few snapshots of the current themes of this vast literature. We will take these snapshots from three angles: (1) the trajectory of studies that have sought to elucidate the etiology of intelligence; (2) the relevance of the selected phenotype (i.e., an observable behavioral trait such as intelligence); and (3) the selection of a particular genetic mechanism as a factor for the genetic bases of intelligence.

The Trajectory of Studies into the Etiology of Intelligence

The trajectory and current status of the inquiry into the genetic bases of intelligence have been captured in a number of recent reviews (Deary, Harris, & Hill, 2019; Hill, Harris, & Deary, 2019; Plomin & von Stumm, 2018). Briefly, from the

The preparation of this chapter was supported by grant No. 14.Z50.31.0027 (PI: Elena L. Grigorenko) from the Government of the Russian Federation. The chapter does not necessarily reflect the position or policies of this funding agency and no official endorsement should be inferred. We are grateful to Mei Tan of the University of Houston for her editorial support.

emergence of the first instruments for the assessment of intelligence, there has been a notion that intelligence scores are distributed in the general population nonindependently and cluster in families (Galton, 1869). As the methodological arsenal available to research on the genetic bases of intelligence has grown in size and sophistication, those early observations have become better qualified and quantified, and have grown exponentially in number. Considered together, these studies can be subdivided into three major categories.

The first category is exemplified by quantitative genetic studies that capitalized on the presence of different types of relatives in the general population. The degree of relatedness varies between near 0 (for two individuals who have no shared relatives in common, although the probability of sharing is never zero) and near 1 (for two identical twins who developed from one fertilized egg, although there are possible genetic effects occurring after postzygotic division that can decrease their similarity), so by combining various pairs of relatives and controlling for the degree to which they share an environment, their similarity on intelligence scores can be attributed, at least in part, to their genetic similarity. The corresponding estimates, when combined across multiple studies, can produce a general population estimate of the percentage of variance in the distribution of intelligence scores that may be attributed to the distribution of genetic similarities. According to meta-analytic estimates conducted on hundreds of quantitative genetic studies, inherited differences in DNA structure (e.g., the sequence of ACGTs) account, on average, for about 50 percent of the differences in intelligence scores (Plomin & von Stumm, 2018).

The second category of research in the etiology of intelligence is exemplified by hypotheses-driven candidate region and gene studies. These studies appeared en masse in the 1980s and were driven by the discoveries of a number of distinct genetic mechanisms for the development of intellectual disability or loss of intellectual functioning. Specifically, such studies include the analysis of the etiology of numerous and various genomic syndromes accompanied by intellectual disability. These syndromes are associated with a range of causes: Some are caused by particular lesions in distinct regions (e.g., the 22p11.2 region, as in DiGiorge syndrome), whereas others are caused by alterations in specific genes (e.g., the *PAH* gene in phenylketonuria). These discoveries, along with the ever-deepening understanding of the genetic architecture of the nervous system, were followed by investigations of both the genetic regions and specific genes for which there are some theoretical or empirical reasons to hypothesize an association with variation in intelligence and related cognitive functions; hundreds of studies were conducted, but the relevant meta-analytic summaries have not converted this research into a reliably replicable observation (e.g., for *COMT*, Barnett, Scoriels, & Munafò, 2008; for *BDNF*, Mandelman & Grigorenko, 2012). Similarly, insights into the genetic bases of degenerative disorders such as Alzheimer's disease have generated candidate genes whose variation has been investigated for their association with individual differences in intelligence and other cognitive functions (e.g., *APOE*, Weissberger et al., 2018).

The third category of studies, genome-wide association studies (GWAS), came to prominence early this century and now dominate the field (Deary et al., 2019; Hill et al., 2019). As the main assumption of these studies is that well-known common variations in the genome, represented by di-allelic single-nucleotide polymorphisms (SNPs), is associated with the variation in intelligence, and that, by definition (Butcher, Kennedy, & Plomin, 2006), the effect sizes associated with every single SNP are small, the main requirement for these types of studies is a large sample size. Since “the first apparently-decently-sized GWAS of intelligence” (Deary et al., 2019, p. 7) was published with an *N* of approximately 3,500 and yielded essentially null results (Davies et al., 2011), the subsequent increase in sample size to an *N* of around 300,000 (Davies et al., 2018) has resulted in the generation of a list of interesting signals throughout the genome (e.g., SNPs in both coding and noncoding regions deemed to be associated with intelligence by virtue of carefully conducted statistical analyses) that has rapidly increased to 150 (Deary et al., 2019). This figure reaches 187 when the findings from two (currently the largest) GWAS of education and intelligence are combined (Hill et al., 2018). Importantly, there are some overlaps among the findings generated in these different categories of studies. For example, the results of one of the GWAS were analyzed further by aligning its list of findings with chromosomal enhancer maps of eight brain regions and a so-called gene-set enrichment analysis was carried out, the point of which is to see which particular genes are overexpressed (i.e., expressed in increased quantity) in these regions (Du et al., 2018). Enhancers are short (50–1500 base pairs [bp]) DNA regulatory sequences that are abundant (400,000–1,000,000) in the human genome (Xie & Ren, 2013) and are known to regulate the expression of target genes through DNA looping (Shlyueva, Stampfel, & Stark, 2014). Multiple biological pathways showing cross-brain regions or brain region–specific association signals for human intelligence were detected (Du et al., 2018). One of these pathways, the systemic lupus erythematosus (SLE) pathway, was associated at $p < 0.01$, indicating that the genetic architecture of intelligence might overlap with that of SLE. Of note is that this pathway involves a number of genes (e.g., *BDNF, COMT*, and *APOE*) that have been considered candidate genes for intelligence in previous studies.

There are two more, among other, important pieces of information that were generated from the studies in this third category of studies into the genetic bases of intelligence. One pertains to the amount of variance in intelligence scores that can be explained by so-called genome-wide polygenic scores (GPS) for intelligence. GPS aggregates the effects of large numbers of SNPs associated with intelligence across the genome. Similarly, although early GWAS-explained variance wavered around 1 percent (Butcher et al., 2005; Davis et al., 2010), subsequent studies have brought this number up to 4 percent (Savage et al., 2018). While this number is four times greater than earlier estimates, it is still only 4 percent; the field yet needs to account for the “other” 96 percent. The other important piece of information pertains to the heritability estimates that can be obtained in these GWAS by comparing individuals’ overall genetic similarity (based on common SNPs) with their phenotypic similarity (based on intelligence scores). When heritability was estimated in this way, results indicated that about 25 percent of the variance in intelligence could be attributed to the variance in common SNPs (Davies et al., 2018; Savage et al., 2018). Such a discrepancy (e.g., 50% versus 25%) between quantitative and molecular genetic estimates has been observed before in studies of other complex human traits (J. Yang et al., 2010).

There are ongoing discussions in the literature regarding the absolute predictive and added value of GPS, compared, for example, with SES (von Stumm & Plomin, 2015), in understanding the sources of individual differences in complex human traits in general (Girirajan, 2017) and intelligence in particular (Sauce & Matzel, 2018). There is also an ongoing investigation into the "missing heritability," that is, the commonly found discrepancy between the heritability estimates obtained in quantitative versus molecular genetic studies on complex human traits (Manolio et al., 2009). The relevant literatures offer a number of pertinent hypotheses regarding these two conundrums. In this chapter we exemplify two. One such explanation relates to the heterogeneity of the phenotypes often used in genetic (quantitative and molecular) studies. The other proposes that, in addition to the genetic mechanism sampled by GWAS (i.e., the common variance-based mechanism), there may be other genetic mechanisms that influence individual differences in intelligence.

Intelligence-Related Phenotypes

Importantly, the overwhelming majority of molecular genetic studies of intelligence, at least all of the most recent GWAS (Benyamin et al., 2014; Davies et al., 2015; Davies et al., 2018; Davies et al., 2011; Kirkpatrick et al., 2014; Lencz et al., 2014; Savage et al., 2018; Sniekers et al., 2017; Trampush et al., 2017), have utilized indicators of general cognitive ability or general intelligence (the *g* or *g*-factor).

Since its introduction to the literature, thousands (and, probably, even hundreds of thousands) of pages have been written regarding the *g*-factor of intelligence. Deemed one of the most researched constructs in psychology, the *g*-factor is based on the noteworthy observation that indicators of performance on variable cognitive tasks are positively and substantially correlated. This means that, on average, people who perform well on a particular cognitive task (e.g., spatial reasoning tasks) tend to do well on all other cognitive tasks (e.g., verbal and numeric reasoning tasks). The existence of the *g*-factor is difficult to deny and is not questioned, really; it is typically derived as the first principal component of all correlated intelligence assessments (Mackintosh, 2011). What is questioned and debated is how this phenomenon arises, what its mechanisms are, and what it means; and on these points opinions vary widely.

The first point of contention pertains to what needs to be included in the assessment battery with which the *g*-factor will be derived. Numerous studies of countless assessments designed to measure different facets of intelligence (from whistling to computerized gamified tasks) have arrived at the replicable but notably obvious conclusion: The more homogeneous the set of the assessments used to derive *g*, the more shared variance is explained by the first principal component; the more heterogeneous these assessments are, the less shared variance is explained. Opinions about what needs to be included in these assessment batteries are diverse, ranging from the argument that *g* comprises primarily working memory (Oberauer et al., 2008), to the argument that multiple cognitive skills (Sternberg, 2003) and domains (Gardner, 2006) need to be sampled in order to derive a *g* that appropriately represents the aspects of human cognitive functioning important for operating in the world at large (Gould, 1981).

The second point of debate arises concerning the hypotheses proposed to explain the positive correlational manifold of different assessments. Most proponents of the *g*-factor assume the singularity of the mechanism that underlies this manifold, although this mechanism might be complex. Yet, although the *g*-factor is a univariate latent construct (i.e., a hidden summative of variance shared by a number of observed indicators that can be extrapolated only by particular analytical manipulations), it does not need to be reduced to a single property of the brain driven by a single genetic mechanism. In fact, there are other convincing explanations of how the positive correlational manifold might emerge (and the factor derived), such as complex coexisting dynamic systems reciprocally influencing each other (Sauce & Matzel, 2018; Van Der Maas et al., 2006).

Needless to say, the position taken regarding the first (construct definition and measurement) and second (etiology) points may bias the design and outcome of genetic studies on the etiology of intelligence. Thus, additional considerations should be mentioned. The landscape of molecular genetic studies today is shaped primarily by the notion that the genetic architecture of the *g*-factor emerges via numerous small effects, exerted by common genetic variants that linearly add to each other. The identification of these small effects requires a lot of statistical power, which needs to be generated by large sample sizes. As such, samples are difficult to amass at a single place and time point. The trend today is to merge multiple samples collected at different points for different reasons in different countries. These samples typically have been characterized using a variety of assessments of intelligence, mostly atheoretical, unsophisticated, and driven by other objectives (e.g., studying aging or psychiatric disorders), and the notion of the *g*-factor offers a convenient way to combine diverse sample-specific assessment batteries. Moreover, these samples are cross-sectional and represent individuals at different stages of the life span. Finally, these samples almost exclusively – if not exclusively – have been ascertained in high-income countries, sampling primarily from middle-class individuals of European descent.

Of note also is the assumption that for the “genetic puzzle” of intelligence to be solved, larger and larger samples are needed. Given that the differential in sample sizes between the studies that resulted in 1 percent (in a sample of 3.5×10^3) versus 4 percent (in a sample of 3.0×10^5) of explained variance in the *g*-factor was about 100 (10^2) times, there appears to be a diminishing return; each percentage of explained variance comes at the cost of larger and larger sample differentials. So, one question we might ask is, what percentage of explained variance could be achieved if all 7.4 billion (7.4×10^9, the current estimated population of Earth) people were genotyped? If the factor of 10 is used, assuming, in the logic of decently sized GWAS (Deary et al., 2019), that each multiplication of the sample size by 10 results in a gain of 1 percent of explained variance, then conducting a GWAS using the complete population of Earth should yield about 10 percent of explained variance for the intelligence of the planet’s population!

It is also important to acknowledge that, along with the massive efforts underway to analyze merged data sets using the *g*-factor as the critical phenotype, there are other efforts focused on the genetic bases of the cognitive processes that constitute that observed positive correlation manifold. These efforts can be subdivided into inquiries based on different typologies. The first typology distinguishes different types of reasoning sampled by conventional theories of intelligence, namely, verbal, numeric, and spatial reasoning. For these processes, there has been a fair amount of quantitative genetic analyses and substantially less molecular genetic work. Importantly, each of these types of reasoning can be and are assessed using multiple tasks that also correlate with each other within each type (substantially) and across the three types (creating the positive manifold of cognition). There is a range of heritability estimates associated with isolated tasks, ranging from 0 percent for a verbal fluency task (Sakakibara et al., 2018) to 69 percent for general spatial ability (Rimfeld et al., 2017). Moreover, different abilities demonstrate different susceptibilities to variability as a function of "other" variables that impact heritability estimates (e.g., SES); for example, heritability estimates have been reported to not vary as a function of parental education for nonverbal and general ability, but to vary for verbal ability (Spengler et al., 2018). Within this typology, verbal reasoning (also referred to as verbal IQ, VIQ) has been studied the most, not only because of its componential association with IQ (i.e., typically, the full scale IQ comprises two components, verbal and nonverbal IQ), but also because of its relevance to understanding the etiology of (a)typical language development. It has been observed in quantitative genetic studies (e.g., Hoekstra et al., 2009) that individual differences in VIQ are strongly influenced by genetic effects (84% and 82%), while performance on tasks tapping into more specific lexicon-related abilities demonstrates more modest genetic influence (29%–55%). Similarly, fluctuations in heritability have been observed in molecular estimates. For example (Davies et al., 2016), SNP-based heritability was highest for verbal–numerical reasoning (0.31), followed by that for reaction time (0.11) and then for memory (0.05).

The second typology is based on subdividing cognitive abilities by academic domains, such as reading and writing, mathematics, and science. Achievement in these domains is typically predicted by cognitive componential processes that are often included in tests of intelligence. For example, vocabulary is considered a componential process important for the acquisition of reading, but is also often assessed as a facet of intelligence. Similarly, many tests of intelligence use arithmetic operations, the mastery of which also predicts academic achievement in mathematics. Similar to studies of different types of reasoning, there have been quantitative and molecular studies of the (a) typical acquisition of reading and mathematics (for a review, see Grigorenko et al., 2019), but not science. As the assessments of these processes contribute to the positive correlational manifold of cognitive variables, they share variance with the *g*-factor and numerous other cognitive indicators. Although there have been several GWAS of reading and mathematics, with some convergent findings, there have not yet been any merged GWAS studies similar to those involving the *g*-factor, although attempts are underway with reading-related phenotypes.

This diversity of findings (i.e., the range of heritability estimates and GWAS signals) with regard to "other" cognitive processes that correlate with the *g*-factor is interpreted quite differently, compared to the interpretation of the proponents of the general cognitive ability, in the field. The following quotes illustrate the distance between various opinions regarding the progress that has been made in understanding the genetic bases of intelligence. The first is optimistic and forward-looking: "Because intelligence is one of the best predictors of educational and occupational outcomes, IQ GPSs will be used for prediction from early in life before intelligence or educational achievement can be assessed" (Plomin & von Stumm, 2018, p. 155). The second quote is much more conservative: "The influence of genes on IQ, are not as powerful or constrictive as might be assumed. . . . intelligence seems to be quite malleable, and changes in the environment can, by *interacting* with genes, explain a great deal of differences in IQ across families, life span, socioeconomic status, and generations" (Sauce & Matzel, 2018, p. 27). Close or not, this gap in opinions indicates that there may be other mechanisms that could be considered in the quest to understand the genetic bases of intelligence. Here we briefly sample a single family of such mechanisms, the epigenetic mechanisms.

Intelligence-Related Genetic Mechanisms

Epigenetic processes are molecular processes that regulate gene expression without altering the DNA sequence (Ptashne, 2007). Ultimately, such epigenetic modifications as DNA methylation, posttranslational modifications of histones, and others control the synthesis of proteins, thereby guiding cell specialization, functioning, and plasticity. Brain cells – neurons – are no exception. Epigenetic marks are dynamic and susceptible to temporal and environmental influences. These dynamics can be partially understood through both the structural characteristics of the genome and lifestyles, or a combination of these (Shah et al., 2014). Collectively, epigenetic marks and processes are referred to as the epigenome – "the regulatory control systems that dynamically alter transcriptional responses of the DNA as a function of cell state" (Winick-Ng & Rylett, 2018, p. 2). The important role of epigenetic modifications in cognitive processes has been exemplified by numerous studies, most of which are being researched in animal models. Overall, two (of many) epigenetic processes, DNA methylation (i.e., molecular modifications of DNA changing its transcriptional properties) and histone acetylation (i.e., molecular modifications of histones – proteins that package and order the DNA into structural units called nucleosomes – changing the folding of DNA and, therefore, its openness for transcription), are the most widely studied epigenetic modifications.

DNA methylation consists of the addition of a methyl group to a cytosine located in the gene promoter, leading to a decrease in its activity (Boyes & Bird, 1992; Hsieh, 1994). According to one hypothesis, methyl groups inhibit the transcription factor from binding to the DNA strand (Bird, 1986). According to another, methylation facilitates the binding of enzymes that suppress transcriptional activity (Bird, 2002). Regardless of the mechanism, DNA methylation is a dynamic process that is related to memory formation and neuronal plasticity. Miller and Sweatt (2007) demonstrated that fear conditioning leads to increased DNA methylation of the phosphatase 1 (*Pp1*) and *reelin* (*Reln*) genes in the rat hippocampus within one hour after training, and that the levels of DNA methylation return to baseline within twenty-four hours of training. Similarly, a decrease in the DNA methylation level in the brain-derived neurotrophic factor (*Bdnf*) gene was registered within two hours after fear conditioning learning; it returned to baseline within twenty-four hours of training (Lubin, Roth, & Sweatt, 2008). Moreover, DNA methylation processes underlie not only fear-related, but also reward-related associative memory (Day et al., 2013). DNA demethylation in the immediate early genes *Egr1* and *Fos* (these genes encode transcription factors activated transiently and rapidly in nerve cells in response to stimulation; Anokhin & Rose, 1991; Brennan, Hancock, & Keverne, 1992) was demonstrated to play an important role in memory formation and the long-term potentiation of brain regions such as the ventral tegmental area (VTA) required for the formation of reward-related associative memories. Furthermore, these changes in DNA methylation patterns were correlated with transcriptional changes and were required for the formation of reward-related associative memories.

In general, levels of DNA methylation in the genome across the life span are maintained primarily by two groups of enzymes: DNA methyltransferases (DNMTs), which add a methyl group to the fifth position of cytosine, and ten-eleven translocation (TET) enzymes, which catalyze the oxidation of 5-methylcytosine, which is removed by base excision repair and substituted back with an unmodified cytosine (Z.-x. Chen & Riggs, 2011; Kohli & Zhang, 2013). There is a growing literature connecting these DNA methylation maintenance enzymes with cognitive functioning. For example, the deficiency of two DNMTs (1 and 3a) in the CNS of adult mice caused deficiencies in synaptic plasticity, learning, and memory (Feng et al., 2010). Whereas the decreased expression of DNMT 3a2 was associated with cognitive functioning decrease, restoring its transcription levels was associated with the restoration of cognitive functioning in both young and aged mice (Oliveira, Hemstedt, & Bading, 2012; Oliveira et al., 2015). Tet-1 deficiency has been demonstrated to be associated with deficient spatial learning and memory (Zhang et al., 2013).

In contrast to DNA methylation, histone acetylation, the addition of a negatively charged acetyl group to a lysine residue on a histone protein, promotes an increase in the transcriptional activation of genes (Wang et al., 2008). There are two hypotheses explaining the intensification of transcription. According to the first hypothesis, this activation is explained by reducing the affinity between the positively charged residue and negatively charged DNA, thereby increasing the availability of DNA sites to transcription factors (Lee et al., 1993). According to an alternative hypothesis, transcription factors can recognize modified lysine residues of histone molecules and interact directly with these histones (Vettese-Dadey et al., 1996; Vitolo, Thiriet, & Hayes, 2000). There is substantial evidence to support the important role of histone acetylation in learning and memory processes. An increase in histone acetylation in the rat hippocampus has been registered after fear conditioning (Levenson et al., 2004; Lubin et al., 2008). Dagnas and Mons (2013) showed that spatial training increased the acetylation of histones H3 and H4 selectively in the dorsal hippocampal CA1 region and the dentate gyrus (DG), and that cued training significantly enhanced acetylation of both histones selectively in the dorsal striatum. The inhibition of histone deacetylases (HDAC), which leads to an increase in the level of histone acetylation in the brain and is accompanied by an increase in transcriptional activity (Sinn et al., 2007), was shown to enhance memory formation in mice (Bredy & Barad, 2008; Stefanko et al., 2009; Vecsey et al., 2007).

It is important to note that there are gene-specific patterns of epigenetic processes. For example, such patterns of DNA methylation and histone acetylation have been described for the *Bdnf* gene and implicated in learning and memory (Lubin et al., 2008). Of note is that DNA methylation and histone modifications are interrelated processes; transcriptional activation is implemented by decreasing the level of DNA methylation, which, in turn, is accompanied by an increase in the level of histone acetylation (Z. J. Chen & Pikaard, 1997; Eden et al., 1998; P. L. Jones et al., 1998).

A separate body of animal data is focused on the role of epigenetic mechanisms in age-related cognitive impairment (Dagnas & Mons, 2013; Ianov et al., 2017; Penner et al., 2011). Specifically, Penner and colleagues (2011) showed that the patterns of methylation in one of the genes important for normal memory function, *Arc*, differ between adult and aged rats in two hippocampal regions, CA1 and the dentate gyrus. These alterations could be responsible for age-related decreases in *Arc* transcription in the hippocampus, leading to the observed impaired spatial memory on the Morris swim task (Penner et al., 2011). Dagnas and Mons showed that, in mice, age-related spatial memory deficits were associated with greater acetylation levels of histone H3 and lower acetylation levels of histone H4 compared to young controls both in CA1 and the DG of the hippocampus. Furthermore, the level of histone H4 acetylation in CA1 did not differ between aged trained rats and control untrained animals in contrast to young animals, in whom these differences were observed. At the same time, aging did not affect memory consolidation in the striatum-dependent cued water maze task with significant increase in the level of acetylation of histones 3 and 4 (Ac-H3 and Ac-H4, respectively) in the striatum of aged trained mice, compared with age-matched untrained controls, even when the levels of Ac-H3 and Ac-H4 were significantly lower in the aged group than in young mice (Dagnas & Mons, 2013). It has been demonstrated that cognitive impairments in aged rats on a set shifting task that depends on the medial prefrontal cortex (mPFC) are associated with DNA methylation in this brain region (Ianov et al., 2017). In general, the impairment of cognitive functions in aging animals has been shown on a number of tasks, and the epigenetic changes accompanying cognitive changes were specific for the brain structures associated with performance on the given task. The administration of HDAC inhibitors (i.e., chemical compounds prohibiting or minimizing histone deacetylation) was shown to enhance aging-associated memory deficits in rats (Reolon et al., 2011).

At this point, the literature on the role of epigenetic modifications in cognitive processes in humans, especially in the brain, is small. However, the presence of correspondence in the patterns of DNA methylation between brain tissue and peripheral tissue, for example, blood (Walton et al., 2016) and saliva (Smith et al., 2015), provides opportunities to obtain interesting and informative data. For example, it has been demonstrated that increased ethylation in the *NR3C1* gene, as assessed in saliva, is associated with reduced performance on a working memory task and increased activation in the right ventrolateral prefrontal cortex and the cuneus (Vukojevic et al., 2014). An association between increased methylation in the *COMT* gene in blood and greater left dorsolateral prefrontal cortex activity was registered during a working memory task; furthermore, the *COMT* methylation patterns positively correlated with working memory task performance at a trend level (Walton et al., 2014). Increased methylation of the Val^{158} allele in the *COMT* gene was associated, conversely, with reduced bilateral prefrontal cortex activity in a working memory task and increased performance on the task (Ursini et al., 2011). Discordance between the results of the latter two studies could be explained by the different gene regions analyzed (promoter versus exon, correspondingly). These studies were performed with healthy young participants.

There are also studies, both prospective and retrospective, that take advantage of large existing cohorts, representative of the general population at the ascertainment location. In one such study, the Avon Longitudinal Study of Parents and Children, ALSPAC, mother-child dyads have been followed for a number of years (www.bristol.ac.uk/alspac); pregnant women were recruited and information was collected on both mothers and their children. Among the multitude of studies generated by this project are studies of early nutrition. One is based on the fact that carbon metabolism includes, among other bioagents, folate and several B vitamins. Folate and B vitamins play an important role in prenatal nutrition. Deficiencies in these bioagents are known to be a risk factor for abnormal neurodevelopment (Rush, Katre, & Yajnik, 2014). In addition, pregnancy supplementation with these bioagents has been associated with reduced risk of severe language delay (Roth et al., 2011) and autism spectrum disorders (Surén et al., 2013). There are also studies that have demonstrated associations between maternal vitamin B12 status during pregnancy and child cognition (Bhate et al., 2008; Villamor et al., 2012). Although the mechanisms for these effects are not fully understood, there is a line of evidence for a pathway involving one carbon metabolism and the donation of a methyl group affecting downstream processes, including DNA methylation. For example, maternal prenatal vitamin B12 status impacts DNA methylation in fetuses at global and site-specific levels (Ba et al., 2011; McKay et al., 2012). These diverse observations were connected to investigate the potential mediating role of DNA methylation in the observed relationship between maternal prenatal vitamin B12 level and child IQ at the age of eight in a large sample from the ALSPAC; the results indicated the presence of causal effects, first, of maternal vitamin B12 levels on cord blood DNA methylation, and, second, of vitamin B12-responsive DNA methylation changes on children's cognition (Caramaschi et al., 2017).

Another line of reasoning connects the epigenome and cognition through the brain. There are six literatures that are relevant here (Kaminski et al., 2018). First, there is a large body of research associating the *g*-factor with the architecture of the brain, specifically, cortical (Karama et al., 2013) and subcortical (Burgaleta et al., 2014; Grazioplene et al., 2015; MacDonald et al., 2014) volume and thickness. Second, there are well-known cortical-subcortical networks that that support goal-directed behavior and decision-making (Alexander, DeLong, & Strick, 1986; Neubert et al., 2015). Third, there is evidence that cognitive training in reasoning impacts resting state connectivity between cortical and subcortical brain structures (Mackey, Miller Singley, & Bunge, 2013). Fourth, the neurotransmitter dopamine plays many roles in the brain, including connecting cortical and subcortical structures and being associated with habitual versus goal-directed behavior and cognition (Deserno et al., 2015; Schlagenhauf et al., 2013). Fifth, there is a well-known and acquirable proxy for dopaminergic neurotransmission, the reward anticipation signal that can be registered with fMRI (Schott et al., 2008). Sixth, dopaminergic signaling is heritable, but is also malleable (Vaessen et al., 2015). Based on all these literatures, it is plausible to assume that epigenetic modifications to the dopaminergic system might serve as a mechanism for the malleability of intelligence via alterations to cortical-subcortical networks of the brain. Moreover, now there is empirical evidence supporting this assumption (Kaminski et al., 2018).

Several studies were also undertaken on healthy older persons in order to examine DNA and the connection between methylation and cognitive aging. In this regard, two epigenetic mechanisms have been prioritized: the epigenetic clock and epigenetic drift (M. J. Jones, Goodman, & Kobor, 2015). The concept of an epigenetic clock is used to refer to functional age-related epigenetic changes at specific sites of the genome common across individuals. The concept of epigenetic drift is used to refer to the global decrease in the stability and precision of DNA methylation with age. Both the epigenetic clock and drift reflect complex genome-wide changes in DNA methylation levels that tend, with age, to rise in regions known for their low methylation (e.g., promoter-associated CpG[1] islands), but decline in regions with high DNA methylation (e.g., intergenic non-island CpG sites), thus leading to a global loss of methylation in the genome (Heyn et al., 2012; Illingworth & Bird, 2009).

No significant global methylation differences in blood were associated with performance on a battery of cognitive tests in older individuals (Gomes et al., 2012; Schiepers et al., 2011). In the first of two studies, cognitive performance was assessed by means of a test battery consisting of a visual verbal word–learning task, the Stroop color-word interference test, a concept-shifting test, a letter-digit substitution test, and a verbal fluency test, which accessed cognitive functioning in the domains of memory, sensorimotor speed, complex speed, information-processing speed, and word fluency (Schiepers et al., 2011). In the second study, the Mini-Mental State Examination was used (Gomes et al., 2012). However, as previously mentioned, these data were obtained at the global level of methylation and do not exclude the possibility of involving more subtle epigenetic mechanisms.

At this point, there is only one meta-analysis of epigenome-wide association studies (EWAS) focusing on cognitive abilities (Marioni et al., 2018). In this study, seven measures of cognitive functioning were considered: Wechsler Logical

[1] CpG is shorthand for 5'—C—phosphate—G—3', cytosine and guanine separated by only one phosphate group. DNA methylation in mammals mainly occurs on the cytosine nucleotide of a CpG site; in fact, 70 percent to 80 percent of CpG cytosines in mammalian genomes are methylated.

Memory as a measure of verbal declarative memory; Wechsler Digit Symbol Test/ Symbol Digit Modalities Test/Letter Digit Substitution Test as a measure of processing speed; Semantic Verbal Fluency as a measure of an aspect of executive function; Phonemic Verbal Fluency as a measure of an aspect of executive function; Trail Making Test Part B as a measure of an aspect of executive function; Boston Naming Test/National Adult Reading Test/any other measure of vocabulary; and Mini-Mental State Examination as a measure of general cognitive function. Across these seven measures of cognitive functioning, epigenome-wide significant associations were registered for global cognitive function and for phonemic verbal fluency. Higher methylation of CpGs located in an intergenic region on chromosome 12 was associated with lower scores on indicators of cognitive function, and in an intergenic region on chromosome 10 with lower indicators of phonemic verbal fluency (Marioni et al., 2018). Phonemic verbal fluency was associated with the DNA methylation of CpGs located in the *INPP5A* gene and the *MAML3* gene. *INPP5A* (inositol polyphosphate-5-phosphatase A) is a gene coding for the protein hydrolyzing inositol 1,4,5-triphosphate, which mobilizes intracellular calcium and acts as a second messenger mediating cell responses to various stimuli. The deletion of this gene is known to cause a deficit in motor coordination in mice running on an accelerating rotarod, and the reduction of cerebellum volume due to Purkinje cell loss (A. W. Yang, Sachs, & Nystuen, 2015). The *MAML3* (mastermind like transcriptional coactivator 3) gene acts as a transcriptional coactivator for NOTCH proteins (Wu et al., 2002) implicated in learning and memory (Costa, Honjo, & Silva, 2003; Presente et al., 2004). Measures of vocabulary were associated with the DNA methylation of CpGs located in the *ESRP2* (epithelial splicing regulatory protein 2) gene; the specifics of this gene's role in neuronal plasticity have not been yet identified.

In a different report (Starnawska et al., 2017), verbal fluency, immediate word recall, and delayed word recall were analyzed not separately, as in previous studies, but as part of a composite score (McGue & Christensen, 2001, 2002). Cognitive decline over ten years in aging individuals was associated with DNA methylation levels in the *AGBL4* and *SORBS1* genes, with the measure of cognitive functioning defined by a composite score based on six brief cognitive tests, namely, for verbal fluency, immediate word recall, delayed word recall, processing speed, attention, and working memory (Starnawska et al., 2017). The former gene is known for its brain-specific expression, at its highest level in the frontal cortex (Lonsdale et al., 2013), and for its function in neuronal survival (Rogowski et al., 2010). Elevated expression of *SOBRS1* was found in the hippocampus of Alzheimer's patients (Blalock et al., 2004). This study also demonstrated some cross-sectional associations. Specifically, cognitive functioning was associated with level of DNA methylation in the *ZBTB46* and *TAF12* genes. The *ZBTB46* gene is most highly expressed in the cerebellum and cerebellar hemisphere (Lonsdale et al., 2013); the rs6062314 polymorphism in this gene was associated with multiple sclerosis (International Multiple Sclerosis Genetics Consortium, 2013). Of note is that these findings did not overlap with any findings in any GWAS completed so far (see section Intelligence-Related Phenotypes).

With regard to verbal IQ, it is important to consider a few more studies. Structural variations in the *CNTNAP2* gene are associated with language impairment (Vernes et al., 2008) and early language acquisition (Whitehouse et al., 2011). Vernes and colleagues (2008) analyzed *CNTNAP2* polymorphisms in children with typical specific language impairment and showed quantitative associations with nonsense-word repetition, a behavioral marker of this disorder. Whitehouse and colleagues (2011) administered the Communication subscale of the Infant Monitoring Questionnaire to parents of children at age two to access the association between *CNTNAP2* variants and early language acquisition. They reported that *CNTNAP2* SNPs were associated with indicators of communicative behavior. Widespread differences in cortex DNA methylation of the *CNTNAP2* gene between humans and chimpanzees might also support a role for epigenetic mechanisms in human-specific language skills (Vernes et al., 2008).

There are also data from songbirds, an established animal model of human speech processing, language acquisition, and speech comprehension (Doupe & Kuhl, 1999). Kelly and colleagues (2018) used either isolated or "tutored" juvenile male zebra finches that had learned to sing from an adult "tutor" male to identify how tutor experience alters the brain and controls the ability to learn. They documented that tutor experiences can induce epigenetic changes in the auditory forebrain, a region required for tutor song learning.

In addition to data on the role of epigenetic modifications in cognitive processes in healthy individuals, large amounts of data originate from patients with various pathological conditions. A number of syndromes, such as Rett syndrome, Angelman syndrome, and Prader-Willi syndrome, are characterized by various forms of mental retardation and are at least in part caused by epigenetic deregulation. Angelman syndrome is a severe form of mental retardation, with most patients having a vocabulary of only one or two words (Laan, Haeringen, & Brouwer, 1999). The syndrome is associated with disturbances in genomic imprinting, one of the epigenetic mechanisms regulating gene expression (Glenn et al., 2000). The same is true for Prader-Willi syndrome, which is also associated with disturbances in genomic imprinting (Glenn et al., 2000); this syndrome presents with mild mental retardation; language, learning, and memory impairments are also observed (Holm et al., 1993). In the case of Rett syndrome, female patients are characterized by severe mental retardation and language impairments (Gold et al., 2018; Weaving et al., 2005). This syndrome is caused by alterations to the function of the methyl CpG binding protein 2 (*MeCP2*) gene, which is responsible for the epigenetic regulation of DNA methylation (Amir et al., 1999). Alzheimer's disease is a chronic neurodegenerative disease affecting mainly older individuals; its symptoms include severe speech and memory impairments (Arshavsky, 2014) with underlying dysregulation of epigenetic processes (Nativio et al., 2018; Winick-Ng & Rylett, 2018).

To summarize, there is considerable evidence of the important role of epigenetic modifications in cognitive processes. Interestingly, there are even some genes that are featured in both the literature on common variants and that on epigenetic modifications. The data on the role of epigenetic mechanisms in learning and cognition in general and componential processes of intelligence in particular not only have fundamental importance, but could also be a promising tool from a therapeutic point of view.

Conclusions

This overview is intended to present selected snapshots of the current accomplishments in the field of the genetic bases of intelligence. The status of the field reflects its 100-plus-year history, with many ideas traceable to the second half of the nineteenth century. Indeed, the idea that intelligence is heritable and transmittable from generation to generation is as old as the field of intelligence. So what is new? From our point of view, there are distinct innovations in the field, some incremental and others with course-changing potential.

Incremental innovations include a continuous effort to build a powerful enough sample to "catch" small genetic effects that, additively, might enhance the understanding of the genetic architecture of intelligence. This attempt is driven by the necessity to converge on what is common to all assessments of intelligence across all included samples, minimizing uniqueness and diversity both in the variation of cognitive abilities and in the sampled populations. Here, the diversity of assessments and the diversity of samples are treated as error; it is assumed that as measures and samples are homogenized, the quintessential genetic root of intelligence will be excavated. It turns out that the root is both wide and deep, and even with a sample size of a third of a million we are far from either identifying or documenting all of the essential branches of the root. As more and more genotyped samples with even minimal (but reasonable) intelligence assessment data become available, the construction of more and more powerful samples to be utilized in this line of research will continue. At this point, these studies have resulted in a list of about 150 genes (out of about 20,000 total), providing an estimate of about 4 percent (out of 100%) of variance explained, so there are more genes and more variance to be investigated in this effort. All of this, in the end, will explain only the *g*-factor, which is an abstract construct itself, accounting only for about 40 percent (i.e., the minority) of the human cognitive repertoire.

Course-changing innovations require explicit deviation from an established research trajectory. From our point of view, the consideration of epigenetic mechanisms in understanding the genetic bases of intelligence may become such an innovation. Epigenetic markers are dynamic and, when sampled multiple times, may provide a window into the dynamics of human learning. And when sampled in multiple environments, such markers may provide a window onto the diversity of human abilities required in these different environments. The logic of the incremental innovation is that we are on the right track and that we are near done (if 4% seems satisfactory); we just need to make a few more steps in the same direction. The logic of the course-changing innovation is that we are at a crossroads, where it may be important to take a step in a different direction.

The current fashion of merging multiple genetic data sets and homogenizing the phenotypes of interest by simplification, although helpfully reductionist, might not necessarily be ecologically valid. Of course there is value in working with large merged data sets. But there is also value in pursuing unfashionable ideas, even though they may not be immediately replicable or easily implemented, and may require not only analytical, but also design and data collection sophistication. The challenges seem insurmountable. But is this not what intelligence is for?

References

Alexander, G. E., DeLong, M. R., & Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience*, *9*, 357–381. https://doi.org/10.1146/annurev.ne.09.030186.002041

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, *23*, 185–188.

Anokhin, K. V., & Rose, S. P. (1991). Learning-induced increase of immediate early gene messenger RNA in the chick forebrain. *European Journal of Neuroscience*, *3*, 162–167.

Arshavsky, Y. I. (2014). Alzheimer disease and cellular mechanisms of memory storage. *Journal of Neuropathology and Experimental Neurology*, *73*, 192–205. https://doi.org/10.1097/NEN.0000000000000043

Ba, Y., Yu, H., Liu, F., Geng, X., Zhu, C., Zhu, Q., et al. (2011). Relationship of folate, vitamin B12 and methylation of insulin-like growth factor-II in maternal and cord blood. *European Journal of Clinical Nutrition*, *65*, 480–485. https://doi.org/10.1038/ejcn.2010.294

Barnett, J. H., Scoriels, L., & Munafò, M. R. (2008). Meta-analysis of the cognitive effects of the Catechol-O-Methyltransferase gene Val158/108Met polymorphism. *Biological Psychiatry*, *64*, 137–144. https://doi.org/10.1016/j.biopsych.2008.01.005

Benyamin, B., Pourcain, B., Davis, O. S., Davies, G., Hansell, N. K., Brion, M. J., et al. (2014). Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. *Molecular Psychiatry*, *19*, 253–258. https://doi.org/10.1038/mp.2012.184

Bhate, V., Deshpande, S., Bhat, D., Joshi, N., Ladkat, R., Watve, S., et al. (2008). Vitamin B12 status of pregnant Indian women and cognitive function in their 9-year-old children. *Food and Nutrition Bulletin*, *29*, 249–254. https://doi.org/10.1177/156482650802900401

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature*, *321*, 209–213.

Bird, A. P. (2002). DNA methylation patterns and epigenetic memory. *Genes and Development*, *16*, 6–21.

Blalock, E. M., Geddes, J. W., Chen, K. C., Porter, N. M., Markesbery, W. R., et al. (2004). Incipient Alzheimer's disease: Microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proceedings of the National Academy of Sciences*, *101*, 2173–2178. https://doi.org/10.1073/pnas.0308512100

Boyes, J., & Bird, A. P. (1992). Repression of genes by DNA methylation depends on CpG density and promoter strength: Evidence for involvement of a methyl-CpG binding protein. *EMBO Journal*, *11*, 327–333.

Bredy, T. W., & Barad, M. (2008). The histone deacetylase inhibitor valproic acid enhances acquisition, extinction, and reconsolidation of conditioned fear. *Learning and Memory*, *15*, 39–45. https://doi.org/10.1101/lm.801108

Brennan, P. A., Hancock, D., & Keverne, E. B. (1992). The expression of the immediate-early genes c-fos, egr-1 and c-jun in the accessory olfactory bulb during the formation of an olfactory memory in mice. *Neuroscience*, *49*, 277–284.

Burgaleta, M., MacDonald, P. A., Martínez, K., Román, F. J., Álvarez-Linera, J., González, A. R., et al. (2014). Subcortical regional morphology correlates with fluid and spatial intelligence. *Human Brain Mapping*, *35*, 1957–1968. https://doi.org/10.1002/hbm.22305

Butcher, L. M., Kennedy, J. K., & Plomin, R. (2006). Generalist genes and cognitive neuroscience. *Current Opinion in Neurobiology*, *16*, 145–151.

Butcher, L. M., Meaburn, E., Knight, J., Sham, P. C., Schalkwyk, L. C., Craig, I. W., et al. (2005). SNPs, microarrays, and pooled DNA: Identification of four loci associated with mild mental impairment in a sample of 6,000 children. *Human Molecular Genetics*, *14*, 1315–1325.

Caramaschi, D., Sharp, G. C., Nohr, E. A., Berryman, K., Lewis, S. J., Davey Smith, G., et al. (2017). Exploring a causal role of DNA methylation in the relationship between maternal vitamin B12 during pregnancy and child's IQ at age 8, cognitive performance and educational attainment: A two-step Mendelian randomization study. *Human Molecular Genetics*, *26*, 3001–3013. https://doi.org/10.1093/hmg/ddx164

Chen, Z-x., & Riggs, A. D. (2011). DNA methylation and demethylation in mammals. *Journal of Biological Chemistry*, 286, 18347–18353. https://doi.org/10.1074/jbc.R110.205286

Chen, Z. J., & Pikaard, C. S. (1997). Epigenetic silencing of RNA polymerase I transcription: A role for DNA methylation and histone modification in nucleolar dominance. *Genes and Development*, *11*, 2124–2136.

Costa, R. M., Honjo, T., & Silva, A. J. (2003). Learning and memory deficits in Notch mutant mice. *Current Biology*, *13*(15), 1348–1354.

Dagnas, M., & Mons, N. (2013). Region- and age-specific patterns of histone acetylation related to spatial and cued learning in the water maze. *Hippocampus*, *23*, 581–591. https://doi.org/10.1002/hipo.22116

Davies, G., Armstrong, N., Bis, J. C., Bressler, J., Chouraki, V., Giddaluru, S., et al. (2015). Genetic contributions to variation in general cognitive function: A meta-analysis of genome-wide association studies in the CHARGE consortium (N=53 949). *Molecular Psychiatry*, *20*, 183–192. https://doi.org/10.1038/mp.2014.188

Davies, G., Lam, M., Harris, S. E., Trampush, J. W., Luciano, M., Hill, W. D., et al. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nature Communications*, *9*(1), 2098. https://doi.org/10.1038/s41467-018-04362-x

Davies, G., Marioni, R. E., Liewald, D. C., Hill, W. D., Hagenaars, S. P., Harris, S. E., et al. (2016). Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Molecular Psychiatry*, *21*, 758–767. https://doi.org/10.1038/mp.2016.45

Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S. E., Liewald, D., et al. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry*, *16*, 996–1005. https://doi.org/10.1038/mp.2011.85

Davis, O. S., Butcher, L. M., Docherty, S. J., Meaburn, E. L., Curtis, C. J., Simpson, M. A., et al. (2010). A three-stage genome-wide association study of general cognitive ability: Hunting the small effects. *Behavior Genetics*, *40*, 759–767.

Day, J. J., Childs, D., Guzman-Karlsson, M. C., Kibe, M., & Moulden, J. (2013). DNA methylation regulates associative reward learning. *Nature Neuroscience*, *16*, 1445–1452.

Deary, I. J., Harris, S. E., & Hill, W. D. (2019). What genome-wide association studies reveal about the association between intelligence and physical health, illness, and mortality. *Current Opinion in Psychology*, *27*, 6–12. https://doi.org/10.1016/j.copsyc.2018.07.005

Deserno, L., Huys, Q. J. M., Boehme, R., Buchert, R., Heinze, H.-J., Grace, A. A., et al. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences*, *112*(5), 1595–1600. https://doi.org/10.1073/pnas.1417219112

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience*, *22*, 567–631. https://doi.org/10.1146/annurev.neuro.22.1.567

Du, Y., Ninga, Y., Wena, Y., Liua, L., Lianga, X., Lia, P., et al. (2018). A genome-wide pathway enrichment analysis identifies brain region related biological pathways associated with intelligence. *Psychiatry Research*, *268*, 238–242. https://doi.org/10.1016/j.psychres.2018.07.029

Eden, S., Hashimshony, T., Keshet, I., Cedar, H., & Thorne, A. W. (1998). DNA methylation models histone acetylation. *Nature*, *394*, 842. https://doi.org/10.1038/29680

Feng, J., Zhou, Y., Campbell, S. L., Le, T., Li, E., Sweatt, J. D., et al. (2010). Dnmt1 and Dnmt3a are required for the maintenance of DNA methylation and synaptic function in adult forebrain neurons. *Nature Neuroscience*, *13*, 423–430.

Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences*. London: Macmillan and Co.

Gardner, H. (2006). *Multiple intelligences: New horizons in theory and practice*. New York: Basic Books.

Girirajan, S. (2017). Missing heritability and where to find it. *Genome Biology*, *18*, 89. https://doi.org/10.1186/s13059-017-1227-x

Glenn, C. C., Deng, G., Michaelis, R. C., Tarleton, J., Phelan, M. C., Surh, L., et al. (2000). DNA methylation analysis with respect to prenatal diagnosis of the Angelman and Prader-Willi syndromes and imprinting. *Prenatal Diagnosis*, *20*, 300–306.

Gold, W. A., Krishnarajy, R., Ellaway, C., & Christodoulou, J. (2018). Rett syndrome: A genetic update and clinical review focusing on comorbidities. *ACS Chemical Neuroscience*, *9*, 167–176. https://doi.org/10.1021/acschemneuro.7b00346

Gomes, M. V. M., Toffoli, L. V., Arruda, D. W., Soldera, L. M., Pelosi, G. G., Neves-Souza, R. D., et al. (2012). Age-related changes in the global DNA methylation profile of leukocytes are linked to nutrition but are not associated with the MTHFR C677T genotype or to functional capacities. *PLoS One*, *7*, e52570. https://doi.org/10.1371/journal.pone.0052570

Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.

Grazioplene, R. G., Ryman, S. G., Gray, J. R., Rustichini, A., Jung, R. E., & DeYoung, C. G. (2015). Subcortical intelligence: Caudate volume predicts IQ in healthy adults. *Human Brain Mapping*, *36*, 1407–1416. https://doi.org/10.1002/hbm.22710

Grigorenko, E. L., Compton, D., Fuchs, L., Wagner, R., Wilcutt, E., Fletcher, J. M. (2019) Understanding, educating, and supporting children with specific learning disabilities: 50 years of science and practice. *American Psychologist*. https://doi.org/10.1002/cad.20290

Heyn, H., Li, N., Ferreira, H. J., Moran, S., Pisano, D. G., Gomez, A., et al. (2012). Distinct DNA methylomes of newborns and centenarians. *Proceedings of the National Academy of Sciences*, *109*, 10522–10527. https://doi.org/10.1073/pnas.1120658109

Hill, W. D., Harris, S. E., & Deary, I. J. (2019). What genome-wide association studies reveal about the association between intelligence and mental health. *Current Opinion in Psychology*, *27*, 25–30. https://doi.org/10.1016/j.copsyc.2018.07.007

Hill, W. D., Marioni, R. E., Maghzian, O., Ritchie, S. J., Hagenaars, S. P., McIntosh, A. M., et al. (2018). A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. *Molecular Psychiatry*, *24*, 169–181. https://doi.org/10.1038/s41380-017-0001-5

Hoekstra, R. A., Bartels, M., van Leeuwen, M., & Boomsma, D. I. (2009). Genetic architecture of verbal abilities in children and adolescents. *Developmental Science*, *12*, 1041–1053. https://doi.org/10.1111/j.1467-7687.2009.00843.x

Holm, V. A., Cassidy, S. B., Butler, M. G., Hanchett, J. M., Greenswag, L. R., Whitman, B. Y., & Greenberg, F. (1993). Prader-Willi syndrome: Consensus diagnostic criteria. *Pediatrics*, *91*, 398–402.

Hsieh, C. L. (1994). Dependence of transcriptional repression on CpG methylation density. *Molecular and Cellular Biology*, *14*(8), 5487–5494.

Ianov, L., Riva, A., Kumar, A., & Foster, T. C. (2017). DNA methylation of synaptic genes in the prefrontal cortex is associated with aging and age-related cognitive impairment. *Frontiers in Aging Neuroscience*, *9*, 249. https://doi.org/10.3389/fnagi.2017.00249

Illingworth, R. S., & Bird, A. P. (2009). CpG islands – "A rough guide." *FEBS Letters*, *583*, 1713–1720. https://doi.org/10.1016/j.febslet.2009.04.012

International Multiple Sclerosis Genetics Consortium. (2013). MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain*, *136*, 1778–1782. https://doi.org/10.1093/brain/awt101

Jones, M. J., Goodman, S. J., & Kobor, M. S. (2015). DNA methylation and healthy human aging. *Aging Cell*, *14*, 924–932. https://doi.org/10.1111/acel.12349

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., et al. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics*, *19*, 187–191. https://doi.org/10.1038/561

Kaminski, J. A., Schlagenhauf, F., Rapp, M., Awasthi, S., Ruggeri, B., Deserno, L., et al. (2018). Epigenetic variance in dopamine D2 receptor: A marker of IQ malleability? *Translational Psychiatry*, *8*, 169. https://doi.org/10.1038/s41398-018-0222-7

Karama, S., Bastin, M. E., Murray, C., Royle, N. A., Penke, L., Muñoz Maniega, S., et al. (2013). Childhood cognitive ability accounts for associations between cognitive ability and brain cortical thickness in old age. *Molecular Psychiatry*, *19*, 555–559. https://doi.org/10.1038/mp.2013.64

Kelly, T. K., Ahmadiantehrani, S., Blattler, A., & London, S. E. (2018). Epigenetic regulation of transcriptional plasticity associated with developmental song learning. *Proceedings of the Royal Society B: Biological Sciences*, 285. https://doi.org/10.1098/rspb.2018.0160

Kirkpatrick, R. M., McGue, M., Iacono, W. G., Miller, M. B., & Basu, S. (2014). Results of a "GWAS Plus": General cognitive ability is substantially heritable and massively polygenic. *PLoS One*, *9*(11), e112390. https://doi.org/10.1371/journal.pone.0112390

Kohli, R. M., & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*, *502*, 472–479. https://doi.org/10.1038/nature12750

Laan, L. A., Haeringen, A., & Brouwer, O. F. (1999). Angelman syndrome: A review of clinical and genetic aspects. *Clinical Neurology and Neurosurgery*, *101*, 161–170.

Lee, D. Y., Hayes, J. J., Pruss, D., & Wolffe, A. P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell*, *72*, 73–84.

Lencz, T., Knowles, E., Davies, G., Guha, S., Liewald, D. C., Starr, J. M., et al. (2014). Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: A report from the Cognitive Genomics ConsorTium (COGENT). *Molecular Psychiatry*, *19*, 168–174. https://doi.org/10.1038/mp.2013.166

Levenson, J. M., O'Riordan, K. J., Brown, K. D., Trinh, M. A., Molfese, D. L., & Sweatt, J. D. (2004). Regulation of histone acetylation during memory formation in the hippocampus. *Journal of Biological Chemistry*, *279*, 40545–40559.

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., et al. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, *45*, 580–585. https://doi.org/10.1038/ng.2653

Lubin, F. D., Roth, T. L., & Sweatt, J. D. (2008). Epigenetic regulation of bdnf gene transcription in the consolidation of fear memory. *Journal of Neuroscience*, *28*, 10576–10586.

MacDonald, P. A., Ganjavi, H., Collins, D. L., Evans, A. C., & Karama, S. (2014). Investigating the relation between striatal volume and IQ. *Brain Imaging and Behavior*, *8*, 52–59. https://doi.org/10.1007/s11682-013-9242-3

Mackey, A. P., Miller Singley, A. T., & Bunge, S. A. (2013). Intensive reasoning training alters patterns of brain connectivity at rest. *Journal of Neuroscience*, *33*, 4796–4803. https://doi.org/10.1523/JNEUROSCI.4141-12.2013

Mackintosh, N. (2011). *IQ and human intelligence*. Oxford: Oxford University Press.

Mandelman, S. D., & Grigorenko, E. L. (2012). BDNF Val66Met and cognition: All, none, or some? A meta-analysis of the genetic association. *Genes, Brain and Behavior*, *11*, 127–136. https://doi.org/10.1111/j.1601-183X.2011.00738.x

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*(7265), 747–753.

Marioni, R. E., McRae, A. F., Bressler, J., Colicino, E., Hannon, E., Li, S., et al. (2018). Meta-analysis of epigenome-wide association studies of cognitive abilities. *Molecular Psychiatry*, *23*(11), 2133–2144. https://doi.org/10.1038/s41380-017-0008-y

McGue, M., & Christensen, K. (2001). The heritability of cognitive functioning in very old adults: Evidence from Danish twins aged 75 years and older. *Psychology and Aging*, *16*, 272–280. https://doi.org/10.1037//0882–7974.16.2.272

McGue, M., & Christensen, K. (2002). The heritability of level and rate-of-change in cognitive functioning in Danish twins aged 70 years and older. *Experimental Aging Research*, *28*, 435–451. https://doi.org/10.1080/03610730290080416

McKay, J. A., Groom, A., Potter, C., Coneyworth, L. J., Ford, D., Mathers, J. C., et al. (2012). Genetic and non-genetic influences during pregnancy on infant global and site specific DNA methylation: Role for golate gene variants and vitamin B12. *PLoS One*, *7*, e33290. https://doi.org/10.1371/journal.pone.0033290

Miller, C. A., & Sweatt, J. D. (2007). Covalent modification of DNA regulates memory formation. *Neuron*, *53*, 857–869. https://dx.doi.org/10.1016/j.neuron.2007.02.022

Nativio, R., Donahue, G., Berson, A., Lan, Y. M., Amlie-Wolf, A., Tuzer, F., et al. (2018). Dysregulation of the epigenetic landscape of normal aging in Alzheimer's disease. *Nature Neuroscience*, *21*, 497–505. https://doi.org/10.1038/s41593-018-0101-9

Neubert, F.-X., Mars, R. B., Sallet, J., & Rushworth, M. F. S. (2015). Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proceedings of the National Academy of Sciences*, *112*, E2695–E2704. https://doi.org/10.1073/pnas.1410767112

Oberauer, K., Süβ, H.-M., Wilhelm, O., & Wittmann, W. W. (2008). Which working memory functions predict intelligence? *Intelligence*, *36*, 641–652. https://doi.org/10.1016/j.intell.2008.01.007

Oliveira, A. M. M., Hemstedt, T. J., & Bading, H. (2012). Rescue of aging-associated decline in Dnmt3a2 expression restores cognitive abilities. *Nature Neuroscience*, *15*, 1111–1113.

Oliveira, A. M. M., Hemstedt, T. J., Freitag, H. E., & Bading, H. (2015). Dnmt3a2: A hub for enhancing cognitive functions. *Molecular Psychiatry*, *21*, 1130–1136. https://doi.org/10.1038/mp.2015.175

Penner, M. R., Roth, T. L., Chawla, M. K., Hoang, L. T., Roth, E. D., Lubin, F. D., et al. (2011). Age-related changes in Arc transcription and DNA methylation within the hippocampus. *Neurobiology of Aging*, *32*, 2198–2210. https://doi.org/10.1016/j.neurobiolaging.2010.01.009

Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature Reviews Genetics*, *19*, 148–159. https://doi.org/10.1038/nrg.2017.104

Presente, A., Boyles, R. S., Serway, C. N., de Belle, J. S., & Andres, A. J. (2004). Notch is required for long-term memory in Drosophila. *Proceedings of the National Academy of Sciences*, *101*(6), 1764–1768. https://doi.org/10.1073/pnas.0308259100

Ptashne, M. (2007). On the use of the word "epigenetic." *Current Biology*, *17*(7), R233–236. https://doi.org/10.1016/j.cub.2007.02.030

Reolon, G. K., Maurmann, N., Werenicz, A., Garcia, V. A., Schroder, N., Wood, M. A., et al. (2011). Posttraining systemic administration of the histone deacetylase inhibitor sodium butyrate ameliorates aging-related memory decline in rats. *Behavioural Brain Research*, *221*(1), 329–332. https://doi.org/10.1016/j.bbr.2011.03.033

Rimfeld, K., Shakeshaft, N. G., Malanchini, M., Rodic, M., Selzam, S., Schofield, K., et al. (2017). Phenotypic and genetic evidence for a unifactorial structure of spatial abilities. *Proceedings of the National Academy of Sciences*, *114*, 2777–2782. https://doi.org/10.1073/pnas.1607883114

Rogowski, K., van Dijk, J., Magiera, M. M., Bosc, C., Deloulme, J.-C., Bosson, A., et al. (2010). A family of protein-deglutamylating enzymes associated with neurodegeneration. *Cell*, *143*, 564–578. https://doi.org/10.1016/j.cell.2010.10.014

Roth, C., Magnus, P., Schjolberg, S., Stoltenberg, C., Suren, P., McKeague, I. W., et al. (2011). Folic acid supplements in pregnancy and severe language delay in children. *Journal of the American Medical Association*, *306*, 1566–1573.

Rush, E. C., Katre, P., & Yajnik, C. S. (2014). Vitamin B12: One carbon metabolism, fetal growth and programming for chronic disease. *European Journal of Clinical Nutrition*, *68*, 2–7.

Sakakibara, E., Takizawa, R., Kawakubo, Y., Kuwabara, H., Kono, T., Hamada, K., et al. (2018). Genetic influences on prefrontal activation during a verbal fluency task in children: A twin study using near-infrared spectroscopy. *Brain and Behavior*, *8*, e00980. https://doi.org/10.1002/brb3.980

Sauce, B., & Matzel, L. D. (2018). The paradox of intelligence: Heritability and malleability coexist in hidden gene-environment interplay. *Psychological Bulletin*, *144*, 26–47. https://doi.org/10.1037/bul0000131

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, *50*, 912–919. https://doi.org/10.1038/s41588-018-0152-6

Schiepers, O. J. G., van Boxtel, M. P. J., de Groot, R. H. M., Jolles, J., Kok, F. J., Verhoef, P., et al. (2011). DNA methylation and cognitive functioning in healthy older adults. *British Journal of Nutrition*, *107*, 744–748. https://doi.org/10.1017/S0007114511003576

Schlagenhauf, F., Rapp, M. A., Huys, Q. J. M., Beck, A., Wüstenberg, T., Deserno, L., et al. (2013). Ventral striatal prediction error signaling is associated with dopamine synthesis capacity and fluid intelligence. *Human Brain Mapping*, *34*, 1490–1499. https://doi.org/10.1002/hbm.22000

Schott, B. H., Minuzzi, L., Krebs, R. M., Elmenhorst, D., Lang, M., Winz, O. H., et al. (2008). Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *Journal of Neuroscience*, *28*, 14311–14319. https://doi.org/10.1523/jneurosci.2058-08.2008

Shah, S., McRae, A. F., Marioni, R. E., Harris, S. E., Gibson, J., Henders, A. K., et al. (2014). Genetic and environmental exposures constrain epigenetic drift over the human life course. *Genome Research*, *24*, 1725–1733. https://doi.org/10.1101/gr.176933.114

Shlyueva, D., Stampfel, G., & Stark, A. (2014). Transcriptional enhancers: From properties to genome-wide predictions. *Nature Reviews Genetics*, *15*, 272–286. https://doi.org/10.1038/nrg3682

Sinn, D. I., Kim, S. J., Chu, K., Jung, K. H., Lee, S. T., Song, E. C., et al. (2007). Valproic acid-mediated neuroprotection in intracerebral hemorrhage via histone deacetylase inhibition and transcriptional activation. *Neurobiology of Disease*, *26*, 464–472. https://doi.org/10.1016/j.nbd.2007.02.006

Smith, A. K., Kilaru, V., Klengel, T., Mercer, K. B., Bradley, B., Conneely, K. N., et al. (2015). DNA extracted from saliva for methylation studies of psychiatric traits: Evidence for tissue specificity and relatedness to brain. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *168B*, 36–44.

Sniekers, S., Stringer, S., Watanabe, K., Jansen, P. R., Coleman, J. R. I., Krapohl, E., et al. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature Genetics*, *49*, 1107–1112. https://doi.org/10.1038/ng.3869

Spengler, M., Gottschling, J., Hahn, E., Tucker-Drob, E. M., Harzer, C., & Spinath, F. M. (2018). Does the heritability of cognitive abilities vary as a function of parental education? Evidence from a German twin sample. *PLoS One*, *13*, 1–15. https://doi.org/10.1371/journal.pone.0196597

Starnawska, A., Tan, Q., McGue, M., Mors, O., Borglum, A. D., Christensen, K., et al. (2017). Epigenome-wide association study of cognitive functioning in middle-aged monozygotic twins. *Frontiers in Aging Neuroscience*, *9*, 413. https://doi.org/10.3389/fnagi.2017.00413

Stefanko, D. P., Barrett, R. M., Ly, A. R., Reolon, G. K., & Wood, M. A. (2009). Modulation of long-term memory for object recognition via HDAC inhibition. *Proceedings of the National Academy of Sciences*, *106*, 9447–9452. https://doi.org/10.1073/pnas.0903964106

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. New York: Cambridge University Press.

Surén, P., Roth, C., Bresnahan, M., Haugen, M., Hornig, M., Hirtz, D., et al. (2013). Association between maternal use of folic acid supplements and risk of autism spectrum disorders in children. *Journal of the American Medical Association*, *309*, 570–577. https://doi.org/10.1001/jama.2012.155925

Trampush, J. W., Yang, M. L. Z., Yu, J., Knowles, E., Davies, G., Liewald, D. C., et al. (2017). GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: A report from the COGENT consortium. *Molecular Psychiatry*, *22*, 336–345. https://doi.org/10.1038/mp.2016.244

Ursini, G., Bollati, V., Fazio, L., Porcelli, A., Iacovelli, L., Catalani, A., et al. (2011). Stress-related methylation of the catechol-O-methyltransferase Val 158 allele predicts human prefrontal cognition and activity. *Journal of Neuroscience*, *31*, 6692–6698. https://doi.org/10.1523/JNEUROSCI.6631-10.2011

Vaessen, T., Hernaus, D., Myin-Germeys, I., & van Amelsvoort, T. (2015). The dopaminergic response to acute stress in health and psychopathology: A systematic review. *Neuroscience and Biobehavioral Reviews*, *56*, 241–251. https://doi.org/10.1016/j.neubiorev.2015.07.008

Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842–861.

Vecsey, C. G., Hawk, J. D., Lattal, K. M., Stein, J. M., Fabian, S. A., Attner, M. A., et al. (2007). Histone deacetylase inhibitors enhance memory and synaptic plasticity via CREB:CBP-dependent transcriptional activation. *Journal of Neuroscience*, *27*, 6128–6140. https://doi.org/10.1523/JNEUROSCI.0296-07.2007

Vernes, S. C., Newbury, D. F., Abrahams, B. S., Winchester, L., Nicod, J., Groszer, M., et al. (2008). A functional genetic link between distinct developmental language disorders. *New England Journal of Medicine*, *359*(22), 2337–2345. https://doi.org/10.1056/NEJMoa0802828

Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane- Robinson, C., Allis, C. D., & Workman, J. L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *EMBO Journal*, *15*(10), 2508–2518.

Villamor, E., Rifas-Shiman, S. L., Gillman, M. W., & Oken, E. (2012). Maternal intake of methyl-donor nutrients and child cognition at 3 years of age. *Paediatric and Perinatal Epidemiology*, *26*, 328–335. https://doi.org/10.1111/j.1365-3016.2012.01264.x

Vitolo, J. M., Thiriet, C., & Hayes, J. J. (2000). The H3-H4 N-terminal tail domains are the primary mediators of transcription factor IIIA access to 5S DNA within a nucleosome. *Molecular and Cellular Biology*, *20*(6), 2167–2175.

von Stumm, S., & Plomin, R. (2015). Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence*, *48*, 30–36. https://doi.org/10.1016/j.intell.2014.10.002

Vukojevic, V., Kolassa, I. T., Fastenrath, M., Gschwind, L., Spalek, K., Milnik, A., et al. (2014). Epigenetic modification of the glucocorticoid receptor gene is linked to traumatic memory and post-traumatic stress disorder risk in genocide survivors. *Journal of Neuroscience*, *34*, 10274–10284. https://doi.org/10.1523/JNEUROSCI.1526-14.2014

Walton, E., Hass, J., Liu, J., Roffman, J. L., Bernardoni, F., Roessner, V., et al. (2016). Correspondence of DNA methylation between blood and brain tissue and its application to schizophrenia research. *Schizophrenia Bulletin*, *42*, 406–414. https://doi.org/10.1093/schbul/sbv074

Walton, E., Liu, J. Y., Hass, J., White, T., Scholz, M., Roessner, V., et al. (2014). MB-COMT promoter DNA methylation is associated with working-memory processing in schizophrenia patients and healthy controls. *Epigenetics*, *9*, 1101–1107. https://doi.org/10.4161/epi.29223

Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature Genetics*, *40*, 897–903.

Weaving, L. S., Ellaway, C. J., Gecz, J., & Christodoulou, J. (2005). Rett syndrome: Clinical review and genetic update. *Journal of Medical Genetics*, *42*, 1–7. https://doi.org/10.1136/jmg.2004.027730

Weissberger, G. H., Nation, D. A., Nguyen, C. P., Bondi, M. W., & Han, S. D. (2018). Meta-analysis of cognitive ability differences by apolipoprotein e genotype in young humans. *Neuroscience and Biobehavioral Reviews*, *94*, 49–58.

Whitehouse, A. J., Bishop, D. V. M., Ang, Q. W., Pennell, C. E., & Fisher, S. E. (2011). CNTNAP2 variants affect early language development in the general population. *Genes, Brain and Behavior*, *10*, 451–456.

Winick-Ng, W., & Rylett, R. J. (2018). Into the fourth dimension: Dysregulation of genome architecture in aging and Alzheimer’s disease. *Frontiers in Molecular Neuroscience*, 11. https://doi.org/10.3389/fnmol.2018.00060

Wu, L., Sun, T., Kobayashi, K., Gao, P., & Griffin, J. D. (2002). Identification of a family of mastermind-like transcriptional coactivators for mammalian notch receptors. *Molecular and Cellular Biology*, *22*(21), 7688–7700.

Xie, W., & Ren, B. (2013). Enhancing pluripotency and lineage specification. *Science*, *341*, 245–247. https://doi.org/10.1126/science.1236254

Yang, A. W., Sachs, A. J., & Nystuen, A. M. (2015). Deletion of Inpp5a causes ataxia and cerebellar degeneration in mice. *Neurogenetics*, *16*(4), 277–285. https://doi.org/10.1007/s10048-015-0450-4

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*, 565–569. https://doi.org/10.1038/ng.608

Zhang, R.-R., Cui, Q.-Y., Murai, K., Lim, Y. C., Smith, Z. D., Jin, S., et al. (2013). Tet1 regulates adult hippocampal neurogenesis and cognition. *Cell Stem Cell*, *13*, 237–245. https://doi.org/10.1016/j.stem.2013.05.006

7 Intelligence in Infancy

Marc H. Bornstein

Introduction

Can little babies be said to be "intelligent"?
You decide after reading this account.

This brief chapter peeks into infancy and addresses the compelling question of whether infants are "intelligent." The chapter also attends to the prediction of future mental development from infancy. First, the chapter tackles the perennially intransigent challenges of defining infancy and intelligence. Next, the chapter reviews the history of infancy study from the point of view of what we thought we knew about infant intelligence. Then, the chapter draws the reader into an intuitive perspective on what might be some everyday intelligent behaviors on the part of infants; that perspective is buttressed with what has been learned from scientific investigations in the laboratory. This work is subsequently elaborated on with passing reference to new paradigms followed by a more intense focus on two prominent interrelated methods and measures of studying cognition in infants: novelty preference and habituation – attention to novel stimulation and the decline in attention to stimulation that is increasingly familiar. Their interpretation qua measures of cognition is supported with evidence from studies of concurrent and predictive validity. The chapter affords a new view of infant intelligence and concludes with some comments and thoughts about the future promise of that new view.

"Infancy" and "Intelligence"

Let's define the two principal terms of the chapter and some surprising aspects of each.

Human Infants are an Altricial Species

Infancy defines the period of extrauterine life between birth and the emergence of language as approximately one and a half to two years into childhood. Infancy appears to be a distinct stage of life based on biological, cognitive, and social data, and people worldwide and throughout the ages have represented infancy as typically an initial age of the human life span. Many outstanding developmental theorists – Freud, Piaget, Erikson, and Werner, for example – championed stage theories of development, and all of them identified infancy as one. Iconographically, infants symbolize origins and beginnings.

Language is a signal marker of infancy and of intelligence. A defining characteristic of infancy is that infants neither understand language nor can they speak. Indeed, our generic terms "infant" and "baby" both have origins in language deficit–related concepts. The word "infant" derives from the Latin *in* + *fans*, translated literally as "nonspeaker," and the word "baby" shares a Middle English root with "babble." The end of infancy is marked traditionally by the advent of language.

First and foremost, then, infants are by definition nonverbal (but not, of course, noncommunicative), and they are also, especially in the early months of life, motorically undeveloped, emotionally labile, and subject to rapid under-regulated fluctuations in behavioral state. Other infant characteristics fuel a view of infant incompetence, such as infants' short attention spans and their limited response repertoires. At the same time, infants are also notoriously uncooperative and perversely unmotivated to perform or conform. Unreliability is, therefore, inherent to this early phase of life. However, human infancy has long offered a certain romantic and simultaneously enigmatic attraction to philosophers, scientists, and parents. All of us have been infants, yet perceptions, thoughts, and feelings of our infancy are seemingly lost to us. These pervasive infant characteristics aggregate to give us pause in attributing intelligence to infants.

Last, but not least, human infants are totally dependent on caregivers . . . even for their survival. The newborn foal will stand in the hour after delivery and soon canter, and the newborn chick will pip on its shell to hatch, feed itself on the internal yolk sac, and forage on its own soon after hatching. However, the "altricial" newborn human cannot walk, talk, thermoregulate, or even nourish themselves without the aid of a competent caregiver. As the psychoanalyst Winnicott (1965, p. 39) once enigmatically mused, "There is no such thing as an infant." That is, infants only exist in a symbiotic system with a caregiver. This altricial status of the human infant seems to close the door on any attribution of "intelligence" to infants.

What Is Intelligence?

No one really knows what intelligence is, and vital as intelligence is to human self-understanding the concept has successfully eluded any universally agreed-on definition (Bornstein & Putnick, 2019; Sternberg, Chapter 1 in this volume). So "intelligence" is open to interpretation. Much, then, is in the eye of the beholder: To the educator, intelligence might be the ability to learn; to the biologist, intelligence might be the ability to adapt. Psychologists have debated the definition and scope of intelligence for more than a century. Alfred Binet (1905) referred to intelligence as "Judgment, otherwise called 'good sense', 'practical sense', 'initiative', the faculty of adapting one's self to circumstances . . . " and David Wechsler (1958, p. 7), author of the most widely used tests of intelligence today, stated that intelligence is "the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his [sic] environment." When two dozen prominent theorists were asked to define intelligence, they offered two dozen definitions (Sternberg & Detterman, 1986). Later, Sattler (1992) discussed more than twenty separate (yet somewhat overlapping) definitions of intelligence. One 1997 "consensus" statement (endorsed by 52 "experts" – out of 131 invited to sign) read as follows: "Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (Gottfredson, 1997, p. 13). Surveying dictionaries, encyclopedias, and professional associations, Legg and Hutter (2007) collected seventy-plus definitions but opined that "compiling a complete list would be impossible" (p. 17). In Wikipedia, intelligence "can be more generally described as the ability to perceive or infer information, and to retain it as knowledge to be applied towards adaptive behaviors within an environment or context" (https://en.wikipedia.org/wiki/Intelligence).

In modern terms, intelligence (cognition) reflects efficiency of information processing and mental representation. Language, reasoning, problem-solving, and memory skills all are integral to human intelligence (Cooper & Regan, 1986; Dempster, 1991; Hunt, 1983). Two characteristics of intelligence agreed on by (virtually) all authorities are, first, that intelligence cannot be seen, smelled, tasted, or palpated but, second, that intelligence can be measured. Intelligence "tests," dating back to Binet and his colleagues, are based on the idea that it is possible to appraise behavior that is consensually recognized as intelligent and that what intelligence tests measure relates conceptually to learning, comprehension, capacity to adjust, and the like, as well as practically to educational attainment, occupational prestige, and other concrete indices of achievement. Correlations between intelligence test scores and formal tests of reading, mathematics, or other subjects, and between intelligence test scores and school performance or grades, range respectively between 0.40 and 0.70 (Brody, 1992; Jensen, 1980; Vernon, 1947); these correlations are obtained, at least partly, because intelligence tests require reading, mathematics, etc., overlapping the contents of achievement tests.

Does Intelligence Begin in Infancy?

Hidden Clues in Full Sight

On first observation, infants appear disorganized and erratic. At any given moment, babies seem to be constantly moving their eyes or hands or feet without apparent purpose. Over longer periods, they appear to shift irregularly and unpredictably between alertness and sleep. However, on close and consistent inspection, infants are actually not quite so irregular and unpredictable. Infant activity is organized at fast, medium, and slow rhythms. Some actions regularly cycle at high frequencies, perhaps once or more every second. Heartbeats, breathing, and sucking exemplify fast biological rhythms that maintain life, and kicking and rocking illustrate other fast-cycling behaviors. General movements of the body cycle at intermediate rates, on the order of once every minute or two. States of waking, quiet sleep, and active sleep cycle at low frequencies in periods of hours. A snapshot of the infant at any one moment reveals apparent inchoate simultaneous and independent schedules of what are complex detectable cyclical patterns. By observing activity and carefully decomposing it, it is possible to detect regularities underlying infants' seeming randomness. In short, apparent irregularity is only just that, apparent, and much infant behavior is characterized by underlying predictable regularity.

Coming back to the preverbal status of infants, judgments of adults often fail to capture significant facts about infant verbal production, for example, where closer (machine spectrographic) analysis provides insight. Macken and Barton (1980) identified contrasts in infant speech that were not detected by parents or other adults, and might have taken a year or more before their production clarity improved to the point where the contrasts that the infants were voicing could be reliably perceived by adults.

These two homilies impart pointed lessons for a revised understanding of infancy and infant intelligence. In short, William James' (1890, p. 488) famous "blooming, buzzing confusion" attribution to infancy mistakes and cloaks order and competence.

A Bit of Infancy History

The first-ever written accounts of children were diary descriptions of infants in their natural settings set down by parents – so-called baby biographies (Darwin, 1877; Hall, 1891; Jaeger, 1985; Lamott, 2013; Mendelson, 1993; Preyer, 1882; Prochner and Doyon, 1997; Rousseau, 1762; Stern, 1990; Taine, 1877; Tiedemann, 1787; Wallace, Franklin, and Keegan, 1994). Darwin, who introduced evolutionary theory in 1859 with the *Origin of Species*, subsequently published observations he had made well before (in the early 1840s) of the infancy of his firstborn son, William Erasmus, nicknamed "Doddy." Darwin's (1877) "Biographical sketch of an infant" gave great impetus to infancy studies (Lerner et al., 2015). In succeeding years, baby biographies grew in popularity around the world – whether they were scientific documents, parents' personal records, or illustrations of educational practices – and they still appear today. Perhaps the most influential baby biographer was Piaget (1952), whose writings and theorizing refer chiefly to observations of his own three infants, Jacqueline, Laurent, and Lucienne.

Baby biographies of the nineteenth and twentieth centuries constituted the first close examinations of infants, and the evidence developed from those records had two important implications relevant here. First, infants of different ages showed themselves to be competent in a variety of ways, and, second, baby biographies uncovered wide variation among infants of the same age, suggesting that individual infants vary among themselves in all domains of development. On these twin bases, the early twentieth century witnessed several attempts in Europe and in the United States to develop standardized tests for infants, and for years the sequences, scales, and schedules of Bayley; Buhler and Hetzer; Cattell; Gesell; Griffiths; Shirley; and

Užgiris and Hunt proved valuable in defining normative infant motor, sensory, cognitive, and socioemotional development. Their traditional infant tests have since evolved as screening tools, useful for diagnostic and clinical purposes. Typically, they represent the infant's developmental status as an aggregate score (the average obtained across different developmental domains). For example, the Bayley Scales of Infant Development III (BSID III; Bayley, 2005) report cognitive, motor, and language as well as social-emotional and adaptive behavior scales.

These constructive developments (important and insightful as they were) paradoxically led to two problems for a thesis of infant intelligence. First, the kinds of items put to young infants in those traditional tests largely tap capacities of questionable cognitive content; sensory discriminations, motor milestones, and emotional reactions do not qualify as really "cognitive" in nature. Many traditional infant test items also rely on imitation of simple actions. Very different items are used in evaluating intelligence after infancy, such as skills related to language, reasoning, and memory. Thus, the underlying constructs or domains assessed by traditional tests in infancy and by intelligence tests in childhood bear little or no conceptual (face-valid) relation to one another. We would not be comparing apples and apples or apples and oranges, but apples and cars.

Second, traditional infant tests were for the most part modeled after Binet-type assessments of intelligence originally developed for older children and adults. Binet's test (and the Stanford-Binet, Wechsler scales, and others that followed it) use external criteria (classroom performance, educational achievement) to validate test performance. In infants, however, there is no obvious external index with which to compare test performance to assess validity. Rather, validity of infant tests has typically been evaluated by documenting associations between infants' performance early in life with their IQ test performance in childhood or even years later in maturity. The presumption in this time-lagged analysis is that, if individuals who perform well on infant tests also do well on standardized intelligence tests as children or adults, then the original tests likely constitute assessments of "intelligence" in infancy. This general approach relies on the so-called predictive validity of the infant measures. But see the first problem.

Unsurprisingly, the predictive validities of infant tests have consistently proved poor in normal populations (Bornstein & Sigman, 1986; Colombo, 1993; Fagan & Singer, 1983; Kopp & McCall, 1980; but see Siegel, 1989). Bayley's (1949) own classic longitudinal study exemplifies both the findings and conclusions that characterized the early tradition in predictive infant testing. Bayley followed twenty-seven Berkeley Growth Study children from the first three months of life to eighteen years of age and then correlated their scores on her California First-Year Mental Scale (later the BSID) with their later intelligence test scores. Bayley reported essentially no correlation between test performance in the first three to four years of life and intelligence test performance at eighteen years. Only after children reached about six to seven years of age (when they were taking the Stanford-Binet) did the correlation between child scores and eventual adult scores rise to a respectable 0.68, eventually reaching about 0.85 between eleven and eighteen years.

Along these lines, Yu and colleagues (2018) examined the stability of a latent construct of intelligence composed of three assessments of "intellectual performance" at each of four developmental periods over a sixteen-year interval from infancy (1, 1.5, and 2 years old), preschool (2.5, 3, and 3.5 years old), childhood (6, 7, and 8 years old), and adolescence (12, 15, and 17 years old). Notably, the latent variable at 1, 1.5, and 2 years was constructed of the Bayley Scales of Infant Development, that at 2.5, 3, and 3.5 years of the McCarthy Scales of Children's Abilities, that at 6, 7, 8, and 12 years of the Wechsler Intelligence Scales for Children – Revised, that at 15 years of the Wechsler Intelligence Scale for Children – Third Edition, and that at 17 years of the Wechsler Adult Intelligence Scale – Revised. The cross-time correlations exhibited stability: For instance, the infant latent variable showed a strong cross-time correlation with that at preschool ($r = 0.91$) and moderate correlations with the childhood and adolescent ones ($r = 0.69$ and 0.57, respectively). It should be noted with respect to the concerns of this chapter, however, that the "infancy" latent variable began at one year and included data up to two years. Furthermore, the zero-order correlations between one-year BSID and the twelve-, fifteen-, and seventeen-year Wechsler scores were all nonsignificant.

Traditional developmental assessments, like the BSID, administered to children older than twelve months possess somewhat higher predictive power than do tests administered during the first year of life; infant scores may predict some kinds of nonverbal performance as opposed to intelligence; and predictive relations are generally stronger for biologically at-risk infants than for typically developing infants (Blaga et al., 2009; Humphreys & Davey, 1988; Siegel, 1989). For example, in preterms the BSID Psychomotor Development Index (PDI; not the Mental Development Index) in the first year of life has been reported to predict expressive language at 2–4 years (Siegel, 1989), eight-month PDI, and seven-year intelligence (Broman, 1989).

Fagan and Singer (1983) reviewed the results of 101 studies published to that date in which attempts had been made to predict IQs in childhood from tests given during the first year of life. They found that the average correlations for fifty groups of normal infants and fifty-one groups of infants expected to be at risk for later intellectual disabilities were 0.14 and 0.21, respectively. These meta-analysis results seem to confirm that infant tests have little validity for the prediction of later IQ, their shared variance being 2 percent and 4 percent, respectively. Subsequent reviews by Bornstein and Sigman (1986), Anastasi and Urbina (1997), Chen and Siegler (2000), Sternberg, Grigorenko, and Bundy (2001), and Hetherington and colleagues (2006) arrived at the same conclusion. With a few exceptions, scores achieved in the first year of life on traditional infant tests fail to provide meaningful information about later intelligence. (Nonetheless, many investigators continue to interpret BSID as a general measure of infant cognition.)

The lack of predictive validity of traditional infant tests led to several significant broad conclusions about intelligence in infancy. Some theorists opined that this dissociation reflects the lack of a stable general intelligence factor in mental life. Others argued that, if such a general factor exists, it is just not fixed or stable across the early part of the life span. Perhaps mental growth follows a stagelike progression wherein intelligence in infancy differs qualitatively from intelligence in maturity. Bayley (1955) and others (McCall, Hogarty, & Hurlburt, 1972) assumed that the growth of intelligence instantiates such a "discontinuous" process. You have no kind, or one kind, of intelligence as an infant and another kind of intelligence as a child. This idea of "discontinuity" in intelligence was conceptually appealing too, because it corresponded with, and was influenced by, similar intellectually transcendent twentieth-century views of intelligence advanced by Piaget (1952), viz. that the very nature of intelligence changes with age. These several interpretations of intelligence differ from one another, but all are based on the same data showing a lack of cross-age association or predictability of intelligence from traditional infant tests. Certainly with development a person's knowledge base (= intelligence?) changes. The case for infant intelligence again seemed closed.

Before throwing out the baby with the bathwater, however, some researchers paused to reconsider the two problems listed earlier: exactly what traditional infant tests measure and whether there is any reason to imbue them with predictive validity in the first place. Recall the first problem. Rather than tap sensory ability, motor requirements, affective components, and imitation, assessments in infancy might tap basic cognitive skills, such as information processing or mental representation, that underlie intelligence and do so in ways that are psychometrically sound and that transcend and conceptually parallel or serve as foundations for mental functions in childhood. We shall come to some candidates shortly.

Sensitivity to and Profiting from Experience qua Intelligence

Although they do not always appear so, even very young infants have an active mental life. They are constantly learning and developing new ideas. This is not the traditional view, but as with their manifest activity and preverbal articulations, it now appears that the myriad capacities infants demonstrate every day in homes and in laboratories around the world betray competencies and have consequences that link to their later mental development. If learning from experience necessarily involves intelligence, if successful adaptation does, then infants can rightly be said to be intelligent because infants are particularly susceptible and responsive to their experiences and so profit from them. Theoreticians and researchers have long supposed that the child's earliest experiences shape mental development (Plato, 1970; ca. 355 BC). Psychoanalysis, ethology, neuropsychology, behaviorism, constructivism, attachment, and systems theories, as well as quotidian and laboratory empiricism, support an "early experience" effects model of mental development.

Theory

No lesser figures in the history of science than Charles Darwin (1877) and Sigmund Freud (1949) initiated scientific observations of infants and theoretical speculations about the role of infancy in the balance of development. Psychoanalytic theory portrayed infant experiences as exerting enduring influences: From oral sensory experiences, Erikson (1950) suggested, infants develop basic trust or mistrust in others, and which infants develop what has implications for the way they negotiate future stages across the life span. For their part, ethologists and embryologists (Lorenz, 1935/1970; Tinbergen, 1951) emphasized the lifelong legacy of early experiences (as in sensitive periods; Bornstein, 1989a). Demonstrations of sensitive periods in lower animals (imprinting, as in the "ugly duckling") award biological and scientific credibility to the lasting potency of experiences in infancy, and this feature has been painted into many portraits of human infant growth and development. Thus, *neoteny* (the prolongation of infancy), which is especially characteristic of human beings, is thought to have special adaptive significance (Gould, 1985) because it allows for enhanced experiential influence and prolonged learning (Bjorklund & Myers, 2019). Insofar as early behavior patterns lay the foundation for later ones, pride of place for infancy was also emphasized by learning theorists who stressed the significance of infant experiences for the balance of the life course (Dollard & Miller, 1950; Watson, 1924/1970). For those theorists, learning in infancy is important because it occurs first and promotes easy and rapid later learning and because early and simpler behavior patterns underlie later and more complex ones. Students of the constructivist school of development, beginning with Piaget (1952), theorized that capacities of later life build on developments earlier in life and that infants actively participate in their own development. Bowlby, Ainsworth, and their successors (Sroufe et al., 2005) in turn theorized that infants develop internal working models of their relationships and, reflecting them, become attached to those persons who are consistent, predictable, and appropriate in responding to their signals as well as to their needs (Cummings & Warmuth, 2019). Attachment experiences in infancy auger future development of cognition, personality, and social relationships. From the perspective of systems theory, development consists of hierarchically organized abilities that incorporate earlier emerging ones (Lewontin, 2005). Thus, an implication of contemporary relational systems perspectives is that earlier emerging characteristics in development lay foundations for, and so likely exert impact on, later appearing characteristics (Overton, 2015).

In brief, the characteristics developed and acquired in infancy are in the view of theorists of many stripes formative and fundamental in the sense that they endure or (at least) constitute building blocks that later developments or experiences exploit or modify.

Quotidian Empiricism

What do those who know infants best – their parents – think about infant intelligence? When their infants are only one month of age, 99 percent of mothers believe that their babies can express interest (Johnson et al., 1982). These judgments likely reflect infants' demonstrative everyday capacities. Somehow, someway, being with an infant on a continuing basis impresses people with something about infants' "competencies."

Caregivers exercise continuing and powerful influences over infants in their roles as ministrators, socializers, and educators. Didactic caregiving consists of the variety of strategies parents use in stimulating infants to engage with and understand the environment. Didactics include focusing the infant's attention on properties, objects, or events in the surround; introducing, mediating, and interpreting the external world; describing and demonstrating; as well as provoking or providing opportunities to observe, to imitate, and to learn. Material caregiving includes those ways in which parents provision and organize the infant's physical world. Adults are responsible for the number and variety of inanimate objects (toys, books) available to the infant, level of ambient stimulation, and overall physical dimensions of babies' experiences. Parents of infants around the globe engage in these didactic and material forms of caregiving and, presumably, would not do so unless they believed at some level that infants benefit from such interactions. Infants do: Mothers encouraging their two-month-olds to attend to properties, objects, and events in the environment predicts infants' tactual exploration of objects at five months over and above two- to five-month stability in infant tactual exploration and contemporaneous five-month maternal stimulation; maternal didactic stimulation when their infant is two through five months cumulates to predict infant non-distress vocalization when the infant is five months over and above similar controls; and mothers who speak more, prompt more, and respond more during the first year of their infants' lives have four-year-old children who score higher in standardized evaluations of language and cognition (Bornstein, 1985; Bornstein & Tamis-LeMonda, 1990; Bornstein, Tamis-LeMonda, & Haynes, 1999; Nicely, Tamis-LeMonda, & Bornstein, 1999). Early (3-month) infant participation in interactions with their mothers predicts their later (6-year) intelligence (Coates & Lewis, 1984), and infants' interactions at seven months predict their expressive and productive vocabulary at fourteen months (Lunden & Silven, 2011). Even features of the parent-outfitted material environment appear to influence infant mental development directly (Wachs & Chan, 1986): New toys and changing room decorations promote child language acquisition in and of themselves, separate and apart from parental language.

Sensitive parents also tailor their behaviors to match their infants' developmental progress (Bornstein, 2013), for example by speaking more and providing more didactic experiences as infants age across the first year (Bornstein et al., 1992; Bornstein & Tamis-LeMonda, 1990; Klein, 1988). Mothers' single-word utterances are just those that appear earliest in their children's vocabularies (Chapman, 1981). Indeed, parents are sensitive to infant age and to increasing infant capacity or performance (Bellinger, 1980): The mean length of mothers' utterances tends to match the mean length of utterances of their one-and-a-half-year-olds (McLaughlin et al., 1983).

A common assumption in parenting is that the overall level of parental involvement or stimulation affects the overall level of infant development (Maccoby & Martin, 1983). An illustration of this simple "main effects" model asserts that the amount of language infants command is determined (at least to some degree) by the amount of language infants hear (Hart & Risley, 1995, 1999). However, increasing evidence suggests that complex and more nuanced processes explain parenting effects on infant development. First, specific (rather than general) parenting cognitions and practices appear to relate concurrently and predictively to specific (rather than general) infant accomplishments. The specificity principle states that specific parent-provided experiences at specific times exert specific effects over specific aspects of infant development in specific ways (Bornstein, 1989b, 2015). This principle gives additional evidence of infant perspicacity.

First language reflects the child's early and rich exposure to the parent-provided target language environment as much as it does competencies that the child brings to learning. Parent-provided experiences swiftly and surely channel early speech development toward the adult target language (de Boysson-Bardies & Vihman, 1991). In the space of approximately two years, infants master rudiments of language, often even without explicit instruction or noticeable effort, but they always speak the language to which they have been exposed. Indian infants adopted by American families and exposed only to English relearn Indian-dialect phonemes more quickly than their American peers who had never heard the Indian phonemes (Singh et al., 2011).

In one long-term prospective study extending from four months to four years, mother-child dyads were seen at three points in children's development (Bornstein, 1985); at four months infant habituation was assessed in the laboratory, at one year infant productive vocabulary was ascertained, and at four years children's intelligence was evaluated using the Wechsler series. At four months and at one year, mothers' didactic interactions with their children were recorded during home observations. Path analysis determined direct and unique longitudinal effects among the variables. Maternal didactic efforts in encouraging infant attention contributed to both one-year and four-year child cognitive outcomes. These findings were subsequently replicated and expanded in two shorter-term longitudinal follow-up studies (Tamis-LeMonda & Bornstein, 1989, 1993). Infants whose mothers show positive responses at twelve months have higher WPPSI (Wechsler Preschool and Primary Scale of Intelligence) IQ at four years (Pearson, Lightman, & Evans, 2011); and fathers' diverse vocabulary in interactions with their infants at six months predicts children's communication skills at fifteen months, after adjusting for infant developmental level at six months and other confounders (Pancsofar, Vernon-Feagans, & Family Life Project Investigators, 2010).

Experiences and habits developed in infancy are of crucial, and perhaps lifelong, significance; that is, the social orientations, personality styles, and intellectual predilections established at the start at least contribute to enduring patterns (Bornstein, 2014). Likely, the invisible first foundation and frame of the edifice are always and forever critical to its surviving structure.

Laboratory Empiricism

Whereas sensory systems are the means through which the brain receives information from the outside world, intelligence is conceptualized as computational systems that make sense of that information once it has been received irrespective of the means of reception (Davis et al., 2011).

The fact that infants are all those "un-s, in-s, and non-s" – unvolitional, uncooperative, unstable, nonverbal, motorically inept, and the like – once warded off all but philosophical speculation about them. However, developmental scientists have overcome multiple vexing problems of infants' formidable and intractable posture to extract information of all sorts from and about them. In the last approximately half-century, a revolution has taken place in infancy studies, fueled by technological and methodological advances. In consequence, we now know a great deal about babies' perceptions, thoughts, and feelings (Bornstein, Arterberry, & Lamb, 2014). Taken together, longitudinal studies are coming to weave infancy more tightly into the tapestry of life-span development (Bornstein, 2018). For good reason, longitudinality has been called a social-science Hubble telescope (Butz & Torrey, 2006). A bottom-up exploration reveals how and why.

The human brain is plastic, experience alters the brain, and brain change is indicative of deep learning, which might be a component of intelligence. Just three months of exposure to their own mother's face shapes occipital cortex evoked response potentials (ERPs) in infants (Bornstein, Arterberry, & Mash, 2013): By three months of age, infants' brains process their mothers' face as different to an appearance-matched stranger face. Infants who show more neural activity to native language contrasts at 7.5 months have larger vocabularies at twenty-four months, suggesting that infants who are more attuned to the sounds in their language are better at learning its words (Kuhl, 2009); auditory ERPs of English-exposed American infants in response to both Spanish and English voicing contrasts at eleven months of age predict the number of words children produce at eighteen through thirty months of age (Kuhl & Rivera-Gaxiola, 2008; see also Garcia-Sierra et al., 2011; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005); auditory ERPs at six and nine months predict language at three and four years (Choudhury & Benasich, 2011); and a discriminant function analysis of the brain waves of newborns to auditory signals predicts the classification of eight-year-old children into normal versus low language performance groups with about 80 percent accuracy (Molfese, 2000). Brain electrical activity at eight months also predicts working memory at four and a half years (Wolfe & Bell, 2007).

During infancy, the capacities to take in information through the major sensory channels and to attribute meaning to perceived information improve dramatically. Newborns are equipped to hear, to orient to, and to distinguish sounds. Newborns also identify particular speakers – notably the mother – right after birth (DeCasper & Spence, 1986; Kisilevsky et al., 2003), based on prenatal exposure to the maternal voice. By their preference reactions, newborns also give good evidence that they possess a developed sense of smell (Delauney-El Allam, Marlier, & Schaal, 2006; Goubet et al., 2002; Goubet, Strasbaugh, & Chesney, 2007; Steiner, 1979), and babies soon suck presumptively at the scent of their mothers (Porter & Levy, 1995; Porter & Winberg, 1999). Infants who nursed for six weeks from mothers who placed a balm with a distinctive odor on their nipples retained a representation of the odor for at least eighteen months after they had stopped nursing (Delauney-El Allam et al., 2010). Thus, infants actively scan the environment, pick up, encode, and process information, aggregate over, and remember their experiences (Bornstein & Colombo, 2012).

In one study, two- to four-month-old infants first viewed real-time images of their mothers interacting with them by means of closed-circuit television (Murray & Trevarthen, 1985). During this period, infants were seen to engage and react with normal interest and pleasure. Immediately afterward, the same infants watched the recording of the same interaction; this time, however, the infants exhibited signs of distress. Infants' negative reactions were considered to arise out of the lack of synchrony with their mothers that the babies expected. Only months-old infants are cognitively sensitive to the presence or the absence of appropriate real-time interactions.

Infants deliberately search for and use others' (parents') emotional (facial, vocal, gestural) expressions to help clarify and evaluate uncertain and novel events, termed social referencing (Kim & Kwak, 2011; Murray et al., 2008). Between nine and twelve months of age, infants look to mothers and fathers for such cues and are influenced to act by both positive and negative adult expressions (Dickstein & Parke, 1988; Hirshberg & Svejda, 1990). That qualities of caregivers' emotional expressions – such as distress, disgust, fear, anger – influence infant behavior seems sensible, given that the overriding message in a parent's emotional expressions is that an event is or is not dangerous or threatening to the baby. Infants not only play less with unusual toys when their mothers display disgust, instead of pleasure, about the toys, but when the same toys are presented later infants show the same responses, even though their mothers no longer pose emotional expressions but are instead silent and neutral (Hornik, Risenhoover, & Gunnar, 1987). Indeed, in uncertain situations infants will position themselves so they can keep their mother's face in view (Sorce & Emde, 1981). Parents embroiled in marital conflict may have difficulty attending to the sometimes-subtle signals infants use to communicate their needs. Infants from these homes learn that their caregivers are unreliable sources of information or assistance in stressful situations. For example, one-year-old infants are less likely to look to their maritally dissatisfied fathers for information or clarification in the face of stress or ambiguity than are infants of maritally satisfied fathers (Parke & Cookston, 2019). Infants respond emotionally to other affective expressions they observe in people, as when, for example, their caregivers are depressed (Manian & Bornstein, 2009). Infants as young as one year of age respond to emotional messages, showing signs of distress when witness to angry interactions between family members (Geangu et al., 2010, 2011; Hutman & Dapretto, 2009; R. A. Thompson, 2006). Furthermore, infants' perception of intentional agency at twelve months predicts their understanding of others' theory of mind, mental states, and beliefs as four-year-olds (Yamaguchi et al., 2009).

As suggested above, the term "infancy" applied to infants may be paradoxically misplaced. Speech perception at six months predicts language acquisition (word understanding, word production, and phrase understanding) at two years (Fernald, Perfors, & Marchman, 2006; Tsao, Liu, & Kuhl, 2004); and speech processing performance (segmenting words from fluent speech) before twelve months predicts language assessed at six years (Newman et al., 2006). Infants' early phonetic perception (Kuhl et al., 2005, 2008; Rivera-Gaxiola et al., 2005; Tsao et al., 2004), their pattern-detection skills for speech (Newman et al., 2006), mismatch responses to native-language sounds (Kuhl et al., 2008), and processing efficiency for words (Fernald et al., 2006) have all been linked to advanced later language abilities. Studies of communication skills and expressive vocabulary at eight and twelve months also show predictive relations to mother-reported child symbolic use of objects at two years (Reilly et al., 2009); and twelve-month-olds' vocabulary predicts their four-year verbal IQ (Blaga et al., 2009; Domsch, Lohaus, & Thomas, 2009). Infants who know more words at one year tend to know more words at two years, and two-year-olds who know more words are at a long-term advantage because knowing more words speeds learning to read, improves verbal comprehension, and eventuates in more advanced written language skills (Marchman & Fernald, 2008).

Intelligence implies memory. Many experiments now document infant recognition and recall memory. As examples, 6.5-month-olds sitting in a dark room who reached out on hearing a sound from the space in front of them reached out again when they returned to the laboratory and were played the same sound two years later; otherwise comparable two-year-olds without the infant experience did not react. In another study infants were exposed to a face when they were four months old and recognized the same face when they were eighteen months old (Bornstein et al., 2004).

In summary, many developmental theories and identified mechanisms of learning and development as well as laboratory and everyday empiricism support the contention that infants possess intelligence (or look like they do) and display their intelligence in many different ways. As indicated earlier, however, strict external criteria against which to judge infant intelligence are lacking. In lieu of such evidence, new and novel demonstrations of capabilities that look more like intelligence have been forged, and developmental scientists have returned to the vital second problem of predictive validity with renewed vigor.

Face and Predictive Validity in Infant Intelligence Redux

It could be that there is little or no predictive validity of intelligence from infancy. Armed with the understanding that infants might behave "intelligently" in many everyday and laboratory situations, however, it would be invalid (not to say shortsighted) to accept this null hypothesis about problem two in human mental development before trying measures that redress the raft of foregoing face-valid first-order problems with traditional infant tests.

It may be that the sensory, motor, affective, and imitative capacities captured by traditional infant tests do not look like "intelligence" or relate conceptually or empirically to later assessed intelligence. However, other domains of infant functioning meet both validity criteria. Students of the infant mind have developed a variety of tasks to discover what infants know and in some cases have tested the predictive validity of such measures for later mental functioning. Many tasks have been created to estimate the infant's ability to take in and to retain information. Such tasks provide a means for investigating classic theoretical issues as to whether intelligence is stable and continuous or takes different forms with development, how to approach the question of whether there is one or many forms of intelligence early in life, and the origins and interplay of genetic and neurobiological with experiential and environmental determinants of intelligence. Among diverse contemporary tasks are some that measure specific cognitive abilities, such as learning to act in a certain way to obtain a reward (Rovee-Collier, 1997), speed of eye movements that anticipate the location of a display following observation of a sequence of events (Dougherty & Haith, 1997), understanding the intentions of others (Woodward, 2009), performance of basic mathematical operations (McCrink & Wynn, 2007), object perception (Amso & Johnson, 2006; Mash, Arterberry, & Bornstein, 2007; Needham, 2009), and understanding of solids and liquids (Hespos, Ferry, & Rips, 2009); others measure more general cognitive capabilities, such as statistical learning (Kirkham, Slemmer, & Johnson, 2002; Kirkham et al., 2007), recognition memory (Bornstein et al., 2004; Keen & Berthier, 2004) and recall (Bauer, 2007), imitation skills (Demiris & Meltzoff, 2008; Legerstee & Markova, 2008; Meltzoff & Moore, 2002), categorizing (Arterberry & Bornstein, 2002; Bornstein & Arterberry, 2003; Oakes et al., 2009), and surprise when an anticipated event does not occur or when an unanticipated one does (Baillargeon, 2004; Hespos & Baillargeon, 2008).

Consider two such general capabilities in more detail. Information processing draws on attention and memory, traditionally viewed as key ingredients of intelligence (Deary, 1995; Stankov, 1983; Vernon, 1987) as well as achievement test scores and grades in school (Alexander, Entwisle, & Dauber, 1993). For example, using six longitudinal data sets, Duncan and colleagues (2007) determined links between key elements of school readiness – attention was one – and later school reading and math achievement. Selection, encoding, and retention processes, presumably measured by attention and memory, assess basic information processing or mental representation that underlie intelligence and do so in ways that are psychometrically sound, transcend sensory ability, are relatively free of motor requirements and affective components, do not rely on simple imitation, and conceptually parallel or serve as foundations for cognitive functions in childhood.

Sit a baby in an infant seat, show the baby a stimulus, and observe and record the baby's looking. When the stimulus first appears, the baby will normally orient and attend to it. If, however, the same stimulus is made available continuously or is presented repeatedly, the baby's visual attention to it usually wanes. The stimulus is no longer new and novel as it first was. Barring artifacts, the decrement in infant attention is thought to comprise processes that reflect the infant's (passive or active) development of some mental representation of stimulus information as well as the infant's ongoing comparison of new information with that representation. Infancy investigators have used such measures of attention to capture two meaningful indices of information processing in human infants: attention to some novel aspect of the environment (*novelty sensitivity or responsiveness*) and the rapidity with which attention declines to, or is withdrawn from, an unchanging aspect of the environment (*habituation*). Relatively more looking at novel stimuli, and less looking at familiar stimuli, are generally believed to index efficient information processing on the grounds that infants who have processed the information, and are now familiar with it, ought to recognize that familiar information later and attend to it less, but discriminate new information from the now familiar, attending to the novel more. Novelty sensitivity or responsiveness is indexed by spontaneous attention or recovery of attention in infants' looking at new information compared with familiar information. Habituation is indexed by the amount or rate of decay in infants' looking; greater decrements, quicker decays, and lower total looking times generally have been considered as indices of more efficient information processing in infants. These twin processes of attention in infants are believed to encompass the classic components of intelligence: selection, engagement, retention, encoding, and memory of information in the environment (Kandel, 2007; R. F. Thompson & Spencer, 1966; see also Rankin et al., 2009).

Consider habituation in more depth. Infant habituation satisfies two prerequisite psychometric criteria reasonably well. Habituation is characterized by adequate individual variation, and it has been shown to be a (moderately) reliable infant response at least over the short term (Bornstein & Benasich, 1986). Beyond these psychometric preliminaries, to arrive at an information-processing interpretation of habituation, investigators must support certain intuitions about the infant mind, and therefore require considerably more evidence. It is now possible to wire a "nomological net" among age, population, stimulus, mental representation, and validity arguments to support such an information-processing – perhaps "intelligence" – interpretative inference about habituation. Referring to the several characteristics just listed, older infants ought to habituate more efficiently than younger infants. They do. Normally developing infants ought to habituate more efficiently than infants born at risk for known cognitive developmental delays. They do. Infants ought to habituate to "simpler" stimuli more efficiently than to more "complex" stimuli. They do. Infants habituated to one stimulus should later be able to distinguish a novel stimulus in comparison with their mental representation of a now familiar stimulus. They can.

Evaluations of habituation as "cognitive" are further buttressed by assessments of its concurrent and predictive validity with respect to other measures of cognition. Infants who habituate efficiently scan and pick up information proficiently, detect information easily, encode and store information quickly, and/or retrieve and recognize information from memory more faithfully. Infant habituation relates to visual discrimination and recognition (Frick & Richards, 2001; Richards, 1997; Shaddy & Colombo, 2004); for example, more efficient habituators recognize perceptually degraded forms and symmetrical stimuli more quickly, where those less efficient require more time (Frick & Colombo, 1996; Stoecker et al., 1998). Children who habituate efficiently also explore their environment rapidly, play in relatively more sophisticated ways, solve problems quickly, attain concepts readily, and excel at operant learning, oddity identification, picture matching, and block configuration (Bjorklund & Schneider, 1996; Deary, 1995; Detterman, 1987; McCall, 1994; Nettelbeck, 1987; Vernon, 1987; for reviews, see Bornstein & Sigman, 1986; Colombo, 1993).

Finally, habituation in infancy predicts intelligence in childhood and after. One longitudinal study (recounted earlier) documented the predictive validity of habituation. At four months, infant habituation was assessed in the laboratory; at one year, infant productive vocabulary was ascertained; and at four years, child intelligence was evaluated using the Wechsler series. At four months and at one year, didactics in mothers' interactions with children were also recorded during home observations. Didactics included mothers' pointing, labeling, showing, demonstrating, and the like. Path analysis determined direct and unique longitudinal effects of independent variables on dependent variables. Maternal didactic efforts in encouraging infant attention contributed to both one-year and four-year child cognitive outcomes. However, infant habituation linked predictively both to toddler productive vocabulary size at one year and to childhood intelligence test performance at four years independent of maternal early and late didactic contributions. These findings (as stated in the section Quotidian Empiricism) were replicated and expanded in two shorter-term longitudinal follow-up studies (Tamis-LeMonda & Bornstein, 1989, 1993). Similarly, Laucht, Esser, and Schmidt (1994) found that three-month habituation accounted for a small but significant proportion of variance in 4.5-year IQ after accounting for the contributions of infants' three-month Bayley MDI, indexes of biological and psychological risk, and parent education. Furthermore, Bornstein and colleagues (2006) determined, in a large-*N* study, that parenting as well as maternal education were predictive of children's Denver and Griffiths developmental scores, MacArthur Communicative Developmental Inventories, and Wechsler Full Scale IQ at six, eighteen, twenty-four, and forty-eight months, respectively, but habituation efficiency at four months predicted Wechsler IQ at forty-eight months independent of those exogenous factors. Neonatal look duration even relates to selective attention at age twelve years (Sigman et al., 1991).

Infant information-processing abilities in the first six months of life in three domains (attention, speed, and memory) relate to language and executive functions (working memory, inhibition, and shifting) at age 1.8 years (Dixon & Smith, 2008), age four years (Courage, Howe, & Squires, 2004; Cuevas & Bell, 2014 and age eleven years (Rose, Feldman, & Jankowski, 2012), academic achievement at age fourteen years (Bornstein, Hahn, & Wolke, 2013), span of apprehension and intelligence at age eighteen years (Sigman, Cohen, & Beckwith, 1997), and IQ and academic achievement at age twenty-one years (Fagan, Holland, & Wheeler, 2007), even after partialing contributions of biological and psychological third variables (Laucht et al., 1994). Smith, Fagan, and Ulvund (2002) investigated the predictive validity of novelty responsiveness at seven and twelve months of age and parental socioeconomic status (SES) on later intellectual functioning at eight years of age in Norwegian families. Measures of a parent's socioeconomic status, such as education and occupation, give a rough estimate of the child's cultural environment, and parental socioeconomic status predicts later child IQ. Infants' novelty responsiveness also predicted later IQ, independent of parental SES.

A series of meta-analyses of predictive relations between habituation and intelligence has now been published, including Bornstein and Sigman (1986), Colombo (1993), McCall and Carriger (1993), and Kavšek (2004); and Kavšek and Bornstein (2010) for preterms. Kavšek (2004) included thirty-eight samples from twenty-five studies. The averaged weighted normalized predictive correlation coefficient (Hedges & Oklin, 1985) across studies of habituation in populations of typically developing infants is 0.48; for at-risk infants, it is 0.36; and for all infants combined, 0.45. Notably, no correlations in any studies have reported opposite results.

It is important to note at this juncture that habituation is not an epiphenomenon of laboratory investigation, but is manifest in infants' everyday interactions with people and objects in the world. A study of habituation in naturally occurring, home-based interactions of American and Japanese infants confirms this (Bornstein & Ludemann, 1989).

Habituation in infancy and intelligence in childhood and adolescence share common characteristics of information processing. Not only do infants assimilate environmental information in habituating, but when infants inhibit attending to familiar stimulation, they also liberate attentional and cognitive resources which can then be deployed in new encounters with new stimulation in the environment.

Comments and Conclusions

Comments

"Development" is readily associated with change (Block & Block, 2006; Kagan, 1976; Wohlwill, 1973), and species-general developmental functions of many characteristics are saltatory if not downright discontinuous (Emmerich, 1964). Moreover, some theories assert that infancy is disconnected from the balance of the life course and that infant characteristics and experiences are peripheral or ephemeral or inconsequential in the sense that they exert few if any enduring later-life effects (Bruer, 2002; Clarke & Clarke, 1972; Kagan, 2009; Lewis, 1997). Infancy may be a period of plasticity and adaptability to transient conditions, and many early effects may not persist or they may alter or be supplanted by subsequent conditions that are more consequential. However, a diverse and impressive array of alternative theories, including psychoanalysis, ethology, neuropsychology, behaviorism, constructivism, attachment, and systems theory, contend that infancy is part of a seamless and united lifeline and that characteristics, abilities, and experiences in infancy are meaningful in themselves and are sustained and crucial to later life. Although what is learned grows and changes with age, some basic abilities and processes to learn and remember are likely consistent in infants and adults (Rovee-Collier & Cuevas, 2009; Wagner & Lakusta, 2009). The assumption that information processing and memory underlie intelligence has made it possible to approach a century's worth of questions and controversies as to the developmental origins and nature of intelligence. Biological functioning, intellectual predilections, personality inclinations, and social orientations in infancy set patterns that endure across some or all of the life span (Bornstein, 2014).

The correlation coefficient (or statistical variant thereof) is often the main index used to authenticate a relation between infant scores and mature scores. Even small effect sizes in prediction from infancy can represent impressive support, however, showing that prediction holds under unlikely and unexpected circumstances, and can be as (or, in some cases, more) striking in simply showing that some infant index accounts for later variance. Small effect sizes are also meaningful from a public-health perspective (Abelson, 1985; Ahadi & Diener, 1989; Cortina & Landis, 2009; Prentice & Miller, 1992; Rosenthal & Rubin, 1982, 1983; Vacha-Haase & Thompson, 2004; Yeaton & Sechrest, 1981). However, it is necessary to interpret effect sizes taking into consideration both their assumptions and their limitations. For example, for any given relation there are likely to be subsets of children for whom the relation is smaller and subsets for whom the relation is larger. Furthermore, some relations between infancy and maturity may actually be obscured because they are studied at very distant points in time, and others may go unnoticed because surface manifestations even of the same characteristics at different developmental periods appear unrelated. We now acknowledge, too, that in contemporary relational systems perspectives, earlier emerging characteristics in development lay foundations for, and so likely exert impact on, later-appearing characteristics (Lewontin, 2005).

Findings of stability from so early in life often entice infancy researchers to believe that endogenous processes are at work (Bornstein, Putnick, & Esposito, 2017). It would be premature to characterize stability of any infant measure as reflecting shared cognitions, however, without considering potential "third-variable" roles of other endogenous processes and of exogenous ones rooted in experience and environment. As the nature of the longitudinal association is at base correlational, so other endogenous or exogenous variables could theoretically carry or mediate predictive associations. For example, habituation and temperament could share variance in infancy, just as intelligence test performance and temperament share variance in childhood (e.g., Guerin et al., 1994; Sigman et al., 1987; Zigler, Abelson, & Seitz, 1973). Besides intelligence, it takes a vigilant or persistent temperament to test well in infancy and in childhood. So, consistency in temperament could carry the predictive relation. However, habituation predicts intelligence over and above temperament. In their longitudinal study, Bornstein and colleagues (2006) found that habituation at four months predicted Wechsler Full Scale IQ at forty-eight months independent of both positive and negative temperament, and this relation extended to adolescent academic achievement at fourteen years independent of temperament and behavior problems. External experiences and family influences, both genetic and experiential, likely also play roles in child mental development, but they do not exclusively mediate prediction (see Bornstein, 1989a; Broman, Nichols, & Kennedy, 1975; Gottfried, 1984; Plomin & DeFries, 1985; Scarr, Weinberg, & Waldman, 1993).

Conclusions

Intelligence and prediction of intelligence from infancy are compelling and long-standing topics of philosophical, biological, psychological, sociological, and clinical interest. Among the perennial and far-reaching questions about human ontogeny, the issues of how infants learn and develop and what connections (if any) obtain between early individual differences and later life loom large (Bornstein et al., 2017). Findings as to the predictive validity of measures of early cognitive functioning for later intelligence and academic achievement support the view that intelligence is in some degree stable. How well infants process the information they are given to think about relates to how well they score as children on intelligence tests as well as what levels of education they eventually attain.

How much we have learned about infants – and infant intelligence – in approximately the last half-century is testimony to the perspicacity, patience, and persistence of researchers in meeting and overcoming the many formidable challenges posed by infants themselves. Through their intense appeal and participation in decades of rigorous scientific investigation, infants have slowly divulged their many secreted competencies. This glimpse into the heretofore private world of infancy, and what it portends for the future development of the child, constitutes a notable achievement of modern developmental science.

Intelligence in infancy has several implications. First, it has developmental meaning in more fully describing individuals and their growth. Second, it has theoretical significance for more completely understanding the nature of intelligence. Studies of intelligence and prediction between infancy and maturity are therefore important to theory building. Third, insofar as new measures of infant intelligence are nonverbal and behavioral, they may serve as "culture-neutral" assessments of early cognition. Fourth, because they show threads of prediction in mental performance from infancy, measures of infant cognitive ability might one day serve as clinical screening tools for early detection of risk status for delayed development and so they have implications for the identification, prevention, and possible treatment of disorders. Likewise, these measures could conceivably, fifth, constitute criteria for assessing deprivation or enrichment. Most generally, last, life-course study has been late arriving to the scholarly developmental literature (Elder, Shanahan, & Jennings, 2015), and these studies illustrate the benefits of longitudinal data (Bornstein, Hahn, & Putnick, 2016; Bornstein & Putnick, 2012; Ferri, Bynner, & Wadsworth, 2003; Hauser, 2009; Phelps, Furstenberg, & Colby, 2002), which Butz and Torrey (2006) extolled as one of the greatest innovations of twentieth-century social science. So, infancy studies have led to insights that have broad significance for psychology as a discipline.

The foregoing account of "infant intelligence" is telling, but it is critical to underscore the facts that it certainly does not mean that intelligence is innate or that the level of a child's intelligence is fixed in early life or that the infant is alone in the journey of becoming intelligent. This account does, however, substantiate a revised view of mental life from infancy. Accumulating theory, data, and experience overturn the argument that intelligence in infancy is not meaningful in itself or that it is unrelated to later life. Selected cognitive capabilities assessed even in the first year relate to selected measures of more mature intelligence assessed at least through adolescence. This perspective calls for a return to Alfred Binet's original ideas in instigating the mental-testing movement to ask how this new view of the infant's ability to know the world can best be used to benefit children. Infancy is a starting point. What we have learned about infant intelligence suggests that infancy may also represent a critical setting point in the life of the person.

Can little babies be said to be "intelligent"?

After surveying the literature, weighing the evidence, and writing this account, my answer is, Yes.

References

Abelson, R. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, *97*, 129–133.

Ahadi, S., & Diener, E. (1989). Multiple determinants and effect sizes. *Journal of Personality and Social Psychology*, *56*, 398–406.

Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (1993). First-grade behavior: Its short-and long-term consequences for school performance. *Child Development*, *64*, 801–814.

Amso, D., & Johnson, S. P. (2006). Learning by selection: Visual search and object perception in young infants. *Developmental Psychology*, *42*, 1236–1245.

Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Upper Saddle River, NJ: Prentice-Hall.

Arterberry, M. E., & Bornstein, M. H. (2002). Variability and its sources in infant categorization. *Infant Behavior and Development*, *25*, 515–528.

Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science*, *13*, 89–94.

Bauer, P. J. (2007). Recall in infancy: A neurodevelopmental account. *Current Directions in Psychological Science*, *16*, 142–146.

Bayley, N. (1949). Consistency and variability in the growth of intelligence from birth to eighteen years. *The Pedagogical Seminary and Journal of Genetic Psychology*, *75*(2), 165–196.

Bayley, N. (1955). On the growth of intelligence. *American Psychologist*, *10*(12), 805–818.

Bayley, N. (2005). *Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III®)*. Pearson. www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Behavior/Adaptive/Bayley-Scales-of-Infant-and-Toddler-Development-%7C-Third-Edition/p/100000123.html

Bellinger, D. (1980). Consistency in the pattern of change in mother's speech: Some discriminant analyses. *Journal of Child Language*, *7*, 469–487.

Binet, A. (1905). *On double consciousness: Experimental psychological studies*. Chicago: Open Court Publishing Company.

Bjorklund, D. F., & Myers, A. (2019). The evolution of parenting. In M. H. Bornstein (Ed.), *Handbook of parenting* (3rd ed., vol. 2, pp. 3–29). New York: Routledge.

Bjorklund, D. F., & Schneider, W. (1996). The interaction of knowledge, aptitude, and strategies in children's memory performance. *Advances in Child Development and Behavior*, *26*, 59–89.

Blaga, O. M., Anderson, C. J., Shaddy, D. J., Kannass, K. N., Little, T. D., & Colombo, J. (2009). Structure and continuity of intelligence during early childhood. *Intelligence*, *37*, 106–113.

Block, J., & Block, J. H. (2006). Venturing a 30-year longitudinal study. *American Psychologist*, *61*(4), 315–327.

Bornstein, M. H. (1985). How infant and mother jointly contribute to developing cognitive competence in the child. *Proceedings of the National Academy of Sciences*, *82*, 7470–7473.

Bornstein, M. H. (1989a). Sensitive periods in development: Structural characteristics and causal interpretations. *Psychological Bulletin*, *105*(2), 179–197.

Bornstein, M. H. (1989b). Between caretakers and their young: Two modes of interaction and their consequences for cognitive growth. In M. H. Bornstein & J. S. Bruner (Eds.), *Interaction in human development* (pp. 197–214). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bornstein, M. H. (2013). Mother-infant attunement: A multilevel approach via body, brain, and behavior. In M. Legerstee, D. W. Haley, & M. H. Bornstein (Eds.), *The infant mind: Origins of the social brain* (pp. 266–298). New York: Guilford.

Bornstein, M. H. (2014). Human infancy . . . and the rest of the lifespan. *Annual Review of Psychology*, *65*, 121–158.

Bornstein, M. H. (2015). Children's parents. In M. H. Bornstein & T. Leventhal (Vol. Eds.), *Ecological settings and processes in developmental systems*. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (7th ed., vol. 4, pp. 55–132). Hoboken, NJ: Wiley.

Bornstein, M. H. (Ed.) (2018). *SAGE encyclopedia of lifespan human development*. Thousand Oaks, CA: Sage.

Bornstein, M. H., & Arterberry, M. E. (2003). Recognition, categorization, and apperception of the facial expression of smiling by 5-month-old infants. *Developmental Science*, *6*, 585–599.

Bornstein, M. H., Arterberry, M. E., & Lamb, M. (2014). *Development in infancy: A contemporary introduction* (5th ed.). New York,: Psychology Press.

Bornstein, M. H., Arterberry, M. E., & Mash, C. (2004). Long-term memory for an emotional interpersonal interaction occurring at 5 months of age. *Infancy*, *6*(3), 407–416.

Bornstein, M. H., Arterberry, M. E., & Mash, C. (2013). Differentiated brain activity in response to faces of "own" versus "unfamiliar" babies in primipara mothers: An electrophysiological study. *Developmental Neuropsychology*, *38*(6), 365–385.

Bornstein, M. H., & Benasich, A. A. (1986). Infant habituation: Assessments of short-term reliability and individual differences at five months. *Child Development*, *57*, 87–99.

Bornstein, M. H., & Colombo, J. (2012). Infant cognitive functioning and mental development. In S. Pauen (Ed.), *Early childhood development and later achievement* (pp. 118–147). New York: Cambridge University Press.

Bornstein, M. H., Hahn, C.-S., Bell, C., Haynes, O. M., Slater, A., Golding, J., Wolke, D., & ALSPAC Study Team. (2006). Stability in cognition from early infancy: A developmental cascade. *Psychological Science*, *17*, 151–158.

Bornstein, M. H., Hahn, C.-H., & Putnick, D. L. (2016). Stability of core language skill across the first decade of life in children at biological and social risk. *Journal of Child Psychology and Psychiatry*, *57*, 1434–1443.

Bornstein, M. H., Hahn, C.-S., & Wolke, D. (2013). Systems and cascades in cognitive development and academic achievement. *Child Development*, *84*, 154–162.

Bornstein, M. H., & Ludemann, P. L. (1989). Habituation at home. *Infant Behavior and Development*, *12*, 525–529.

Bornstein, M. H., & Putnick, D. L. (2012). Stability of language in childhood: A multiage, multidomain, multimeasure, and multisource study. *Developmental Psychology*, 48, 477–491.

Bornstein, M. H., & Putnick, D. L. (2019). *The architecture of the child mind: g, F, and their hierarchy*. New York: Routledge.

Bornstein, M. H., Putnick, D. L., & Esposito, G. (2017). Continuity and stability in development. *Child Development Perspectives*, *11*, 113–119.

Bornstein, M. H., & Sigman, M. D. (1986). Continuity in mental development from infancy. *Child Development*, *57*, 251–274.

Bornstein, M. H., Tal, J., Rahn, C., Galperín, C. Z., Pêcheux, M.-G., Lamour, M., et al. (1992). Functional analysis of the contents of maternal speech to infants of 5 and 13 months in four cultures: Argentina, France, Japan, and the United States. *Developmental Psychology*, *28*, 593–603.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1990). Activities and interactions of mothers and their firstborn infants in the first six months of life: Covariation, stability, continuity, correspondence, and prediction. *Child Development*, *61*, 1206–1217.

Bornstein, M. H., Tamis-LeMonda, C. S., & Haynes, O. M. (1999). First words in the second year: Continuity, stability, and models of concurrent and predictive correspondence in vocabulary and verbal responsiveness across age and context. *Infant Behavior and Development*, *22*, 65–85.

Brody, N. (1992). *Intelligence* (2nd ed.). New York: Academic.

Broman, S. H. (1989). Infant physical status and later cognitive development. In M. H. Bornstein & N. A. Krasnegor (Eds.), *Stability and continuity in mental development: Behavioral and biological perspectives* (pp. 45–62). Hillsdale, NJ: Lawrence Erlbaum Associates.

Broman, S. H., Nichols, P. L., & Kennedy, W. A. (1975). *Preschool IQ: Prenatal and early developmental correlates*. Hillsdale, NJ: Erlbaum.

Bruer, J. (2002). *The myth of the first three years*. New York: The Free Press.

Butz, W. P., & Torrey, B. B. (2006). Some frontiers in social science. *Science*, *312*(5782), 1898–1900.

Chapman, R. S. (1981). Cognitive development and language comprehension in 10 to 21-month-olds. In R. E. Stark (Ed.), *Language behavior in infancy and early childhood* (pp. 359–394). New York: Elsevier North Holland.

Chen, Z., & Siegler, R. (2000). Intellectual development in childhood. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 92–116). New York: Cambridge University Press.

Choudhury, N., & Benasich, A. A. (2011). Maturation of auditory evoked potentials from 6 to 48 months: Prediction to 3 and 4 year language and cognitive abilities. *Clinical Neurophysiology*, *122*(2), 320–338.

Clarke, A. M., & Clarke, A. D. B. (Eds.) (1976). *Early experience: Myth and evidence*. New York: Free Press.

Coates, D. L., & Lewis, M. (1984). Early mother-infant interaction and infant cognitive status as predictors of school performance and cognitive behavior in six-year-olds. *Child Development*, *55*, 1219–1230.

Colombo, J. (1993). *Infant cognition: Predicting later intellectual functioning*. Newbury Park, CA: Sage.

Cooper, L. A., & Regan, D. T. (1986). Attention, perception, and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 123–169). Cambridge, UK: Cambridge University Press.

Cortina, J. M., & Landis, R. S. (2009). When small effect sizes tell a big story, and when large effect sizes don't. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and* *methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 287–308). New York: Taylor & Francis.

Courage, M. L., Howe, M. L., & Squires, S. E. (2004). Individual differences in 3.5-month-olds' visual attention: What do they predict at 1 year?. *Infant Behavior and Development*, *27*(1), 19–30.

Cuevas, K., & Bell, M. A. (2014). Infant attention and early childhood executive function. *Child Development*, *85*(2), 397–404.

Cummings, M. E., & Warmuth, K. M. (2019). Parenting and attachment. In M. H. Bornstein (Ed.), *Handbook of parenting* (3rd ed., vol. 4, pp. 374–400). New York: Routledge.

Darwin, C. R. (1859). *The origin of species*. New York: Modern Library.

Darwin, C. R. (1877). A biographical sketch of an infant. *Mind*, *2*, 286–294.

Davis, K., Christodoulou, J. A., Seider, S., & Gardner, H. (2011). The theory of multiple intelligences. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 485–503). New York: Cambridge University Press.

de Boysson-Bardies, B., & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language*, *67*(2), 297–319.

Deary, I. J. (1995). Auditory inspection time and intelligence: What is the causal direction? *Developmental Psychology*, *31*, 237–250.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, *9*, 133–150.

Demiris, Y., & Meltzoff, A. (2008). The robot in the crib: A developmental analysis of imitation skills in infants and robots. *Infant and Child Development*, *17*, 43–53.

Delauney-El Allam, M., Marlier, L., & Schaal, B. (2006). Learning at the breast: Preference formation for an artificial scent and its attraction against the odor of maternal milk. *Infant Behavior and Development*, *29*, 308–321.

Delauney-El Allam, M., Soussignan, R., Patris, B., Marlier, L., & Schaal, B. (2010). Long-lasting memory for an odor acquired at the mother's breast. *Developmental Science*, *13*(6), 849–863.

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence*, *15*, 157–173.

Detterman, D. K. (1987). Theoretical notions of intelligence and mental retardation. *American Journal of Mental Deficiency*, *92*, 2–11.

Dickstein, S., & Parke, R. D. (1988). Social referencing in infancy: A glance at fathers and marriage. *Child Development*, *59*, 506–511.

Dixon, W. E., & Smith, P. H. (2008). Attentional focus moderates habituation–language relationships: Slow habituation may be a good thing. *Infant and Child Development*, *17*(2), 95–108.

Dollard, J., & Miller, N. (1950). *Personality and psychotherapy*. New York: McGraw-Hill.

Domsch, H., Lohaus, A., & Thomas, H. (2009). Prediction of childhood cognitive abilities from a set of early indicators of information processing capabilities. *Infant Behavior and Development*, *32*(1), 91–102.

Dougherty, T. M., & Haith, M. (1997). Infant expectations and reaction time as predictors of childhood speed of processing and IQ. *Developmental Psychology*, *33*, 146–155.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446.

Elder, G. H., Shanahan, M. J., & Jennings, J. A. (2015). Human development in time and place. In M. H. Bornstein & T. Leventhal (Eds.), *Handbook of child psychology and* *developmental science*, vol. 4, *Ecological settings and processes in developmental systems* (7th ed., pp. 6–54). Hoboken, NJ: Wiley.

Emmerich, W. (1964). Continuity and stability in early social development. *Child Development*, *35*(2), 311–332.

Erikson, E. H. (1950). *Childhood and society*. New York: W. W. Norton and Company

Fagan, J. F., Holland, C. R., & Wheeler, K. (2007). The prediction, from infancy, of adult IQ and achievement. *Intelligence*, *35*(3), 225–231.

Fagan, J. F., & Singer, L. T. (1983). Infant recognition memory as a measure of intelligence. In L. P. Lipsitt (Ed.), *Advances in infancy research* (vol. 2, pp. 31–79). Norwood, NJ: Ablex.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116.

Ferri, E., Bynner, J., & Wadsworth, M. (2003). *Changing Britain, changing lives*. London: Institute of Education, University of London.

Freud, S. (1949). *An outline of psycho-analysis*. New York: Norton

Frick, J. E., & Colombo, J. (1996). Individual differences in infant visual attention: Recognition of degraded visual forms by four-month-olds. *Child Development*, *67*, 188–204.

Frick, J. E., & Richards, J. E. (2001). Individual differences in infants' recognition of briefly presented visual stimuli. *Infancy*, *2*, 331–352.

Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., et al. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics*, *39*(4), 546–557.

Geangu, E., Benga, O., Stahl, D., & Striano, T. (2010). Contagious crying beyond the first days of life. *Infant Behavior and Development*, *33*, 279–288.

Geangu, E., Benga, O., Stahl, D., & Striano, T. (2011). Individual differences in infants' emotional resonance to a peer in distress: Self-other awareness and emotion regulation. *Social Development*, *20*, 450–470.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*(1), 13–23.

Gottfried, A. W. (Ed.) (1984). *Home environment and early cognitive development*. Orlando, FL: Academic.

Goubet, N., Rattaz, C., Pierrat, V., Allemann, E., Bullinger, A., & Lequien, P. (2002). Olfactory familiarization and discrimination in preterm and full-term newborns. *Infancy*, *3*, 53–76.

Goubet, N., Strasbaugh, K., & Chesney, J. (2007). Familiarity breeds content? Soothing effect of a familiar odor on full-term newborns. *Journal of Developmental and Behavioral Pediatrics*, *28*, 189–194.

Gould, S. J. (1985). *Ontogeny and phylogeny*. Cambridge, MA: Belknap.

Guerin, D. W., Gottfried, A. W., Oliver, P. H., & Thomas, C. W. (1994). Temperament and school functioning during early adolescence. *Journal of Early Adolescence*, *14*, 200–225.

Hall, G. S. (1891). Notes on the study of infants. *Pedagogical Seminary*, *1*, 127–138.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. New York: Paul H. Brookes.

Hart, B., & Risley, T. R. (1999). *The social world of children learning to talk*. New York: Paul H. Brookes.

Hauser, R. M. (2009). The Wisconsin longitudinal study. In G. H. Elder & J. Z. Giele (Eds.), *The craft of life course research* (pp. 29–50). New York: Guilford Press.
Hedges, L. L., & Oklin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
Hespos, S. J., & Baillargeon, R. (2008). Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings. *Cognition*, *107*, 304–316.
Hespos, S. J., Ferry, A. L., & Rips, L. J. (2009). Five-month-old infants have different expectations for solids and liquids. *Psychological Science*, *20*, 603–611.
Hetherington, E. M., Parke, R. D., Gauvain, M., & Locke, V. O. (2006). *Childhood psychology: A contemporary viewpoint* (6th ed.). Boston: McGraw-Hill.
Hirshberg, L. M., & Svejda, M. (1990). When infants look to their parents: I. Infants' social referencing of mothers compared to fathers. *Child Development*, *61*, 1175–1186.
Hornik, R., Risenhoover, N., & Gunnar, M. (1987). The effects of maternal positive, neutral, and negative affective communications on infant responses to new toys. *Child Development*, *58*, 937–944.
Humphreys, L. G., & Davey, T. C. (1988). Continuity in intellectual growth from 12 months to 9 years. *Intelligence*, *12*, 183−197.
Hunt, E. B. (1983). On the nature of intelligence. *Science*, *219*, 141–146.
Hutman, T., & Dapretto, M. (2009). The emergence of empathy during infancy. *Cognition, Brain, Behavior*, *13*, 367–390.
Jaeger, S. (1985). The origin of the diary method in developmental psychology. In G. Eckhardt, W. G. Bringmann, & L. Sprung (Eds.), *Contributions to a history of developmental psychology* (pp. 63–74). Berlin: Mouton.
James, W. (1890). *Principles of psychology*. New York: Holt.
Jensen, A. R. (1980). Chronometric analysis of intelligence. *Journal of Social and Biological Structures*, *3*(2), 103–122.
Johnson, W., Emde, R. N., Pennbrook, B., Stenberg, C., & Davis, M. (1982). Maternal perception of infant emotion from birth through 18 months. *Infant Behavior and Development*, *5*, 313–322.
Kagan, J. (1976). Emergent themes in human development. *American Scientist*, *64*(2), 186–196.
Kagan, J. (2009). *Three seductive ideas*. Cambridge, MA: Harvard University Press.
Kandel, E. C. (2007). *In search of memory: The emergence of a new science of mind*. New York: W. W. Norton.
Kavšek, M. (2004). Predicting later IQ from infant visual habituation and dishabituation: A meta-analysis. *Journal of Applied Developmental Psychology*, *25*, 369–393.
Kavšek, M., & Bornstein, M. H. (2010). Visual habituation and dishabituation in preterm infants: A review and meta-analysis. *Research in Developmental Disabilities*, *31*(5), 951–975.
Keen, R. E., & Berthier, N. E. (2004). Continuities and discontinuities in infants' representation of objects and events. *Advances in Child Development and Behavior*, *32*, 243–279.
Kim, G., & Kwak, K. (2011). Uncertainty matters: Impact of stimulus ambiguity on infant social referencing. *Infant and Child Development*, *20*, 449–462.
Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*, B35–B42.

Kirkham, N. Z., Slemmer, J. A., Richardson, D. C., & Johnson, S. P. (2007). Location, location, location: Development of spatiotemporal sequence learning in infancy. *Child Development*, *78*, 1559–1571.

Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Huang, H., Ye, H.-H., et al. (2003). Effects of experience on fetal voice recognition. *Psychological Science*, *14*, 220–224.

Klein, P. S. (1988). Stability and change in interaction of Israeli mothers and infants. *Infant Behavior and Development*, *11*, 55–70.

Kopp, C. B., & McCall, R. B. (1980). Stability and instability in mental test performance among normal, at-risk, and handicapped infants and children. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (vol. 4, pp. 33–61). New York: Academic.

Kuhl, P. K. (2009). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In P. McCardle, J. Colombo, & L. Freund (Eds.), *Infant pathways to language: Methods, models, and research directions*. New York: Erlbaum.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the "critical period." *Language Learning and Development*, *1*(3–4), 237–264.

Kuhl, P. K., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annual Review of Neuroscience*, *31*, 511–534.

Lamott, A. (2013). *Some assembly required: A journal of my son's first son*. New York: Riverhead Books.

Laucht, M., Esser, G., & Schmidt, M. H. (1994). Contrasting infant predictors of later cognitive functioning. *Journal of Child Psychology and Psychiatry*, *35*(4), 649–662.

Legerstee, M., & Markova, G. (2008). Variations in 10-month-old infant imitation of people and things. *Infant Behavior and Development*, *31*, 81–91.

Legg, S., & Hutter, M. (2007). A collection of definitions of intelligence. *Frontiers in Artificial Intelligence and Applications*, *157*, 17–24.

Lerner, R. M., Hershberg, R. M., Hilliard, L. J., & Johnson, S. K. (2015). Concepts and theories of human development. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental science: An advanced textbook* (7th ed., pp. 3–41). New York: Psychology Press.

Lewis, M. (1997). *Altering fate: Why the past does not predict the future*. New York: Guilford.

Lewontin, R. (2005). *The triple helix*. Cambridge, MA: Harvard University Press

Lorenz, K. (1935/1970). *Studies in animal and human behavior* (Trans. R. Martin). London: Methuen.

Lunden, M., & Silven, M. (2011). Balanced communication in mid-infancy promotes early vocabulary development: Effects of play with mother and father in mono-and bilingual families. *International Journal of Bilingualism*, *15*(4), 535–559.

Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family: Parent–child interaction. In E. M. Hetherington (Ed.), *Handbook of child psychology*, vol. 4, *Socialization, personality, and social development* (3rd ed., pp. 1–101). New York: Wiley.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *7*(1), 41–74.

Manian, N., & Bornstein, M. H. (2009). Dynamics of emotion in infants of clinically depressed and nondepressed mothers. *Journal of Child Psychology and Psychiatry*, *50*, 1410–1418.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16.

Mash, C., Arterberry, M. E., & Bornstein, M. H. (2007). Mechanisms of visual object recognition in infancy: Five-month-olds generalize beyond the interpolation of familiar views. *Infancy*, *12*, 31–43.

McCall, R. B. (1994). What process mediates prediction of childhood IQ from infant habituation and recognition memory? Speculations on the roles of inhibition and rate of information processing. *Intelligence*, *18*, 107–125.

McCall, R. B., & Carriger, M. S. (1993). A meta-analysis of infant habituation and recognition memory performance as predictors of later IQ. *Child Development*, *64*, 57–79.

McCall, R. B., Hogarty, P. S., & Hurlburt, N. (1972). Transitions in infant sensorimotor development and the prediction of childhood IQ. *American Psychologist*, 27(8), 728–748.

McCrink, K., & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychological Science*, *18*, 740–745.

McLaughlin, B., White, D., McDevitt, T., & Raskin, R. (1983). Mothers' and fathers' speech to their young children: Similar or different? *Journal of Child Language*, *10*, 245–252.

Meltzoff, A. N., & Moore, M. K. (2002). Imitation, memory, and the representation of persons. *Infant Behavior and Development*, *25*, 39–61.

Mendelson, M. J. (1993). *Becoming a brother: A child learns about life, family, and self*. Cambridge, MA: MIT Press.

Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain and Language*, *72*(3), 238–245.

Murray, L., de Rosnay, M., Pearson, J., Bergeron, C., Schofield, E., Royal-Lawson, M., et al. (2008). Intergenerational transmission of social anxiety: The role of social referencing processes in infancy. *Child Development*, *79*, 1049–1064.

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social perception in infants* (pp. 177–197). Norwood, NJ: Ablex.

Needham, A. (2009). Learning in infants' object perception, object-directed action, and tool use. In A. Woodward & A. Needham (Eds.), *Learning and the infant mind* (pp. 208–226). New York: Oxford University Press.

Nettelbeck, T. (1987). Inspection time and intelligence. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 295–346). Norwood, NJ: Ablex.

Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology*, *42*(4), 643–655.

Nicely, P., Tamis-LeMonda, C. S., & Bornstein, M. H. (1999). Mothers' attuned responses to infant affect expressivity promote earlier achievement of language milestones. *Infant Behavior and Development*, *22*, 557–568.

Oakes, L. M., Horst, J. S., Kovack-Lesh, K. L., & Perone, S. (2009). How infants learn categories. In A. Woodward & A. Needham (Eds.), *Learning and the infant mind* (pp. 144–171). New York: Oxford University Press.

Overton, W. F. (2015). Process and relational developmental systems. In W. F. Overton & P. C. Molenaar (Eds.), *Handbook of child psychology and developmental science*, vol. 1, *Theory and method* (7th ed., pp. 9–62). Hoboken, NJ: Wiley.

Pancsofar, N., Vernon-Feagans, L., & Family Life Project Investigators. (2010). Fathers' early contributions to children's language development in families from low-income rural communities. *Early Childhood Research Quarterly*, *25*(4), 450–463.

Parke, R. D., & Cookston, J. (2019). Fathers and families. In M. H. Bornstein (Ed.), *Handbook of parenting* (3rd ed., vol.3, pp. 64–136). New York: Routledge.

Pearson, R. M., Lightman, S. L., & Evans, J. (2011). Attentional processing of infant emotion during late pregnancy and mother-infant relations after birth. *Archives of Women's Mental Health*, *14*(1), 23–31.

Phelps, E., Furstenberg, F. F., Jr., & Colby, A. (Eds.) (2002). *Looking at lives: American longitudinal studies of the twentieth century.* Troy, NY: Russell Sage Foundation.

Piaget, J. (1952). *The origins of intelligence in children* (Trans. M. Cook). New York: W. W. Norton.

Plato (1970). *The laws* (Trans. T. J. Saunders). Harmondsworth, Middlesex: Penguin. (Original work written ca. 355 BC)

Plomin, R., & DeFries, J. C. (1985). *Origins of individual differences in infancy: The Colorado Adoption Project*. New York: Academic Press.

Porter, R. H., & Levy, F. (1995). Olfactory mediation of mother–infant interactions in a selected mammalian species. In R. Wong (Ed.), *Biological perspectives on motivated activities* (pp. 77–110). Norwood, NJ: Ablex.

Porter, R. H., & Winberg, J. (1999). Unique salience of maternal breast odors for newborn infants. *Neuroscience and Biobehavioral Reviews*, *23*, 439–449.

Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin*, *112*(1), 160–164.

Preyer, W. (1882). *Die seele des kindes*. Leipzig: Grieben. Published in English in 1888–1889 as *The mind of the child*, Parts 1 and 2. (trans. H. W. Brown). New York: Appleton. (Reprinted by Arno Press, 1973.)

Prochner, L., & Doyon, P. (1997). Researchers and their subjects in the history of child study: William Blatz and the Dionne quintuplets. *Canadian Psychology*, *38*, 103–110.

Rankin, C., Abrams, T., Barry, R., Bhatnagar, S., Cerruti, D., Fang, C.-W., & Thompson, R. (2009). Habituation: An evaluation and revision of Thompson and Spencer (1966). *Neurobiology of Learning and Memory*, *92*, 135–138.

Reilly, S., Bavin, E. L., Bretherton, L., Conway, L., Eadie, P., Cini, E., et al. (2009). The Early Language in Victoria Study (ELVS): A prospective, longitudinal study of communication skills and expressive vocabulary development at 8, 12 and 24 months. *International Journal of Speech-Language Pathology*, *11*(5), 344–357.

Richards, J. E. (1997). Effects of attention on infants' preference for briefly exposed visual stimuli in the paired-comparison recognition-memory paradigm. *Developmental Psychology*, *33*, 22–31.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7-and 11-month-old American infants. *Developmental science*, *8*(2), 162–172.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2012). Implications of infant cognition for executive functions at age 11. *Psychological Science*, *23*(11), 1345–1355.

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, *74*(2), 166–169.

Rosenthal, R., & Rubin, D. B. (1983). A note on percent variance explained as a measure of the importance of effects. *Journal of Applied Social Psychology*, *9*(5), 395–396.

Rousseau, J. J. (1762). *Emile*. New York: Barron's Educational Series.

Rovee-Collier, C. (1997). Dissociations in infant memory: Rethinking the development of implicit and explicit memory. *Psychological Review*, *104*(3), 467–498.

Rovee-Collier, C., & Cuevas, K. (2009). Multiple memory systems are unnecessary to account for infant memory development: An ecological model. *Developmental Psychology*, *45*, 160–174.

Sattler, J. M. (1992). Assessment of children's intelligence. In C. E. Walker & M. C. Roberts (Eds.), *Handbook of clinical child psychology* (2nd ed., pp. 85–100). Oxford: John Wiley & Sons.

Scarr, S., Weinberg, R. A., & Waldman, I. D. (1993). IQ correlations in transracial adoptive families. *Intelligence*, *17*, 541–555.

Shaddy, D. J., & Colombo, J. (2004). Developmental changes in infant attention to dynamic and static stimuli. *Infancy*, *5*, 355–365.

Siegel, L. S. (1989). A reconceptualisation of prediction from infant test scores. In M. H. Bornstein & N. A. Krashnegor (Eds.), *Stability and continuity in mental development: Behavioral and biological perspectives* (pp. 89–103). Hillsdale, NJ: Erlbaum.

Sigman, M., Cohen, S. E., & Beckwith, L. (1997). Why does infant attention predict adolescent intelligence? *Infant Behavior and Development*, *20*(2), 133–140.

Sigman, M., Cohen, S. E., Beckwith, L., Asarnow, R., & Parmelee, A. H. (1991). Continuity in cognitive abilities from infancy to 12 years of age. *Cognitive Development*, *6*(1), 47–57.

Sigman, M., Cohen, S. E., Beckwith, L., & Topinka, C. (1987). Task persistence in 2-year-old preterm infants in relation to subsequent attentiveness and intelligence. *Infant Behavior and Development*, *10*, 295–305.

Singh, L., Liederman, J., Mierzejewski, R., & Barnes, J. (2011). Rapid reacquisition of native phoneme contrasts after disuse: You do not always lose what you do not use. *Developmental science*, *14*(5), 949–959.

Smith, L., Fagan, J. F., & Ulvund, S. E. (2002). The relation of recognition memory in infancy and parental socioeconomic status to later intellectual competence. *Intelligence*, *30*(3), 247–259.

Sorce, J. F., & Emde, R. N. (1981). Mother's presence is not enough: Effect of emotional availability on infant exploration. *Developmental Psychology*, *17*, 737–745.

Sroufe, L. A., Egeland, B., Carlson, E., & Collins, W. A. (2005). *The development of the person: The Minnesota study of risk and adaptation from birth to adulthood*. New York: Guilford Press.

Stankov, L. (1983). Attention and intelligence. *Journal of Educational Psychology*, *75*, 471–490.

Steiner, J. E. (1979). Human facial expressions in response to taste and smell stimulation. In H. Reese & L. Lipsitt (Eds.), *Advances in child development and behavior* (vol. 13, pp. 257–295). New York: Academic.

Stern, D. (1990). *Diary of a child*. New York: Basic Books.

Sternberg, R. J. (Ed.) (2020). *Cambridge handbook of intelligence* (2nd ed.). New York: Cambridge University Press.

Sternberg, R. J., & Detterman, D. K. (Eds.) (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Santa Barbara, CA: Praeger Pub Text.

Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, *47*, 1–41.

Stoecker, J. J., Colombo, J., Frick, J. E., & Ryther, J. S. (1998). Long- and short-looking infants' recognition of symmetrical and asymmetrical visual forms. *Journal of Experimental Child Psychology*, *71*, 63–78.

Taine, H. A. (1877). Taine on the acquisition of language by children. *Mind*, *2*, 252–259.

Tamis-LeMonda, C. S., & Bornstein, M. H. (1989). Habituation and maternal encouragement of attention in infancy as predictors of toddler language, play, and representational competence. *Child Development*, *60*, 738–751.

Tamis-LeMonda, C. S., & Bornstein, M. H. (1993). Antecedents of exploratory competence at one year. *Infant Behavior and Development*, *16*(4), 423–439.

Thompson, R. A. (2006). The development of the person: Social understanding, relationships, conscience, self. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (6th ed., vol. 2, pp. 24–98). Hoboken, NJ: Wiley.

Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, *73*, 16–43.

Tiedemann, D. (1787).Beobachtungen über die Entwicklung der Seelenfähigkeiten bei Kindern. *Hessische Beiträge zur Gelehrsamkeit und Kunst*, 2, 313–315 and 3, 486–488. (Observations on the development of the mental faculties of children, *Pedagogical Seminary*, 1927, *34*, 205–230.)

Tinbergen, N. (1951). *The study of instinct*. Oxford: Oxford University Press.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*(4), 1067–1084.

Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology*, *51*(4), 473–481.

Vernon, P. E. (1947). The variations of intelligence with occupation, age, and locality. *British Journal of Mathematical and Statistical Psychology*, *1*(1), 52–63.

Vernon, P. E. (Ed.) (1987). *Speed of information-processing and intelligence*. Norwood, NJ: Ablex.

Wachs, T. D., & Chan, A. (1986). Specificity of environmental action, as seen in environmental correlates of infants' communication performance. *Child Development*, *57*, 1464–1474.

Wagner, L., & Lakusta, L. (2009). Using language to navigate the infant mind. *Perspectives on Psychological Science*, *4*, 177–184.

Wallace, D. B., Franklin, M. B., & Keegan, R. T. (1994). The observing eye: A century of baby diaries. *Human Development*, *37*(1), 1–29.

Watson, J. B. (1924/1970). *Behaviorism* (Rev. ed.) New York: Norton.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore, MD: Williams & Wilkins Co.

Winnicott, D. W. (1965). *The maturational processes and the facilitating environment: Studies in the theory of emotional development*. New York: International Universities Press.

Wohlwill, J. F. (1973). *The study of behavioral development*. New York: Academic Press.

Wolfe, C. D., & Bell, M. A. (2007). The integration of cognition and emotion during infancy and early childhood: Regulatory processes associated with the development of working memory. *Brain and Cognition*, *65*(1), 3–13.

Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current Directions in Psychological Science*, *18*, 53–57.

Yamaguchi, M., Kuhlmeier, V. A., Wynn, K., & VanMarle, K. (2009). Continuity in social cognition from infancy to childhood. *Developmental Science*, *12*(5), 746–752.

Yeaton, W., & Sechrest, L. (1981). Meaningful measures of effect. *Journal of Consulting and Clinical Psychology*, *49*, 766–767.

Yu, H., McCoach, D. B., Gottfried, A. W., & Gottfried, A. E. (2018). Stability of intelligence from infancy through adolescence: An autoregressive latent variable model. *Intelligence*, 69, 8–15.

Zigler, E., Abelson, W. D., & Seitz, V. (1973). Motivational factors in the performance of economically disadvantaged children on the Peabody Picture Vocabulary Test. *Child Development*, *44*, 294–303.

8 Intelligence in Childhood

Susan A. Gelman and Jasmine M. DeJesus

Introduction

In this chapter, we argue that children provide a unique and valuable window onto understanding human intelligence. Indeed, intelligence cannot be fully understood without a consideration of how it develops in childhood. This claim might at first seem surprising. In many respects, children seem like the prototypical case of *un*intelligent thought. In the preschool years, children need help with the most basic tasks of survival, from donning clothes to preparing a meal to finding their way around their neighborhood. Even older children fare substantially worse than adults on the very tasks that are standardly used to measure intelligence, such as digit span (Cowan, 2016), speed of processing (Kail, Lervåg, & Hulme, 2016), logical reasoning (Moshman & Tarricone, 2016), and analogical reasoning (Vendetti et al., 2015). Indeed, if the average ten-year-old were scored on the same scale as sixteen-year-olds, their IQ would be 76 – in the lowest fifth percentile of the population (Kaufman, 1997). One might conclude, in brief, that children's intelligence is distinctive in what it lacks.

Why, then, take a developmental perspective? What does the study of children reveal about human intelligence? We note three contributions.

First, intelligence goes hand in hand with learning, and children are (arguably) the best learners on the planet. In the five-year span from the ages of two to seven, dramatic changes are taking place. A seven-year-old has learned to speak in grammatically complex sentences, engage in arithmetic calculations, read and write, reason accurately about others' false beliefs, compute another person's visual perspective when different from their own, and so on. In non-Western cultures, seven-year-olds have learned to weave, to use a machete, or to herd sheep (Rogoff, 2003). These changes are qualitative leaps compared to the learning that typically takes place in any five-year period in adulthood (e.g., ages 22–27). Anyone who wishes to understand what gives human intelligence its flexibility and power would be wise to attend to those impressive learners.

Second, intelligence is adaptive change, and children are the paradigm case of an organism that changes and adapts. Studying children necessitates studying process and change – not just *what* but also *how*. Put differently: The study of intelligence does not just entail measuring and understanding capacities and talent in a fully functioning adult, but also the puzzle of how that adult endpoint comes to be. A careful consideration of intelligence in childhood permits insights into rather deep philosophical questions, including: How much continuity versus change is there in human intelligence? Where does intelligence come from? How much do we depend on our social or cultural context to reason in intelligent ways? These questions are difficult, perhaps impossible, to answer by studying adults. Therefore, a focus on children is instructive.

Third, historically, the study of children's intelligence has taken seriously the reasoning behind their errors. An incorrect answer is not just a failure to achieve the predicted response; instead it provides a window onto how the mind works. It is instructive to consider the experience of the great Swiss psychologist Jean Piaget, who is widely considered the father of cognitive development (Flavell, 1963). Piaget began his post-PhD career working in the lab of Théodore Simon, the French psychometrician who, with Alfred Binet, developed the Binet-Simon scale for measuring intelligence in children. Piaget's job was to administer IQ tests to young French children to help develop standardized age norms. During this period, he found himself drawn more to understanding children's errors than their correct responses (Furth, 1973). He found that children of a given age often made similar errors, and that these errors suggested ways in which they construed the world in distinctively their own way. The point is that intelligence is not only about "right" or "wrong," but also about how an intelligent human mind takes in information and attempts to make sense of it. When the intelligent mind is lacking in information (such as in a child), the conclusions may be lacking, but that does not mean that the intelligence is lacking.

Although this chapter is about children, the lessons learned from the study of children have broad significance for intelligence at any age. By studying how a child learns to solve math problems, constructs a theory of biology, or makes inferences about their friends' thoughts and feelings, we can learn about what makes the human species intelligent. It is a truism that all adults were once children. Studying children reveals the cognitive capacities that ultimately support the most impressive achievements of our species, from designing cities, to mapping the human genome, to sending people to outer space.

The chapter is organized into five sections, each addressing a key theme of childhood intelligence: Continuity amid developmental change, multiple modes of reasoning, when children outperform adults, the role of social context, and policy implications. Each section focuses on a few content areas that illustrate the theme and how it relates to intelligence. We draw on literature from the full childhood period from two years to eighteen years, though the primary focus is on children in preschool and elementary school (2–10 years), where the majority of research has been done.

Continuity Amid Developmental Change

Debates about the continuity versus discontinuity of thought run through the study of children's reasoning. Whereas classic views proposed that children undergo qualitative, stagelike shifts over development, and that children are irrational, illogical, and egocentric until adolescence (e.g., Piaget, 1968), more recent evidence suggests that even infants and young children approach learning in a manner that is fundamentally rational, evidence based, and sensitive to patterns of data in the environment (Xu & Kushnir, 2013). This point has been demonstrated with tasks in which children learn new words (Xu & Tenenbaum, 2007), make generalizations from samples (Rhodes, Gelman, & Brickman, 2010), determine the causes of their own failed actions (Gweon & Schulz, 2011), or interpret others' actions (Gergely & Csibra, 2003). For example, 5–6-year-old children interpret the choices that others make in terms of an intuitive utility calculus involving (invisible) costs and rewards, enabling them to make consistent inferences about a specific character's preferences and abilities (Jara-Ettinger et al., 2015). Thus, when observing someone who is given a choice between an easily accessible apple and a hard-to-reach orange, 5–6-year-old children are more confident that someone who chooses the orange and thus has to climb up high to reach it has a general preference for oranges, whereas someone who chooses the apple is less likely to have a general preference for apples (they just selected the most convenient option). Similarly, when four-year-olds are tested on their ability to resist temptation in order to obtain a more rewarding treat in the future, they are swayed by rational calculations regarding the likelihood that the experimenter is reliable and will actually provide the better treat (Kidd, Palmeri, & Aslin, 2013).

This continuity in thought shows not only that children are more capable than previously suspected, but also that intelligence can appear in surprisingly humble forms. Consider, for example, the action of grasping a spoon (Keen, 2011). For an adult, picking up a spoon would not appear on any IQ test. The behavior makes use of visual cues and motor actions, but would not ordinarily be considered to require higher-level thought. For a one-year-old child, though, picking up a spoon is a problem to be solved. The angle, orientation, and directionality of the spoon relative to a child's handedness all pose variations that require representing the trajectory of one's hand relative to ever-changing factors. And in response to these challenges, infants and toddlers generate an action plan that determines the type of grip and manner of transportation to the mouth. In a way, carrying out this humble activity is the baby equivalent of solving a spatial reasoning task, where one has to rotate an unfamiliar block of cubes to decide which two items are the same. (That's on the IQ test!)

Similarly, learning to locomote requires an analysis of the infant's own (constantly growing) body, the environment that they are in (e.g., traversing a smooth path, a slope, or an edge), and the task at hand (e.g., crawling, creeping, or walking) in order to flexibly select the appropriate actions (Adolph & Hoch, 2019). These everyday behaviors are intelligent, revealing an ability to take in information and adjust expectations in an adaptive, responsive manner. An intriguing finding is that motor maturity and object exploration at five months of age predict intelligence scores at four and ten years of age, and academic achievement at fourteen years of age (Bornstein, Hahn, & Suwalsky, 2013). This result suggests cascading effects in development, whereby initial advances (e.g., in reaching or exploration) provide infants with opportunities to gain additional new experiences, which themselves provide further kinds of stimulation and learning (see also Campos et al., 2000; Libertus, Joh, & Needham, 2016). For example, three-month-old infants who are given Velcro-backed mittens that enable them to reach for and manipulate objects earlier than they could do on their own engage in more manipulation of objects at fifteen months of age (Libertus et al., 2016) and are better able to understand other people's goals than infants who were not given the "sticky" mittens (Sommerville, Woodward, & Needham, 2005).

Children's number concepts provide a particularly rich example of continuity amid developmental change. Preverbal infants and many nonhuman animals possess two core frameworks for reasoning about number: an approximate number system (ANS) and an object file system (OFS) (Carey, 2009; Feigenson, Dehaene, & Spelke, 2004). The ANS represents large numbers (i.e., greater than 3) in an inexact way, whereas the OFS represents small numbers (usually up to 3) in a precise way. An example of a task measuring ANS is to rapidly flash a display of yellow and blue dots and ask the participant to report which color has more dots. The larger the ratio between a larger and smaller set, the more accurately a person can detect which set is larger in a glance (i.e., without counting). The ratio needed for success decreases with age: six-month-olds can distinguish sets differing in a 2:1 ratio; two-month-olds can distinguish sets differing in a 3:2 ratio, and adults can distinguish sets differing in an 8:7 ratio. An example of a task measuring OFS in infants is to place treats one at a time into two containers, and then allow the baby to choose one of the two containers. By ten months of age, babies select the container with more items, as long as there are no more than three items (i.e., 1 vs. 2, 1 vs. 3, or 2 vs. 3). When the set of treats in one of the containers exceeds three, thereby exceeding the limits of the OFS, performance falls to chance levels. This results in the counterintuitive result that babies can successfully retrieve the larger set when the contrast is one versus three, but not when the contrast is one versus four.

What these two core systems cannot represent are large, precise numbers (such as 8 or 275) – those that are the foundation of formal mathematical computations required for success in school. Reasoning about large numbers in a precise way seems to require cultural support in the form of a linguistic counting system (Frank et al., 2008). Even for children who have learned a symbolic counting system, it takes years to understand the linear relation of large numbers to one another, as revealed by their errors when asked to show where numbers should go on a number line where only the end points are labeled, such as 0 to 1,000 (Siegler & Opfer, 2003). On such tasks, children before about sixth grade rely on logarithmic rather than linear representations of numerical magnitudes, for example, second-graders tend to place the number 75 at about one-third of the way between 0 and 1,000 on the number line – shifting it much farther to the right than it should be. Figuring out the number line does not happen all at once, but is solved for lower numbers (e.g., 0 to 100) before larger numbers (e.g., 0 to 1,000), suggesting that number magnitude knowledge is acquired gradually over a period of years (Siegler & Lortie-Forgues, 2014; but see Cohen & Sarnecka, 2014, for an alternative interpretation of these errors).

Despite what appears to be a developmental discontinuity between core systems of number and symbolic representations, evidence suggests that the core systems of number provide a foundation for later mathematical reasoning (Starr, Libertus, & Brannon, 2013). These observations follow from linking individual differences in ANS with individual differences in standardized math scores. For example, children's standardized math performance at age five years correlates with their ANS scores at age fourteen (Halberda, Mazzocco, & Feigenson, 2008), even when controlling for general intelligence, rapid lexical access, and a range of other standardized test scores. Moreover, children with dyscalculia (a learning disability involving specific impairment in reasoning about numbers and arithmetic) are severely impaired in their performance on tests of the ANS (Piazza et al., 2010). Precisely how ANS and formal mathematics performance are linked is an open question that is the topic of ongoing research (Feigenson, Libertus, & Halberda, 2013).

Ultimately, issues of continuity and discontinuity also need to consider timescales, as changes that appear discontinuous from one birthday to the next may (or may not) be more continuous from one day to the next. The microgenetic method can be an important tool for providing insights into processes and mechanisms of change (Rhodes & Wellman, 2013; Siegler, 1996). This approach involves gathering densely spaced measurements using the same task over a period of rapid change (e.g., giving children a multiplication test once a week over eight weeks, to examine patterns of change as children acquire new strategies; van der Ven et al., 2012). These insights can reveal both continuity (i.e., gradual improvement over the time period under investigation) and discontinuity (i.e., on study day 8 the child did not understand the concept, whereas on study day 9 they did) in children's reasoning.

Multiple Modes of Reasoning

Human intelligence is not just the accumulation of more facts or quicker solutions; rather, at its best, human intelligence permits and promotes new problem-solving strategies (e.g., how to reason about balancing items on a scale; Hofman et al., 2015), new concepts (e.g., differentiating heat and temperature; Carey, 2009), and new theories (e.g., natural selection; Rosengren et al., 2012). A question that has recently received much attention is the extent to which these new ways of thinking displace older frameworks or strategies. In other words, we may ask whether advances in skills, concepts, explanatory frameworks, or problem-solving approaches consistently "win out," or whether instead a child uses both older and newer conceptual strategies/concepts/systems, perhaps differentially primed with different contexts or tasks.

A consistent theme is that multiple modes of reasoning coexist at any point in development. For example, even after children acquire a more effective means of approaching a problem, they do not necessarily make use of this strategy consistently, or in optimal ways. This has been documented in the domain of memory. Children's memory can be enhanced by the use of deliberate, metacognitive strategies, such as rehearsal (repeating the information to be recalled), organization (putting items into meaningful groups or categories), or elaboration (adding details or creating a new image). Yet much research documents that children at first don't spontaneously use strategies during learning to boost recall, that the use and breadth of strategies increases over the elementary school years, that inefficient strategies continue alongside more efficient ones, and that even when children are trained to use strategies and can see that they are effective, they tend to stop using them once they are no longer required to do so, a phenomenon known as "production deficiency" (Bjorklund, Dukes, & Brown, 2008; Morrison & Chein, 2011).

A similar pattern can be seen in children's procedures for solving problems. Siegler (1996) notes that children make use of a repertoire of different strategies at any point in time, and that the use of particular strategies increases and decreases, in a pattern of overlapping waves. For example, in one study of Dutch elementary school children learning single-digit multiplication, children used a combination of retrieval, derived facts, repeated addition, and counting to solve these problems (van der Ven et al., 2012). Some strategies were considered more mature than others (e.g., solving 4×7 by stating $4 \times 5 = 20$ and then counting fingers from 20 to 28, was more mature than adding $7 + 7 + 7 + 7$). With increasing ability, children tended to use more mature strategies, but nonetheless their progress through these strategies was not a matter of wholesale replacement of a less efficient strategy with one that is more efficient. This is in contrast to classic arguments that children undergo steplike changes with age, with shifts from one strategy to the next, ever-increasing in efficiency and effectiveness.

Children's causal-explanatory systems (often referred to as folk or commonsense theories; Gopnik & Wellman, 1992; Wellman & Gelman, 1998) likewise reveal that different modes of reasoning may coexist within an individual (Gelman & Legare, 2011; Legare et al., 2012). A rich source of examples can be found in studies of children's "naïve biology" – that is, their inferences, predictions, and explanations regarding living things (Hatano & Inagaki, 1994). At the same time that young children understand that people, nonhuman animals, and plants participate in natural processes that are lawlike and beyond the control of human action (e.g., growth is inevitable and unidirectional; illness is caused by "germs"; death is permanent; Wellman & Gelman, 1998), they also endorse supernatural explanations: Illness can result from witchcraft (Legare & Gelman, 2008), magic (Nguyen & Rosengren, 2004), or bad behavior (such as lying or cheating; Raman & Gelman, 2004); the dead can continue to think, feel emotionally connected to others, and have spiritual existence (Bering, Blasi, & Bjorklund, 2005; Harris, 2011; Rosengren et al., 2014); God is not subject to biological constraints (Giménez-Dasí, Guerrero, & Harris, 2005); and prayer permits transcending human limitations (Woolley & Phelps, 2001). Although preschool-age children are not experts at reasoning about germs or avoiding illness (e.g., Blacker & LoBue, 2016; DeJesus, Shutts, & Kinzler, 2015; Solomon & Cassimatis, 1999), their naïve biological theories may provide an important foundation on which children can build more advanced theories.

Even when focused on natural explanations, children are prone to cognitive biases that distort their scientific reasoning. One topic that has received much attention is

Darwin's theory of natural selection, which is rejected at high rates (e.g., about 50% of U.S. adults; Rosengen et al., 2012). One obstacle to understanding evolution is adherence to psychological essentialism, an intuitive belief that natural categories are immutable, that there are sharp boundaries between such categories, and that all members of a natural category are fundamentally alike and share the same internal essence (Gelman, 2003). These assumptions run counter to several concepts required for appreciating Darwin's theory of natural selection, including species change over time, within-species variability, and population-level causal processes (Gelman & Rhodes, 2012). Adults who endorse essentialist assumptions about categories are more likely to hold erroneous beliefs about the process of evolutionary change (Shtulman & Calabi, 2012). Essentialism has also been linked to misconstruals of the biological processes of growth/metamorphosis and genetics, again in adults (Dar-Nimrod & Heine, 2011; Gelman & Marchak, in press).

A second obstacle to understanding biological processes – including not just evolutionary theory but also inheritance – is the tendency to assume a teleological perspective when reasoning about species (the idea that characteristics of an animal exist for a purpose; Kelemen, 2012). Thus, for example, four- and five-year-old children overextend a teleological understanding, reporting (for example) that animals are "for" walking around. Importantly, misconceptions may lead to inappropriate inferences, and not just in childhood. Adults, too, fall back on heuristics such as essentialism and "promiscuous" (overgeneralized) teleology when under time pressure (Eidson & Coley, 2014; Kelemen, Rottman, & Seston, 2013). For example, adults interpret evolution as a teleologically driven process: "Giraffes have long necks to reach food on treetops" (with the function of reaching treetops serving as the cause of giraffes' long necks) rather than "Giraffes evolved long necks because of the differential reproductive success of earlier animals whose progenitor necks happened to offer marked, heritable, tree-top-food-reaching benefits" (Kelemen, 2012; Sánchez Tapia et al., 2016). Consequently, adults report that animals can inherit properties that they need, or that have value in their environment (Ware & Gelman, 2014).

In summary, achieving certain intellectual insights does not mean that more intuitive systems disappear, or that children (or even adults) use them in all contexts. Intellectual oversights may be an overlay on top of intuition, much like learning about genealogically based kinds does not prevent a botanist from seeing "trees," despite its lack of biological reality (because genealogical units, such as legumes, crosscut our folk distinctions between tree, vine, and bush). In Scott Atran's (1998, p. 563) words, "science cannot simply subvert common sense." The coexistence of different perspectives (be they problem-solving strategies or conceptual frameworks) not only provides insight into the nature of human intelligence, but also has translational implications for how to elicit more intelligent solutions to problems that we face (e.g., evaluating scientific claims; sifting well-reasoned arguments from fallacious arguments; Horne et al., 2015; Shtulman & Harrington, 2016). Smart people do not always think in smart ways, and it is important to consider the conditions that encourage thoughtful decision-making (e.g., Kahneman, 2011).

When Children Outperform Adults

A truism in developmental psychology is that children get better with age, and certainly this is the case for many aspects of intelligence. As noted in the introduction, tasks that assess or correlate with intelligence (e.g., memory span, processing speed, reasoning skills, problem-solving strategies, vocabulary) show improvement as children get older. One capacity that has far-reaching implications for a wide swath of developmental skills is that of executive function (also known as cognitive control), which involves monitoring and regulating one's own behavior (Diamond, 2013). Components of executive control include working memory, inhibitory control, and the ability to engage in task-switching. Tasks measuring executive function reveal the maximum complexity in the rules that one can use to solve problems (Zelazo et al., 2003). These are effortful processes that undergo substantial developmental changes during the elementary and middle school years. While these skills are critical throughout life, the early years may be an especially important window for intervention (see Diamond, 2013). Children who arrive at school with higher executive function skills are better prepared to succeed in a traditional school environment – they will receive positive attention from their teachers, others will hold higher expectations for their performance and capabilities, and they will develop a positive feeling toward school and their intellectual capacity. The opposite is true for children who arrive in school with lower executive function capabilities, creating a cycle that is difficult to reverse, even though it is possible to intervene to improve executive function later in life. Entire volumes have been devoted to the topic of executive function, but the important takeaway here is that these skills undergo important changes in the early years, and are predictors of children's learning outcomes, even beyond their scores on IQ tests and reading/math skills at school entry.

Similarly, analogical reasoning – the capacity to detect relationships between superficially dissimilar items – improves markedly as children get older. Analogical reasoning has been implicated in "general fluid intelligence, creativity, and adaptive learning capacities" (Richland & Burchinal, 2013). Analogical reasoning improves alongside executive function and other developmental changes (Vendetti et al., 2015). For example, children who performed better on an executive function task (the Tower of Hanoi) in first grade performed better on a verbal analogies test at age fifteen, even after controlling for children's vocabulary scores and demographic variables (Richland & Burchinal, 2013). Additionally, analogical reasoning skills develop as children become better able to ignore irrelevant features of the problems at hand (Richland, Morrison, & Holyoak, 2006), as they gain more content knowledge (Gentner & Rattermann, 1991), and as the functional connectivity between brain regions active during analogical problem-solving improves (Wendelken et al., 2015).

Despite these compelling examples of improvements with age, "older is better" is not always the case. In the remainder of this section we note three respects in which children may outperform adults: learning languages, detecting unexpected patterns in data, and reasoning in expert domains (compared to adult novices).

Children are particularly skilled at learning languages, and those who learn a language as a child ultimately reach higher levels of grammatical competence than those who learn a language as an adult. This critical or sensitive period has been found in both first-language acquisition (Newport, 1990) and second-language acquisition (e.g., Hartshorne, Tenenbaum, & Pinker, 2018). How and why children outperform adults is a matter of debate. One possibility is that language-specific perception and learning mechanisms decay with decreasing neural flexibility as children's brains mature (Petitto et al., 2012; Werker & Tees, 2002). In contrast, Newport (1990) has suggested a counterintuitive "less-is-more" hypothesis, whereby the enhanced domain-general information-processing abilities of adults actually stand in the way of detecting patterns in language. On this view, children's more limited capacity breaks down the input into smaller units that allow for more effective componential analysis. Children's inability to perceive and remember larger, more complex "chunks" of language may (paradoxically) be an advantage, given that language learning requires breaking down the speech input into smaller component pieces that combine and recombine into ever-larger units in a hierarchically organized structure. In support of this theory, Kam and Newport (2009) used an artificial-language learning paradigm to test how children and adults process inconsistent language input, and found that children – but not adults – impose patterns that regularize the system (see also Finn et al., 2014; Senghas, Kita, & Özyürek, 2004). An interesting open question is whether there might be other aspects of thought in which increased information-processing abilities stand in the way of detecting or abstracting structures.

The distinctive life history of humans (with a much longer period of childhood compared to that of other species) is also thought to contribute to differences in how children versus adults learn (e.g., Bjorklund & Green, 1992). Gopnik and colleagues (2017) propose that there is a trade-off between exploration (searching broadly, flexibly, and in a rather unconstrained way through a hypothesis space) and exploitation (focusing in on familiar hypotheses that have been successful in the past). Children seem to be particularly good at learning new structures, whereas adults are particularly good at more efficient, rapid, and skilled performance. Specifically, children exhibit a surprisingly high ability to detect unexpected patterns in data. One task in which children show better performance than adults involves watching a machine that provides information, on a trial-by-trial basis, about the cause of a physical event (the machine lighting up and playing music). On each trial, one or more blocks are placed on the device, on some trials leading to the machine activating, and on other trials leading to the machine not responding. After seeing a series of trials, participants are then asked to determine which blocks will make the machine activate. The results show that younger children rely on the evidence and respond in accordance with the data that they see, whereas adolescents and adults rely more on their prior beliefs about likely causes. Thus, with age, participants are more likely to guess that causes are disjunctive (where only one block causes the machine to activate), even when the evidence favors a conjunctive cause (where two blocks together are required to cause the machine to activate), which younger participants are more willing to consider. In brief, young children appear to be more open and flexible learners. Gopnik and colleagues (2017) argue that this openness to new patterns is particularly useful early on in development, when children need to learn the structure of a new environment. It is interesting to note that, in contrast, tasks that typically yield improvements with age often are those that favor more efficient, rapid, and skilled performance (e.g., executive function, attentional focus).

Another area where age per se does not tell us about a person's performance level involves expertise. Childhood expertise can start early, with some children exhibiting intense special interests (e.g., vehicles, dinosaurs, blenders) as early as two years of age (DeLoache, Simcock, & Macari, 2007). By 4–10 years of age, children can attain expert levels of knowledge that approach or even exceed those of most adults. Expertise in childhood is associated with improved memory performance and reduction of the standard age differences, though these benefits are restricted to the domain of expertise, thus ruling out individual differences in broad cognitive abilities (e.g., Chi, 1978; Schneider, 1996). For example, child chess experts display better memory for pieces on a chessboard than adult novices, but this advantage doesn't carry over to their memory with non-chess information (e.g., digit span).

The Role of Social Context

As we have noted, a distinctive feature of human intelligence is the capacity for flexible learning that is highly adaptive to variable contexts (see Mayr, 1974). This flexibility leaves open a substantial role for early learning and intellectual development to be shaped by the social world, including information from other people, and parent-child relationships.

Much of intelligent thought and behavior involves learning from those around us, permitting increasingly sophisticated solutions over generations (Tomasello, 1999). Rather than accepting any information indiscriminately from any person, children evaluate the individuals who provide information, as well as the information itself, as part of their learning process (Harris et al., 2018). Preschool-age children tend to trust people who are familiar, such as their parents or familiar teachers (Corriveau & Harris, 2009; Corriveau et al., 2009b). However, even when children meet someone for the first time, they quickly form impressions of the informant and the quality of their information. Studies that examine preschool-age children's spontaneous learning from others are often conducted in the laboratory, in which different, unfamiliar informants present different types of information. For instance, in some studies children view informants who vary on a key feature, and then children are asked from whom they would prefer to learn. In this paradigm, preschool-aged children tend to prefer to learn new information from people who have been reliable in the past – they prefer to learn a word for an unfamiliar object from a person who correctly labeled a ball as "a ball" than a person who incorrectly labeled a ball as "a shoe" (Brosseau-Liard & Birch, 2010). By age four, children can make an even more sophisticated judgment about others' knowledge states – they consider whether individuals received help from others or could generate information without help

(Einav & Robinson, 2011), the number of people supporting a particular claim (Corriveau, Fusaro, & Harris, 2009a), and whether someone might have special expertise in a domain (Koenig & Jaswal, 2011). Altogether, children display a striking early capacity to evaluate the information presented to them by social partners.

In addition to the content and veracity of an informant's claims, children are sensitive to an informant's more enduring properties when deciding from whom to learn. Children prefer to learn from adults over children (Nguyen, 2012), those that are nice or honest over those that are mean or dishonest (Lane, Wellman, & Gelman, 2012), people of normal weight or not physically disabled over those who are overweight or physically disabled (Jaffer & Ma, 2015), people who are physically attractive or strong over those who are not (Bascandziev & Harris, 2013), and people who are part of the children's own social group, such as people who speak the children's native language or match children's racial or ethnic group (Gaither et al., 2014; Kinsler, Corriveau, & Harris, 2011). The variety of factors that children use when judging a person's information suggests that they are motivated from an early age to learn accurate and culturally relevant information to build their own store of knowledge.

In some ways, preschool-age children are impressively savvy learners. They display an early sensitivity to the characteristics of potential informants and the quality of their information. These properties could be a signal of expertise or cultural knowledge – children may rightly infer that an adult would know more detailed or complex information (e.g., information about nutrition) than a child (VanderBorght & Jaswal, 2009), or that someone who speaks the child's language would know how things are done in the child's community, compared to someone from a different background (Kinsler, Corriveau, & Harris, 2011). However, there are also stark limitations in children's early learning from other people. In the absence of alternative testimony, children will accept testimony from an unreliable informant (Vanderbilt, Heyman, & Liu, 2014), and children do not always realize the limits to their own knowledge. For instance, even though four- and five-year-olds can correctly assign questions based on their topic to relevant experts (e.g., questions about medicine to a doctor or about firefighting to a firefighter), they often try to answer the questions themselves (despite their lack of knowledge and the availability of relevant experts) when given the option to do so (Aguiar, Stoess, & Taylor, 2012). In addition, many of the properties that children use to make these judgments are unrelated to a person's expertise or the quality of information they provide. For instance, though someone who is physically strong may have expertise in certain domains (e.g., physical fitness or weight lifting), there is no reason to believe that they would be more knowledgeable than anyone else as a general rule. Though children are expert learners, this evidence suggests that children are susceptible to bias in their learning process.

Children's early ability to selectively learn from other people is central to the broader process of cultural transmission, a key feature of human learning. An interesting manifestation of this tendency is the phenomenon of over-imitation – children across cultures and contexts often imitate nonessential or inefficient features of a task (Nielsen et al., 2014). For instance, in one study children were presented with a complex puzzle box and watched a researcher engage in several behaviors to open the box and retrieve a prize (Lyons, Young, & Keil, 2007), including behaviors that were relevant to retrieving the prize (e.g., using a wand to push out a bolt holding the box shut) and some that were irrelevant (e.g., tapping a wand on the top of the box). Children tended to imitate all actions, even when explicitly warned to ignore unnecessary actions. In a related study, after observing a person use their head (rather than their hands) to turn on a light, children imitated that behavior (Gergely, Bekkering, & Király, 2002). However, children did not imitate the behavior if there was an apparent reason the person used their head to turn on the light, such as having their hands occupied; in this case, children used the more typical and efficient motion. Relatedly, infants view these ritual actions as a cue to social relationships: sixteen-month-old infants expect people who perform the same (unnecessary) ritual action (i.e., using their head to turn on a light) to affiliate with each other (Liberman, Kinzler, & Woodward, 2018).

These behaviors may appear silly – why imitate extraneous or highly inefficient behaviors? Nonetheless, over-imitation is observed robustly and across cultures with different experiences and ecologies (Nielsen et al., 2014). At least two explanations can account for this seemingly erroneous behavior. First, instead of building children's knowledge base, these behaviors may instead serve an important social function (Wen, Herrmann, & Legare, 2016). For instance, children produce more irrelevant actions when the model who demonstrated those actions is present (Nielsen & Blank, 2011), are more likely to imitate after experiencing social exclusion (Over & Carpenter, 2009), and report stronger in-group affiliation after engaging in causally irrelevant rituals (Wen et al., 2016). Second, imitation facilitates skill acquisition that, in some cases, serves as an important precursor to innovation (Legare & Nielsen, 2015). For instance, children struggled to innovate a tool (e.g., bending a straight pipe cleaner) to solve a task (e.g., pull a bucket out of a narrow tube to retrieve a sticker) until 9–10-years of age, but children ranging from three to ten years of age were immediately able to imitate an adult who solved this problem (Nielsen, 2013). In sum, a series of basic social learning processes, including the tendency to determine from whom to learn and to imitate the behaviors of others, play an important role in the development of children's learning across cultures and contexts.

In addition to learning from their broader social community, the parent-child bond has important implications for children's intellectual development. As noted earlier, children tend to trust familiar and accurate individuals. However, their trust in their mother's claims varies depending on their attachment status (Corriveau et al., 2009b). Children who had been classified as insecure-avoidant at fifteen months were less reliant on their mother's testimony at 4–5 years, whereas those who had been classified as insecure-resistant were more reliant on their mother's testimony, even when that testimony was misleading.

Children also learn the attitudes and stereotypes of those around them, even when not intending to and when not explicitly taught. For instance, the extent to which parents count or label sets of objects predicts their children's cardinal number knowledge, particularly when they are counting large sets that go beyond children's ability to individuate and track objects (Gunderson & Levine, 2011), and everyday conversations during meals can support children's math skills (Susperreguy & Davis-Kean, 2016). On the other hand, parents and teachers can also transmit their anxieties about math, with negative consequences for children's math learning. In one study, parents' math anxiety predicted their children's math learning across the school year in first and second grade, but only if parents who were high in math anxiety helped their children with their homework (Maloney et al., 2015). Similarly, when their teachers were more math-anxious, first- and second-grade girls (but not boys) were more likely to endorse the stereotype "boys are good at math and girls are good at reading" and to have lower math achievement (Beilock et al., 2010). These studies highlight just a few ways in which parents and other adults can both support and (unintentionally) hinder children's learning and achievement, with important consequences for their intellectual development across domains.

Policy Implications

The development of intelligence in childhood, and the factors that contribute to that development, have important real-world consequences. Specifically, children's capacity to learn and think critically is relevant to consider their participation in the justice system, how they consider their own intellectual potential (e.g., how they react to societal stereotypes), and how to promote their health and well-being (see also Nickerson, Chapter 10 in this volume, for a review of developing intelligence through instruction, and Barnett et al., Chapter 40 in this volume, regarding society and intelligence).

Children are important actors in the justice system, raising questions as to when and how children are capable of testifying in legal settings or being held responsible for their actions (see Lamb, Malloy, Hershkowitz, & La Rooy, 2015). Children's ability to accurately remember the source of their memories improves with age: Although participants in all age groups tested (ranging from first-graders to college students) made some source misattribution errors (i.e., claiming to remember seeing something they did not actually see in a video), the error rate was much higher among the youngest participants and was especially pronounced after a one-week delay (Ackil & Zaragoza, 1995). As such, relying on children's eyewitness testimony (particularly the testimony of young children) raises questions as to the accuracy of their testimony and their susceptibility to persuasion. Outside of criminal investigations, children are also asked to testify in family matters, specifically to give testimony on the parent with whom they should live with in custody disputes (Amato & Dorius, 2010). In addition to consideration of children's participation in the legal system as witnesses or parties of interest, important debates surround the age at which children should be punished for breaking the law and the appropriate nature of punishment. Until recently, people could be sentenced to death for crimes they committed before the age of eighteen, but this practice was barred by the U.S. Supreme Court as cruel and unusual punishment, and was extended to age twenty-one in Kentucky based on recent evidence about the still-developing adolescent brain (see Dahl, 2004). These actions suggest that social science researchers can make important contributions to support and protect children who interact with the legal system.

In contrast to questions pertaining to children's ability to participate in the justice system, recent scholarship has turned to intuitive jurisprudence – an approach to legal scholarship that focuses on our earliest concepts of obligations, rights, and punishment – to understand how intuitions about our legal system develop, change, and influence public policy (see Bregant, Shaw, & Kinzler, 2016). Developmental psychologists in this field consider the beliefs of children in the context of morality and punishment as providing key insights into adults' intuitions and reasoning. For instance, by the age of six, children believe that punishment has the function of deterrence, but in a specific way: Children reported that a character who stole was less likely to steal again if the character was punished, but there was no effect on victims and punishers who saw that the thief was punished (i.e., children expected the punishment to reduce recidivism, but not to broadly deter that crime; Bregant et al., 2016). People who accidentally harm others are treated differently by the legal system than people who harm others on purpose (e.g., manslaughter vs. murder), yet judges and juries still punish accidental transgressors. It is not clear that young children grasp this distinction, however. For example, before young children pass a false-belief task, they tend to attribute negative intentions to accidental transgressors (e.g., a person who throws away a bag, not realizing that it contains their classmate's cupcake) and are more willing to punish accidental transgressors (Killen et al., 2011).

In addition to the processes we have described thus far, children's conception of their own intellectual potential may have important consequences for their achievement (see also Carr & Dweck, Chapter 44 in this volume). Beliefs about the ability to succeed in an academic discipline have far-reaching and early-emerging consequences. In a recent study, academics from a variety of disciplines, including STEM (science, technology, engineering, and mathematics) fields and the humanities, were surveyed about their beliefs about what it takes to succeed in their discipline (Leslie et al., 2015). Fields in which its members believed that inherent brilliance was required to be successful were less likely to include women and African Americans, and these beliefs better explained the distribution of women in those fields than other common explanations for these distributions, including whether the field is a STEM discipline or not and the number of hours people tended to work on campus (rather than working remotely) per week. This pattern was also found in teaching evaluations: Students tended to use the words "brilliant" or "genius" in their reviews on RateMyProfessors.com in fields in which women and African Americans are underrepresented (Storage et al., 2016). Signatures of these beliefs can be observed as early as six years of age (Bian, Leslie, & Cimpian, 2017). In one study, five- to seven-year-old children heard a story about a person who was "really, really smart" (without any information about that person's gender) and selected the story's protagonist from an array of four pictures (two men and two women). By age six, girls were less likely to select a picture of a woman as the "really, really smart" person, even though they were more likely to select a picture of a girl than a boy when asked to pick out who "gets the best grades in school." By age six, girls were also less likely to choose a game for "children who are really, really smart" compared to a game for "children who try really, really hard." These findings suggest that very early in development, early in their school careers, girls have already developed the belief that girls are less likely to demonstrate brilliance than boys, regardless of their actual intellectual ability or willingness to try hard at a task. These beliefs may have cascading effects across the life span that contribute to later educational attainment and explain underrepresentation in fields that are thought to require "brilliance."

Finally, understanding children's early learning capacities has important consequences for child health. In the United States, thousands of children are hospitalized annually from influenza and related consequences (Centers for Disease Control and Prevention, 2016), childhood obesity is alarmingly high (Cunningham, Kramer, & Narayan, 2014), and children are not meeting recommendations for fruit and vegetable intake (Kim et al., 2014). As such, studying the development of children's learning and reasoning about health, illness, and food has important practical as well as theoretical consequences. Children's early reasoning about illness unfolds over a protracted period of time in early childhood, as the causes of illness are complex and opaque. Young children are often willing to ingest things that older children and adults would avoid: A majority of 16–29-month-old children in one study were willing to put disgusting or dangerous items in their mouths, such as imitation feces (made by mixing peanut butter with limburger cheese; Rozin et al., 1986). Three- and four-year-old children also struggle to make rational predictions about who might get sick in a series of stories about contamination (Legare, Wellman, & Gelman, 2009) and are willing to eat food that appeared to be sneezed on by another person (DeJesus et al., 2015). Four- and five-year-old children fail to avoid interacting with a person who visibly demonstrated signs of illness, and children's understanding that one could get a cold (but not a broken arm) from another person was a better predictor of whether children avoided a sick person than their age (Blacker & LoBue, 2016).

In the food domain, social cues shape early learning and food choices (see DeJesus, Shutts, & Kinzler, 2018; Shutts, Kinzler, & DeJesus, 2013). For example, young children are attuned to the food choices of others and make food selections based on what they observe others doing (Birch, 1980; Cruwys, Bevelander, & Hermans, 2015; DeJesus, Shutts, & Kinzler, 2018). Children are also susceptible to marketing strategies that include social cues. For instance, children are more likely to select foods that include popular cartoon characters (Roberto et al., 2010). Although these marketing strategies are not inherently harmful to children's health, the advertisements that children view are far more likely to feature processed foods with high levels of fat, sugar, and sodium, than fruits, vegetables, and other healthy foods (Batada & Wootan, 2007; Chapman et al., 2006). Despite attempts to regulate and reduce the number of advertisements for junk foods that children view (Abbasi, 2017), children are more likely to eat the foods featured in the advertisements that they see, and their advertisement viewing had a unique effect even after controlling for socioeconomic status and total television viewing time (Dalton et al., 2017; Longacre et al., 2017). In addition, the cognitive capacities to understand the persuasive intent of advertising develop during early childhood (with the understanding that the purpose of adverting is to sell a product coming online earlier and that advertisers may provide positively biased information to sell that product coming online later), suggesting that children may be especially susceptible to advertisers' messages (Moses & Baldwin, 2005). Children's ability to understand lying and to privilege information from "helpers" over information from "trickers" emerges over a similar developmental period (Vanderbilt, Liu, & Heyman, 2011). Not only do advertisements have effects on children's preferences and requests in the moment, but adults evaluate products more positively the earlier they were exposed to characters used to advertise those products (Connell, Brucks, & Nielsen, 2014). Ultimately, these findings suggest that understanding children's intellectual capacities is important to contextualize their understanding of health-related behaviors.

Conclusions

We argue in this chapter that a key feature of childhood is the capacity to take in, organize, and process information in a manner that gives rise to a variety of intelligent behaviors and modes of reasoning. Although children lack content knowledge and experience, they are experts at learning – and sometimes demonstrate even better learning potential than adults. Children's learning is situated in the social world, which allows children to selectively learn from other people and engage in the process of cultural transmission. The development of intelligence in childhood also has important implications for several domains of public policy, including their participation in the legal system, their beliefs about their own capabilities, and the development of healthy behaviors.

Throughout, we have primarily referred to capabilities or challenges demonstrated by "children" in a generic way – that is, we have described the abilities of children as a broad group. Of course, there are a variety of factors that contribute to children's learning and reasoning in the domains we have described, and this variability can be observed at many levels of analysis (Siegler, 2007). As noted earlier, an individual child may switch between different strategies to solve problems, ranging from how to descend a ramp (Adolph et al., 1997), which tool to use to achieve a goal (Chen & Siegler, 2000), and what strategy to use to succeed at a false-belief task (Flynn, O'Malley & Wood, 2004), rather than using the same strategy consistently. In addition to variability within a child, intelligence exhibits important between-child variability as well. Sources of such variability are the focus of several other chapters in this volume, most notably interactions between genes and the environment (Chapter 6), intellectual disabilities (Chapter 11), biological factors that give rise to intelligence (Chapter 19), and environmental effects (Chapter 41). In all these cases, childhood is an especially important time period to consider. Children are especially at risk from exposure to environmental contaminants based on their intake of food, water, and air as a proportion of their body weight (compared to adults), and the effect of exposure to these contaminants early in development cascades into adulthood (Trentacosta et al., 2016).

Finally, we point to the need for more research that examines a wider range of cultural contexts when characterizing intelligence in children. This chapter is a largely Western perspective on intelligence, as is the majority of research in child development (Nielsen et al., 2017) and psychology more broadly (Henrich, Heine, & Norenzayan, 2010). Yet, this focus on "WEIRD" populations (Western, educated, industrialized, rich, and democratic; Henrich et al., 2010) overlooks the experiences and perspectives of the majority of the world's population. Cross-cultural studies reveal that different aspects of intelligence may be valued differently by different communities (Rogoff, 2003). For instance, one study found that adults in Vanuatu (a Melanesian island nation) were relatively more likely than adults in the United States to evaluate children's intelligence on the basis of conformity, and relatively less likely to evaluate their intelligence on the basis of creativity (Clegg, Wen, & Legare, 2017). Examining intelligence from different cultural perspectives is critical to develop a more nuanced understanding of the cultural values, practices, and experiences of children around the world.

References

Abbasi, J. (2017). Junk food ads reach children despite food industry self-regulation. *Journal of the American Medical Association*, *317*(23), 2359–2361. https://doi.org/10.1001/jama.2017.4653

Ackil, J. K., & Zaragoza, M. S. (1995). Developmental differences in eyewitness suggestibility and memory for source. *Journal of Experimental Child Psychology*, *60*(1), 57–83. https://doi.org/10.1006/jecp.1995.1031

Adolph, K. E., Bertenthal, B. I., Boker, S. M., Goldfield, E. C., & Gibson, E. J. (1997). Learning in the development of infant locomotion. *Monographs of the Society for Research in Child Development*, *162*(3), i–vi, 1–158. https://doi.org/10.2307/1166199

Adolph, L. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. *Annual Review of Psychology*, *70*, 141–164. https://doi.org/10.1146/annurev-psych-010418-102836

Aguiar, N. R., Stoess, C. J., & Taylor, M. (2012). The development of children's ability to fill the gaps in their knowledge by consulting experts. *Child Development*, *83*(4), 1368–1381. https://doi.org/10.1111/j.1467-8624.2012.01782.x

Amato, P. R., & Dorius, C. (2010). Fathers, children, and divorce. In M. E. Lamb (Ed.), *The role of the father in child development* (5th ed., pp. 177–200). Hoboken, NJ: Wiley.

Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, *21*(4), 547–569.

Bascandziev, I., & Harris, P. L. (2013). In beauty we trust: Children prefer information from more attractive informants. *British Journal of Developmental Psychology*, *32*(1), 94–99. https://doi.org/10.1111/bjdp.12022

Batada, A., & Wootan, M. G. (2007). Nickelodeon markets nutrition-poor foods to children. *American Journal of Preventive Medicine*, *33*(1), 48–50. https://doi.org/10.1016/j.amepre.2007.02.035

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860–1863. https://doi.org/10.1073/pnas.0910967107

Bering, J. M., Blasi, C. H., & Bjorklund, D. F. (2005). The development of afterlife beliefs in religiously and secularly schooled children. *British Journal of Developmental Psychology*, *23*(4), 587–607. https://doi.org/10.1348/026151005X36498

Bian, L., Leslie, S.-J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, *355*(6323), 389–391. https://doi.org/10.1126/science.aah6524

Birch, L. L. (1980). Effects of peer models' food choices and eating behaviors on preschoolers' food preferences. *Child Development*, *51*(2), 489–496. https://doi.org/10.2307/1129283

Bjorklund, D. F., Dukes, C., & Brown, R. D. (2008). The development of memory strategies. In M. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 157–188). Hove, UK: Psychology Press.

Bjorklund, D. F., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist*, *47*, 46–54. https://doi.org/10.1037/0003-066X.47.1.46

Blacker, K.-A., & LoBue, V. (2016). Behavioral avoidance of contagion in childhood. *Journal of Experimental Child Psychology*, *143*, 162–170. https://doi.org/10.1016/j.jecp.2015.09.033

Bornstein, M. H., Hahn, C. S., & Suwalsky, J. T. (2013). Physically developed and exploratory young infants contribute to their own long-term academic achievement. *Psychological Science*, *24*(10), 1906–1917. https://doi.org/10.1177/0956797613479974

Bregant, J., Shaw, A., & Kinzler, K. D. (2016). Intuitive jurisprudence: Early reasoning about the functions of punishment. *Journal of Empirical Legal Studies*, *13*(4), 693–717. https://doi.org/10.1111/jels.12130

Brosseau-Liard, P. E., & Birch, S. A. (2010). "I bet you know more and are nicer too!": What children infer from others' accuracy. *Developmental Science*, *13*(5), 772–778. https://doi.org/10.1111/j.1467-7687.2009.00932.x

Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. (2000). Travel broadens the mind. *Infancy*, *1*(2), 149–219.

Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.

Centers for Disease Control and Prevention (2016). Influenza (Flu): Information for Schools & Childcare Providers. www.cdc.gov/flu/school.

Chapman, K., Nicholas, P., Banovic, D., & Supramaniam, R. (2006). The extent and nature of food promotion directed to children in Australian supermarkets. *Health Promotion International*, *21*(4), 331–339. https://doi.org/10.1093/heapro/dal028

Chen, Z., & Siegler, R. S. (2000). *Intellectual development in childhood*. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 92–116). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511807947.006

Chi, M. T. (1978). Knowledge structures and memory development. In R. Siegler (Ed.), *Children's thinking: What develops?* (pp. 75–96). Mahwah, NJ: Erlbaum.

Clegg, J. M., Wen, N. J., & Legare, C. H. (2017). Is non-conformity WEIRD? Cultural variation in adults' beliefs about children's competency and conformity. *Journal of Experimental Psychology: General*, *146*(3), 428–441. https://doi.org/10.1037/xge0000275

Cohen, D. J., & Sarnecka, B. W. (2014). Children's number-line estimation shows development of measurement skills (not number representations. *Developmental Psychology*, *50*(6), 1640–1652. https://doi.org/10.1037/a0035901

Connell, P. M., Brucks, M., & Nielsen, J. H. (2014). How childhood advertising exposure can create biased product evaluations that persist into adulthood. *Journal of Consumer Research*, *41*(1), 119–134. https://doi.org/10.1086/675218

Corriveau, K. H., Fusaro, M., & Harris, P. L. (2009). Going with the flow: Preschoolers prefer nondissenters as informants. *Psychological Science*, *20*(3), 372–377. https://doi.org/10.1111/j.1467-9280.2009.02291.x

Corriveau, K. H., & Harris, P. L. (2009). Choosing your informant: Weighing familiarity and recent accuracy. *Developmental Science*, *12*(3), 426–437. https://doi.org/10.1111/j.1467-7687.2008.00792.x

Corriveau, K. H., Harris, P. L., Meins, E., Fernyhough, C., Arnott, B., Elliott, L., et al. (2009b). Young children's trust in their mother's claims: Longitudinal links with attachment security in infancy. *Child Development*, *80*(3), 750–761. https://doi.org/10.1111/j.1467-8624.2009.01295.x

Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science*, *11*(2), 239–264. https://doi.org/10.1177/1745691615621279

Cruwys, T., Bevelander, K. E., & Hermans, R. C. (2015). Social modeling of eating: A review of when and why social influence affects food intake and choice. *Appetite*, *86*, 3–18. https://doi.org/10.1016/j.appet.2014.08.035

Cunningham, S. A., Kramer, M. R., & Narayan, K. V. (2014). Incidence of childhood obesity in the United States. *New England Journal of Medicine*, *370*(5), 403–411. https://doi.org/10.1056/NEJMoa1309753

Dahl, R. E. (2004). Adolescent brain development: A period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences*, *1021*(1), 1–22. https://doi.org/10.1196/annals.1308.001

Dalton, M. A., Longacre, M. R., Drake, K. M., Cleveland, L. P., Harris, J. L., et al. (2017). Child-targeted fast-food television advertising exposure is linked with fast-food intake among pre-school children. *Public Health Nutrition*, *20*(9), 1548–1556. https://doi.org/10.1017/S1368980017000520

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: On the deceptive determinism of DNA. *Psychological Bulletin*, *137*(5), 800–818. https://doi.org/10.1037/a0021860

DeJesus, J. M., Kinzler, K. D., & Shutts, K. (2018). Food cognition and nutrition knowledge. In J. O. Fisher & J. C. Lumeng (Eds.), *Pediatric food preferences and eating behaviors* (pp. 271–288). New York: Elsevier. https://doi.org/10.1016/B978-0-12-811716-3.00014-2

DeJesus, J. M., Shutts, K., & Kinzler, K. D. (2015). Eww she sneezed! Contamination context affects children's food preferences and consumption. *Appetite*, *87*, 303–309. https://doi.org/10.1016/j.appet.2014.12.222

DeJesus, J. M., Shutts, K., & Kinzler, K. D. (2018). Mere social knowledge impacts children's consumption and categorization of foods. *Developmental Science*. https://doi.org/10.1111/desc.12627

DeLoache, J. S., Simcock, G., & Macari, S. (2007). Planes, trains, automobiles – and tea sets: Extremely intense interests in very young children. *Developmental Psychology*, *43*(6), 1579–1586. https://doi.org/10.1037/0012-1649.43.6.1579

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Eidson, R. C., & Coley, J. D. (2014). Not so fast: Reassessing gender essentialism in young adults. *Journal of Cognition and Development*, *15*(2), 382–392. https://doi.org/10.1080/15248372.2013.763810

Einav, S., & Robinson, E. J. (2011). When being right is not enough: Four-year-olds distinguish knowledgeable informants from merely accurate informants. *Psychological Science*, *22*(10), 1250–1253. https://doi.org/10.1177/0956797611416998

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives*, *7*(2), 74–79. https://doi.org/10.1111/cdep.12019

Finn, A. S., Lee, T., Kraus, A., & Kam, C. L. H. (2014). When it hurts (and helps) to try: The role of effort in language learning. *PloS One*, *9*(7), e101806. https://doi.org/10.1371/journal.pone.0101806

Flavell, J. H. (1963). *The developmental psychology of Jean Piaget*. Princeton: D. Van Nostrand.

Flynn, E., O'Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science*, *7*(1), 103–115. https://doi.org/10.1111/j.1467-7687.2004.00326.x

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824. https://doi.org/10.1016/j.cognition.2008.04.007

Furth, H. G. (1973). Piaget, IQ and the nature-nurture controversy. *Human Development*, *16* (1–2), 61–73. https://doi.org/10.1159/000271267

Gaither, S. E., Chen, E. E., Corriveau, K. H., Harris, P. L., Ambady, N., & Sommers, S. R. (2014). Monoracial and biracial children: Effects of racial identity saliency on social learning and social preferences. *Child Development*, *85*(6), 2299–2316. https://doi.org/10.1111/cdev.12266

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.

Gelman, S. A., & Legare, C. H. (2011). Concepts and folk theories. *Annual Review of Anthropology*, *40*, 379–398. https://doi.org/10.1146/annurev-an40

Gelman, S. A., & Marchak, K. A. (in press). How does intuition mislead? The role of human bias in scientific inquiry. In K. McCain & K. Kampourakis (Eds.), *What is scientific knowledge? An introduction to contemporary epistemology of science*. London: Routledge.

Gelman, S. A., & Rhodes, M. (2012). "Two-thousand years of stasis": How psychological essentialism impedes evolutionary understanding. In K. S. Rosengren, S. Brem, E. M. Evans, & G. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution* (pp. 3–21). New York: Oxford University Press.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: interrelations in development* (pp. 225–277). New York: Cambridge University Press.

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, *415*(6873), 755. https://doi.org/10.1038/415755a

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve theory of rational action. *Trends in Cognitive Sciences*, *7*(7), 287–292. https://doi.org/10.1016/S1364-6613(03)00128-1

Giménez-Dasí, M., Guerrero, S., & Harris, P. L. (2005). Intimations of immortality and omniscience in early childhood. *European Journal of Developmental Psychology*, *2*(3), 285–297. https://doi.org/10.1080/17405620544000039

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., et al. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences*, 201700811. https://doi.org/10.1073/pnas.1700811114

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language*, *7*(1–2), 145–171. https://doi.org/10.1111/j.1468-0017.1992.tb00202.x

Gunderson, E. A., & Levine, S. C. (2011). Some types of parent number talk count more than others: Relations between parents' input and children's cardinal-number knowledge. *Developmental Science*, *14*(5), 1021–1032. https://doi.org/10.1111/j.1467-7687.2011.01050.x

Gweon, H., & Schulz, L. (2011). 16-month-olds rationally infer causes of failed actions. *Science*, *332*(6037), 1524–1524. https://doi.org/10.1126/science.1204493

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Harris, P. L. (2011). Conflicting thoughts about death. *Human Development*, *54*(3), 160–168. https://doi.org/10.1159/000329133

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, *69*, 251–273. https://doi.org/10.1146/annurev-psych-122216-011710

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263–277. https://doi.org/10.1016/j.cognition.2018.04.007

Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition*, *50* (1–3), 171–188. https://doi.org/10.1016/0010-0277(94)90027-2

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466* (7302), 29. https://doi.org/10.1038/466029a

Hofman, A. D., Visser, I., Jansen, B. R., & van der Maas, H. L. (2015). The balance-scale task revisited: A comparison of statistical models for rule-based and information-integration theories of proportional reasoning. *PloS One*, *10*(10), e0136449. https://doi.org/10.1371/journal.pone.0136449

Horne, Z., Powell, D., Hummel, J. E., & Holyoak, K. J. (2015). Countering antivaccination attitudes. *Proceedings of the National Academy of Sciences*, *112*(33), 10321–10324. https://doi.org/10.1073/pnas.1504019112

Jaffer, S., & Ma, L. (2015). Preschoolers show less trust in physically disabled or obese informants. *Frontiers in Psychology*, *5*, 1524. https://doi.org/10.3389/fpsyg.2014.01524

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition*, *140*, 14–23. https://doi.org/10.1016/j.cognition.2015.03.006

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.

Kail, R. V., Lervåg, A., & Hulme, C. (2016). Longitudinal evidence linking processing speed to the development of reasoning. *Developmental Science*, *19*(6), 1067–1074. https://doi.org/10.1111/desc.12352

Kam, C. L. H., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*(1), 30–66. https://doi.org/10.1016/j.cogpsych.2009.01.001

Kaufman, A. S. (1997). *KBIT-2: Kaufman Brief Intelligence Test*. Minneapolis, MN: NCS Pearson.

Keen, R. (2011). The development of problem solving in young children: A critical cognitive skill. *Annual Review of Psychology*, *62*, 1–21. https://doi.org/10.1146/annurev.psych.031809.130730

Kelemen, D. (2012). Teleological minds. In K. S. Rosengren, S. Brem, E. M. Evans, & G. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution* (pp. 66–92). New York: Oxford University Press.

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General*, *142*(4), 1074–1083. https://doi.org/10.1037/a0030399

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, *126*(1), 109–114. https://doi.org/10.1016/j.cognition.2012.08.004

Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition*, *119*(2), 197–215. https://doi.org/10.1016/j.cognition.2011.01.006

Kim, S. A., Moore, L. V., Galuska, D., Wright, A. P., Harris, D., Grummer-Strawn, L. M., et al. (2014). Vital signs: Fruit and vegetable intake among children – United States, 2003–2010. *Morbidity and Mortality Weekly Report*, *63*(31), 671–676.

Koenig, M. A., & Jaswal, V. K. (2011). Characterizing children's expectations about expertise and incompetence: Halo or pitchfork effects? *Child Development*, *82*(5), 1634–1647. https://doi.org/10.1111/j.1467-8624.2011.01618.x

Lamb, M. E., Malloy, L. C., Hershkowitz, I., & La Rooy, D. (2015). Children and the law. In M. E. Lamb (Ed.), *Handbook of child psychology and developmental science*, vol. *3*, *Social, emotional, and personality development* (7th ed., pp. 464–512). Hoboken NJ: Wiley. https://doi.org/10.1002/9781118963418.childpsy312

Lane, J. D., Wellman, H. M., & Gelman, S. A. (2012). Informants' traits weigh heavily in young children's trust in testimony and in their epistemic inferences. *Child Development*, *84*(4), 1253–1268. https://doi.org/10.1111/cdev.12029

Legare, C. H., Evans, E. M., Rosengren, K. S., & Harris, P. L. (2012). The coexistence of natural and supernatural explanations across cultures and development. *Child Development*, *83*(3), 779–793. https://doi.org/10.1111/j.1467-8624.2012.01743.x

Legare, C. H., & Gelman, S. A. (2008). Bewitchment, biology, or both: The co-existence of natural and supernatural explanatory frameworks across development. *Cognitive Science*, *32*(4), 607–642. https://doi.org/10.1080/03640210802066766

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Sciences*, *19*(11), 688–699. https://doi.org/10.1016/j.tics.2015.08.005

Legare, C. H., Wellman, H. M., & Gelman, S. A. (2009). Evidence for an explanation advantage in naïve biological reasoning. *Cognitive Psychology*, *58*(2), 177–194. https://doi.org/10.1016/j.cogpsych.2008.06.002

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 262–265. https://doi.org/10.1126/science.1261375

Liberman, Z., Kinzler, K. D., & Woodward, A. L. (2018). The early social significance of shared ritual actions. *Cognition*, *171*, 42–51. https://doi.org/0.1016/j.cognition.2017.10.018

Libertus, K., Joh, A. S., & Needham, A. W. (2016). Motor training at 3 months affects object exploration 12 months later. *Developmental Science*, *19*, 1058–1066. https://doi.org/10.1111/desc.12370

Longacre, M. R., Drake, K. M., Titus, L. J., Harris, J., Cleveland, L. P., Langeloh, G., et al. (2017). Child-targeted TV advertising and preschoolers' consumption of high-sugar breakfast cereals. *Appetite*, *108*, 295–302. https://doi.org/10.1016/j.appet.2016.10.014

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, *104*(50), 19751–19756. https://doi.org/10.1073/pnas.0704452104

Maloney, E. A., Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2015). Intergenerational effects of parents' math anxiety on children's math achievement and anxiety. *Psychological Science*, *26*(9), 1480–1488. https://doi.org/10.1177/0956797615592630

Mayr, E. (1974). Behavior programs and evolutionary strategies: Natural selection sometimes favors a genetically "closed" behavior program, sometimes an "open" one. *American Scientist*, *62*(6), 650–659.

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin and Review*, *18*(1), 46–60. https://doi.org/10.3758/s13423-010-0034-0

Moses, L. J., & Baldwin, D. A. (2005). What can the study of cognitive development reveal about children's ability to appreciate and cope with advertising? *Journal of Public Policy and Marketing*, *24*, 186–201. https://doi.org/10.1509/jppm.2005.24.2.186

Moshman, D., & Tarricone, P. (2016). Logical and causal reasoning. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 54–67). New York: Routledge.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11–28. https://doi.org/10.1207/s15516709cog1401_2

Nguyen, S. P. (2012). The role of external sources of information in children's evaluative food categories. *Infant and Child Development*, *21*(2), 216–235. https://doi.org/10.1002/icd.745

Nguyen, S. P., & Rosengren, K. S. (2004). Causal reasoning about illness: A comparison between European and Vietnamese-American children. *Journal of Cognition and Culture*, *4*(1), 51–78. https://doi.org/10.1163/156853704323074750

Nielsen, M. (2013). Young children's imitative and innovative behaviour on the floating object task. *Infant and Child Development*, *22*(1), 44–52. https://doi.org/10.1002/icd.1765

Nielsen, M., & Blank, C. (2011). Imitation in young children: When who gets copied is more important than what gets copied. *Developmental Psychology, 47*, 1050–1053. https://doi.org/10.1037/a0023866

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Nielsen, M., Mushin, I. Tomaselli, K., & Whiten, A. (2014). Where culture takes hold: "Overimitation" and its flexible deployment in Western, Aboriginal, and Bushmen children. *Child Development*, *85*(6), 2169–2184. https://doi.org/10.1111/cdev.12265

Over, H., & Carpenter, M. (2009). Priming third-party ostracism increases affiliative imitation in children. *Developmental Science*, *12*(3). https://doi.org/10.1111/j.1467-7687.2008.00820.x

Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The "Perceptual Wedge Hypothesis" as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, *121*(2), 130–143. https://doi.org/10.1016/j.bandl.2011.05.003

Piaget, J. (1968). Piaget's point of view. *International Journal of Psychology*, *3*(4), 281–299. https://doi.org/10.1080/00207596808246651

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41. https://doi.org/10.1016/j.cognition.2010.03.012

Raman, L., & Gelman, S. A. (2004). A cross-cultural developmental analysis of children's and adults' understanding of illness in South Asia (India) and the United States. *Journal of Cognition and Culture*, *4*(2), 293–317. https://doi.org/10.1163/1568537041725088

Rhodes, M., Gelman, S. A., & Brickman, D. (2010). Children's attention to sample composition in learning, teaching and discovery. *Developmental Science*, *13*(3), 421–429. https://doi.org/10.1111/j.1467-7687.2009.00896.x

Rhodes, M., & Wellman, H. (2013). Constructing a new theory from old ideas and new evidence. *Cognitive Science*, *37*(3), 592–604. https://doi.org/10.1111/cogs.12031

Richland, L. E., & Burchinal, M. R. (2013). Early executive function predicts reasoning development. *Psychological Science*, *24*(1), 87–92. https://doi.org/10.1177/0956797612450883

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, *94*(3), 249–273. https://doi.org/10.1016/j.jecp.2006.02.002

Roberto, C. A., Baik, J., Harris, J. L., & Brownell, K. D. (2010). Influence of licensed characters on children's taste and snack preferences. *Pediatrics*, *126*(1), 88–93. https://doi.org/10.1542/peds.2009-3433

Rogoff, B. (2003). *The cultural nature of human development*. New York: Oxford University Press.

Rosengren, K. S., Brem, S. K., Evans, E. M., & Sinatra, G. M. (Eds.) (2012). *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. New York: Oxford University Press.

Rosengren, K. S., Miller, P. J., Gutiérrez, I. T., Chow, P. I., Schein, S. S., Anderson, K. N., et al. (Eds.) (2014). *Children's understanding of death: Toward a contextualized and integrated account*. Monographs of the Society for Research in Child Development, vol. 79.1. Boston: John Wiley & Sons.

Rozin, P., Hammer, L., Oster, H., Horowitz, T., & Marmora, V. (1986). The child's conception of food: Differentiation of categories of rejected substances in the 16 months to 5 year age range. *Appetite*, *7*(2), 141–151. https://doi.org/10.1016/S0195-6663 (86) 80014-9

Sánchez Tapia, I., Gelman, S. A., Hollander, M., Manczak, E. M., Mannheim, B., & Escalante, C. (2016). Development of teleological explanations in Peruvian Quechua-speaking and US English-speaking preschoolers and adults. *Child Development*, *87*(3), 747–758. https://doi.org/10.1111/cdev.12497

Schneider, W. (1996). The effects of expertise and IQ on children's memory: When knowledge is, and when it is not enough. *International Journal of Behavioral Development*, *19*(4), 773–796. https://doi.org/10.1080/016502596385578

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*(5691), 1779–1782. https://doi.org/10.1126/science.1100199

Shtulman, A., & Calabi, P. (2012). Cognitive constraints on the understanding and acceptance of evolution. In K. S. Rosengren, S. Brem, E. M. Evans, & G. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution* (pp. 47–65). New York: Oxford University Press.

Shtulman, A., & Harrington, K. (2016). Tensions between science and intuition across the lifespan. *Topics in Cognitive Sciences*, *8*(1), 118–137. https://doi.org/10.1111/tops.12174

Shutts, K., Kinzler, K. D., & DeJesus, J. M. (2013). Infants' and children's social learning about foods: Previous research and new prospects. *Developmental Psychology*, *49*(3), 419–425. https://doi.org/10.1037/a0027551

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. https://doi.org/10.1111/j.1467-7687.2007.00571.x

Siegler, R. S., & Lortie-Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives*, *8*(3), 144–150. https://doi.org/10.1111/cdep.12077

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–250. https://doi.org/10.1111/1467-9280.02438

Solomon, G. E. A., & Cassimatis, N. L. (1999). On facts and conceptual systems: Young children's integration of their understandings of germs and contagion. *Developmental Psychology*, *35*(1), 113–126. https://doi.org/10.1037/0012-1649.35.1.113

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*(1), B1–B11. https://doi.org/10.1016/j.cognition.2004.07.004

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, *110*(45), 18116–18120. https://doi.org/10.1073/pnas.1302751110

Storage, D., Horne, Z., Cimpian, A., & Leslie, S. J. (2016). The frequency of "brilliant" and "genius" in teaching evaluations predicts the representation of women and African Americans across fields. *PloS One*, *11*(3), e0150194. https://doi.org/10.1371/journal.pone.0150194

Susperreguy, M. I., & Davis-Kean, P. E. (2016). Maternal math talk in the home and math skills in preschool children. *Early Education and Development*, *27*(6), 841–857. https://doi.org/10.1080/10409289.2016.1148480

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Trentacosta, C. J., Davis-Kean, P., Mitchell, C., Hyde, L., & Dolinoy, D. (2016). Environmental contaminants and child development. *Child Development Perspectives*, *10*(4), 228–233. https://doi.org/10.1111/cdep.12191

van der Ven, S. H. G., Boom, J., Kroesbergen, E. H., & Leseman, P. P. M. (2012). Microgenetic patterns of children's multiplication learning: Confirming the overlapping waves model by latent growth modeling. *Journal of Experimental Child Psychology*, *113*, 1–19. https://doi.org/10.1016/j.jecp.2012.02.001

Vanderbilt, K. E., Heyman, G. D., & Liu, D. (2014). In the absence of conflicting testimony young children trust inaccurate informants. *Developmental Science*, *17*(3), 443–451. https://doi.org/10.1111/desc.12134

Vanderbilt, K. E., Liu, D., & Heyman, G. D. (2011). The development of distrust. *Child Development*, *82*(5), 1372–1380. https://doi.org/10.1111/j.1467-8624.2011.01629.x

VanderBorght, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development*, *18*(1), 61–71. https://doi.org/10.1002/icd.591

Vendetti, M. S., Matlen, B. J., Richland, L. E., & Bunge, S. A. (2015). Analogical reasoning in the classroom: Insights from cognitive science. *Mind, Brain, and Education*, *9*(2), 100–106. https://doi.org/10.1111/mbe.12080

Ware, E. A., & Gelman, S. A. (2014). You get what you need: An examination of purpose-based inheritance reasoning in undergraduates, preschoolers, and biological experts. *Cognitive Science*, *38*(2), 197–243. https://doi.org/10.1111/cogs.12097

Wellman, H. M., & Gelman, S. A. (1998). Knowledge acquisition. In D. Kuhn & R. Siegler (Eds.), *Handbook of child psychology*, vol. *2*, *Cognitive development* (5th ed., pp. 523–573). New York: Wiley.

Wen, N. J., Herrmann, P. A., & Legare, C. H. (2016). Ritual increases children's affiliation with in-group members. *Evolution and Human Behavior*, *37*(1), 54–60. https://doi.org/10.1016/j.evolhumbehav.2015.08.002

Wendelken, C., Ferrer, E., Whitaker, K. J., & Bunge, S. A. (2015). Fronto-parietal network reconfiguration supports the development of reasoning ability. *Cerebral Cortex*, *26*(5), 2178–2190. https://doi.org/10.1093/cercor/bhv050

Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *25*(1), 121–133. https://doi.org/10.1016/S0163-6383(02)00093-0

Woolley, J. D., & Phelps, K. E. (2001). The development of children's beliefs about prayer. *Journal of Cognition and Culture*, *1*(2), 139–166. https://doi.org/10.1163/156853701316931380

Xu, F., & Kushnir, T. (2013). Infants are rational constructivist learners. *Current Directions in Psychological Science*, *22*(1), 28–32. https://doi.org/10.1177/0963721412469396

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272. https://doi.org/10.1037/0033-295X.114.2.245

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., et al. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*(3), i–151.

9 Intelligence in Adulthood

Christopher Hertzog

The field of gerontology – the scientific study of aging – emerged as a major scientific discipline in the twentieth century (e.g., Birren, 1964). Research on intelligence and intellectual development played a major role in shaping the field of psychological gerontology (e.g., Botwinick, 1977). This chapter reviews what is known and not yet known about adult intellectual development after decades of research on the topic. Most of the information we have available concerns aspects of what Sternberg (1985) has defined as academic intelligence (based on traditional psychometric tests of human abilities). This chapter focuses on what is known about these types of human abilities and their correlates, although I also briefly treat other aspects of intellect, such as practical intelligence and tacit knowledge.

Descriptive Research on Adult Age Differences

Early studies of psychometric intelligence prior to 1940 determined that there were large differences in performance on general tests of intellectual aptitude (see Salthouse, 1982 for an excellent summary and review). Wechsler (1939) characterized the performance tests on the Wechsler Adult Intelligence Scale (WAIS) as "don't-hold" tests because of the lower performance on those subscales (e.g., WAIS Block Design) by older adults in his cross-sectional norming studies of the test. Conversely, Wechsler found that tests like WAIS Vocabulary were typically shown to have much smaller age differences, causing them to be characterized as "hold" tests. This basic idea – that one class of intellectual ability tests manifests age decline, whereas others do not, has been widely replicated and studied across a variety of intelligence tests, and today represents a virtual "truism" about aging and intelligence. These findings mirrored outcomes of studies using other tests to evaluate age differences in human abilities, studies that spanned much of the twentieth century (Salthouse, 1982).

The concept of contrasting maintenance of knowledge and verbal abilities, relative to other types of human abilities, has therefore figured prominently in theoretical treatment of how aging affects intelligence. Cattell (1971) developed the theory of fluid and crystallized intelligence, arguing that this basic pattern reflected two prototypic classes of intellectual abilities. Fluid intelligence was seen as the fundamental ability to think, reason, and process information, and prone to adult age decline as a function of biological aging processes (Horn & Cattell, 1967; Horn & Hofer, 1992).

Crystallized intelligence, on the other hand, was seen as determined by investment of fluid intelligence in knowledge acquisition, which was largely maintained or even improved into old age (Horn & Cattell, 1967).

Baltes and his colleagues characterized the distinction as involving a decline in basic information-processing mechanisms labeled the mechanics of cognition (e.g., Baltes, 1997). In contrast, experience with a culture leads to acquisition of a broad class of declarative and procedural knowledge and skills about how to achieve goals in a cultural context, labeled the pragmatics of intelligence. Figure 9.1 shows the idealized functions thought to exist for the two classes of intellectual ability (Baltes, Staudinger, & Lindenberger, 1999). Although Baltes' conceptualization emphasized mechanisms that influence observed abilities, similar arguments were being made by Horn (e.g., Horn & Hofer, 1992) in extended versions of fluid-crystallized theory. As a consequence, the differences between these theoretical viewpoints are subtle at best.

Can a two-curve model actually account for most of the age-related variance in adult intellectual development? If so, it would be surprising, for several reasons. First, theories of psychometric abilities generally acknowledge that there are a large number of intellectual abilities. Theoretical approaches based on the work by Thurstone on primary mental abilities (e.g., Thurstone, 1938) typically argue for thirty or more primary abilities (Carroll, 1993; Horn & Hofer, 1992). It would be surprising if all these abilities declined at the same rate in adulthood. Second, contemporary hierarchical models of abilities typically acknowledge that fluid and crystallized intelligence are distinct from other higher-order ability factors. Horn (1985; Horn & Hofer, 1992) argued that, for example, general visualization abilities, general auditory abilities, speediness, and secondary memory are all empirically distinct from fluid intelligence. To the extent that these second-order factors are indeed differentiable from fluid intelligence, then one might expect their developmental curves in adulthood to also differ. Third, theories of biological aging identify a large number of potential biological clocks, operating at different levels of basic physiology, that appear to be associated with variations in rates of biological aging.

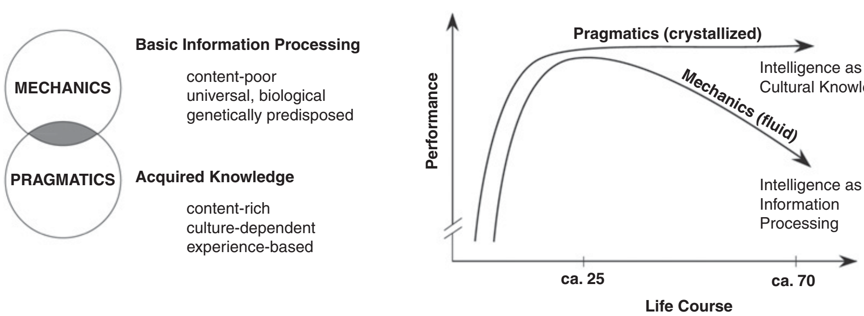

Figure 9.1 *Baltes' conception of the two curves of intellectual development, based on negative changes in cognitive mechanics and preservation of pragmatics during adulthood and aging (from Baltes et al., 1999).*

What do the empirical data tell us? The cross-sectional age curves for episodic memory, spatial visualization, and measures of fluid intelligence and general processing speed vary somewhat as a function of factors such as how the tests are constructed and scaled, their processing requirements, and the like. Yet there is surprising similarity in the curves across these different classes of abilities. Certainly the ability that is typically found to have the largest cross-sectional age differences is speed of processing, such as that identified by the perceptual-speed factor (Carroll, 1993). Salthouse (1996) has evaluated age differences in perceptual speed in a large number of studies, typically finding the largest cross-sectional age differences for that factor (see also Schaie, 1989). However, fluid intelligence shows considerable similarity in magnitude of estimated decline to measures of episodic memory, working memory, and spatial visualization (e.g., Hertzog, 1989; Hultsch et al., 1998; Park et al., 1996; Salthouse, Pink, & Tucker-Drob, 2008). No one study has examined all the relevant abilities in a truly representative sample of the adult population, and most studies observe at least some variation in cross-sectional age slopes across abilities. Nevertheless, the available cross-sectional evidence on the mechanics of cognition is more or less consistent with the argument that abilities emphasizing cognitive mechanics produce similar patterns of average decline in adulthood. There are important exceptions – not all processing mechanisms decline, and not all aspects of pragmatics are maintained (see Hertzog, 2008). Also, cross-sectional data disagree as to whether the cross-sectional curves are linear, or curvilinear – accelerating the magnitude of estimated decline in old age (e.g., compare Hultsch et al., 1998 with Park et al., 1996 regarding episodic memory). Nevertheless, the negative correlation of age with fluid intelligence, working memory, spatial visualization, and the like from early adulthood to old age is about −0.4.

There is evidence that the cross-sectional age curves for crystallized intelligence may differ as a function of the type of knowledge being assessed. Work by Ackerman and colleagues has focused on tracking domain-specific knowledge that may occur during and after the time that young adults begin to specialize their vocational and personal interests, crystallizing them into a pattern of preferences for information sought, acquired, digested, and assimilated into existing knowledge structures (e.g., Ackerman, 2000; Beier & Ackerman, 2005). Ackerman's argument is that measures of crystallized intelligence, as manifested in general cultural knowledge tests (like WAIS Information) or in recognition vocabulary tests, underestimate acquisition of new knowledge during adulthood. Thus, although the existing psychometric data suggesting long-term stability in verbal abilities and cultural knowledge diverges from the pattern of negative age differences seen with fluid intelligence and other human abilities, the stable plateau seen for vocabulary tests may be insensitive to the life-long learning that occurs in the specific domains in which people invest time and effort to acquire knowledge. Even within the domain of vocabulary, there may be activity-dependent differences in the types of word knowledge that are acquired. Frequent crossword puzzle players show different cross-sectional age differences in esoteric vocabulary terms they can correctly recognize (compared to non-puzzlers), probably as a direct function of actual experience with encountering these terms while solving puzzles (Hambrick, Meinz, & Salthouse, 1999). Be that as it may, there is little question that abilities that reflect specific knowledge acquisition are maintained or improved, at least into the sixties.

Beier and Ackerman's (2005) work on specificity of knowledge acquisition resonates with other evidence that people of different ages at a particular historical time point also differ in historical life contexts that have produced cohort differences in knowledge-based abilities. Schaie (2012) has studied adult intellectual development for over fifty years, using hybrid cross-sectional and longitudinal designs known as sequential strategies. This approach enables an evaluation of age changes across different birth cohorts and epochs of historical time. One of Schaie's findings is that there are large cohort differences in vocabulary, which helps to explain why studies of age and cognition that use older vocabulary tests – particularly with "advanced" and perhaps dated if not obsolete words – tend to find that older adults perform better than younger adults. Such age differences probably reflect a combination of improvement with experience in older adults and lower knowledge of esoteric word meanings in younger generations, which can affect test validity (Fox, Berry, & Freeman, 2014). By the same token, it is likely to be true that younger adults have more word knowledge in domains they commonly employ, such as technical terms and jargon associated with advanced technology (older adults are less likely to use new technology such as smartphones or tablets; Czaja et al., 2006). Schaie (2012) has also shown that there are cohort differences favoring earlier-born generations in simple mental calculations such as two-column addition. One could view this effect as being a societal consequence of the prevalent use of computers and calculators in more recently born cohorts, slowing the efficiency of mental arithmetic.

In sum, the distinction in developmental functions between knowledge and experience-based abilities, on the one hand, and fluid-like abilities, on the other hand, is consistent with a large body of cross-sectional evidence.

Longitudinal Evidence Regarding Levels of Adult Intellectual Development

As noted earlier, Schaie and colleagues (e.g., Schaie, 2012) have assembled the largest extant data base with combined longitudinal and cross-sectional intelligence test data. A reasonable question to ask, then, is whether these data produce radically different conclusions regarding age changes in adult intellectual development, relative to the cross-sectional data.

On the one hand, Schaie's (2012) data clearly indicate that cohort differences are not confined to aspects of knowledge and crystallized intelligence. He also observes substantial generational differences on tests of fluid reasoning and spatial relations. Others have noted the changes during the twentieth century in performance on tests of reasoning and fluid intelligence, as manifested in the so-called Flynn effect (Flynn, 2007; Raven, 2000). Time-lag studies of test norms in different countries suggest that more recently born cohorts are improving level of performance on tests related to fluid intelligence, including the Raven's Progressive Matrices test (Raven, 2000). The causes of these shifts are unknown and may reflect societal change in a number of variables, including the nature of schooling, nutritional practices, family size, and urbanization (see Pietschnig & Voracek, 2015). Fox and Mitchum (2013) presented item analysis of inductive reasoning tests to argue for cohort differences in cognitive strategies for representing and solving reasoning test items.

The impact of these cohort effects is primarily in attenuating the estimated changes in intelligence from ages twenty to fifty, but they also reduce the magnitude of estimated age change in late life as well (Zelinski et al., 2009). Figure 9.2 shows cumulative longitudinal curves assembled by Schaie (2012) for the five STAMAT (Schaie-Thurstone Adult Mental Abilities Test) tests that have been in his study from its inception. Note that these curves suggest that, on average, intellectual abilities start showing decline after age fifty. It is also interesting to note that the five curves are more similar than different in their altered shapes, relative to cross-sectional trends.

Certainly the STAMAT Verbal Meaning test shows a prolonged period of maintenance, relative to the other abilities, but it too manifests evidence of longitudinal decline in old age. Separate evidence, however, suggests that this pattern of apparent decline is an artifact of the speeded properties of the STAMAT Verbal Meaning test (e.g., Hertzog, 1989). In fact, all of the STAMAT tests are substantially influenced by speed of processing, in part because of limited item difficulty, even for the tests of inductive reasoning and spatial ability.

The pattern of mean ability changes based on sequential data can be summarized in three parts. The first is the similarity of age changes across different aspects of cognitive mechanics. The second is the conclusion that meaningful age-related changes in cognitive mechanics occur after midlife and accelerate in magnitude in

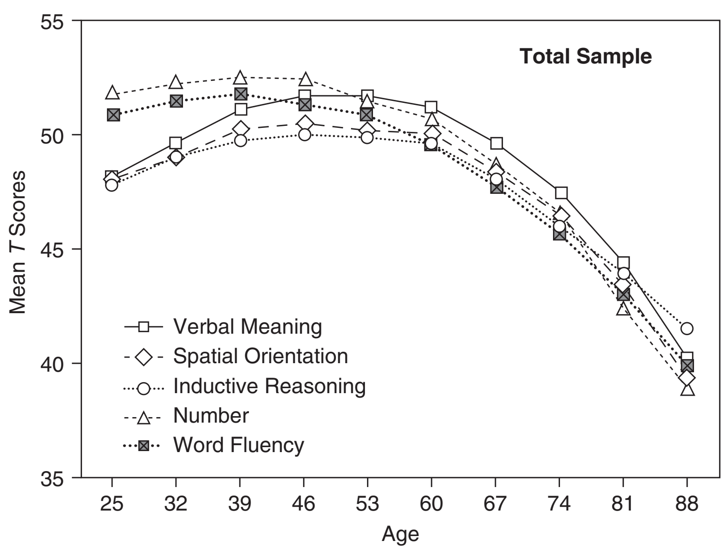

Figure 9.2 *Cumulated longitudinal gradients from seven-year longitudinal data on five primary abilities measured by Schaie (2012).*

late life. The third is the presence of substantial cohort effects on variables measuring different aspects of cognitive mechanics that inflate estimates of age changes made from cross-sectional data.

Regarding cohort effects, there is broad agreement across studies that there are few cohort effects in general information-processing speed, including the perceptual-speed factor identified by psychometric tests (e.g., Hultsch et al., 1998; Schaie, 1989). However, the limited available data from studies other than the Seattle Longitudinal Study confirm substantial cohort effects on tests of reasoning (Raven, 2000; Rönnlund & Nilsson, 2008; Zelinski & Kennison, 2007) and visuospatial ability (Rönnlund & Nilsson, 2008; Zelinski & Kennison, 2007). These effects attenuate estimated age changes in cognition. For example, Zelinski and Kennison (2007) found that six-year effect sizes in reasoning, spatial ability, and episodic memory were reduced in old age by between 0.2 and 0.3 standard deviations by controlling on cohort differences. Interestingly, some studies report few cohort effects on crystallized intelligence while finding larger effects on abilities more related to cognitive mechanics (see Zelinski et al., 2009; cf. Alwin, 2009), even though the fluid-crystallized theory would lead one to expect the opposite – i.e., larger effects on tests subject to cultural influences on knowledge and cognitive pragmatics.

The conclusion that declines in cognitive mechanics are subtle before age fifty, and accelerating thereafter, is broadly consistent with reported results from a number of other longitudinal studies of cognition and intellectual abilities in adulthood, including the Long Beach Longitudinal Study (Zelinski et al., 2009), the Victoria Longitudinal Study (Hultsch et al., 1998), and the Betula Longitudinal Study (Rönnlund et al., 2005). These studies all suggest curvilinear patterns of average age changes from the period of midlife through old age, with an acceleration in the rate of aging effects on fluid intelligence, episodic memory, and spatial visualization and other fluid-like abilities after age sixty-five.

Salthouse (2009) has argued that the type of longitudinal gradients produced by Schaie (2012) are contaminated by practice effects on the tests, an internal validity threat (Shadish, Cook, & Campell, 2002) that is problematic for longitudinal designs (Schaie, 1977). Because individuals are repeatedly given the same tests, they may show some savings in generating problem answers. If it were the case that younger adults manifest larger practice effects (an age x practice interaction), perhaps due to retention of prior test answers, then the contamination by practice would produce shallower age slopes. One way to address the problem of practice effects has been to incorporate effects of number of occasions of measurement as a proxy for exposure that would benefit from practice. Models which use this approach also tend to increase the magnitude of age-related decline and estimate an earlier onset of reliable age-related decline (e.g., Ferrer et al., 2004; Rabbitt et al., 2004).

However, this modeling approach is controversial (see the exchange between Salthouse, 2009; Schaie, 2009; and Nilsson et al., 2009). A model that uses all available data in a standard longitudinal panel and then jointly estimates age changes and practice effects (under the convergence assumption – see McArdle & Bell, 2001) confounds the estimates of practice effects with other influences that are not modeled, including historical period (time), experimental mortality (attrition), and selection x period interactions. It also assumes that age changes are not moderated by personal characteristics of people who are likely to drop out of the study. Sliwinski, Hoffman, and Hofer (2010) argued that such models inevitably assign within-person changes that deviate from cross-sectional trends to estimates of practice effects, skewing the estimated age effects away from within-person change toward (cross-sectional) between-person differences. As pointed out by Nilsson et al. (2009), studies that use independent-samples comparison groups to estimate practice effects report far less impressive practice adjustments than studies like Ferrer and colleagues (2004).

Age Changes in the Factor Structure of Intelligence Tests

Another important question about aging is whether it influences the underlying factor structure of human abilities. A leading developmental hypothesis has been the dedifferentiation hypothesis (e.g., deFrias et al., 2007). It states that shared causes of age effects across different kinds of human abilities will produce increased correlations among ability factors. In the extreme, such changes could lead to a reduced number of distinct human abilities.

Factor-analytic questions of this type cannot be separated from issues of how broadly or narrowly tests are selected. A unifying perspective on this issue derives from hierarchical models of abilities, such as that of Carroll (1993). This view suggests that one can evaluate factor structure at a relatively narrow level (how different tests define primary abilities, such as inductive reasoning or working memory), at a second-order level (how different primary abilities define higher-order factors like fluid intelligence, general speed of processing, or spatial visualization), or at the highest levels (how second-order factors define a highest-order general intelligence factor). At the primary-ability or second-order level, one can also evaluate the correlations among ability factors, treating these correlations as an index of differentiation. In addressing these questions, one can run into difficulty separating measurement invariance and suboptimal measurement properties of tests from changes in relationships among constructs. For example, use of speeded tests of intelligence may produce a substantial degree of dedifferentiation that is attributable to the global effects of speed of processing on test performance, rather than because the underlying ability constructs are becoming more correlated (Hertzog & Bleckley, 2001).

The best available evidence suggests that the factor structure of intelligence is not materially affected by aging. A large number of confirmatory factor-analytic studies, using both cross-sectional and longitudinal data, indicate that the same human abilities can be identified in young adulthood, middle age, and old age (e.g., Anstey, Hofer, & Luszcz, 2003; Brickley, Keith, & Wolfe, 1995; Hertzog et al., 2003; Hertzog & Schaie, 1986; Hultsch et al., 1998; Lane & Zelinski, 2003; Schaie et al., 1998). In all cases, the hypothesis of configural invariance (i.e., that the same variables load on the same factors at all ages; Meredith & Horn, 2001) has been supported. In most cases, the evidence supports the stronger hypothesis of metric invariance, that the unstandardized factor pattern weights, or factor loadings (Meredith & Horn, 2001), are numerically equivalent across time in longitudinal studies or across age groups. This is a broad generalization, and there are some interesting exceptions. Nevertheless, the developmental changes that occur in adulthood do not appear to radically alter the underlying nature of human abilities.

On the other hand, the evidence regarding whether adult development results in increasing correlations among human ability factors is mixed. Some studies have not found such effects (e.g., Brickley et al., 1995; Zelinski et al., 2009), whereas other studies have (deFrias et al., 2007; Hertzog & Bleckley, 2001; Hertzog et al., 2003; Hultsch et al., 1998; Schaie et al., 1998; Verhaeghen & Salthouse, 1997). However, major increases in factor correlations may be most likely to emerge past age seventy (deFrias et al., 2007; Schaie et al., 1998).

One methodological concern with age-comparative factor analysis is that aggregation over long epochs of age is often needed to generate sufficient sample sizes for factor analysis of cross-sectional data. For example, one might pool data from people within the age ranges of twenty to thirty-nine, forty to fifty-nine, and sixty to seventy-nine to create "young," "middle-aged," and "old-age" groups. Aggregation over wide age spans (such as 20 years) can create spurious increases in factor correlations because of the inflating influence of age heterogeneity within each group on correlations among ability tests (Hofer, Flaherty, & Hoffman, 2006). That is, some of the positive correlation may arise from similar cross-sectional age trends for different variables. Given the evidence for greater average age change after age sixty (change that is similar across different abilities), factor correlations in the oldest group would be inflated more than other age groups. Forming narrower age spans, if possible given the sample size, helps to avoid this effect.

In sum, factor-analytic evidence indicates subtle changes, if any, in the factor structure of human abilities. Thus, quantitative comparisons of ability test scores are unlikely to be compromised either by age-related changes in the organization of human abilities or age-related shifts in the measurement properties of the tests (Baltes & Nesselroade, 1970).

Individual Differences in Cognitive Change

One of the remarkable features of human intelligence is its relative stability of individual differences over years, even decades. When longitudinal data are collected on the same persons over time, it is possible to compute correlations of individual differences in ability test scores across that interval. These correlations – also termed stability coefficients – can be remarkably high. For example, Ian Deary and colleagues discovered large-sample data on a general ability test for cohorts of Scottish schoolchildren in multiple cohorts, and readministered the test over sixty years later to those who could be located. Test-retest correlations were approximately 0.65 across the different cohorts (e.g., Deary et al., 2004). Similar findings have been reported in long-term longitudinal studies using a wider range and variety of intelligence tests and cognitive tasks (e.g., Schaie, 2012). Moreover, when statistical corrections are possible to correct for attenuation of the stability estimates for measurement error, the correlations are even higher. Hertzog and Schaie (1986) reported that the latent seven-year stability of a general intelligence factor formed from primary ability tests was about 0.9. Hence it is reasonable to conclude that individual differences in abilities are to a substantial degree preserved as a function of aging. Those individuals who perform well in a particular domain are likely to continue to do so across their adult lives.

There is a caveat here. Longitudinal studies may overestimate the stability of individual differences. Selective attrition has been universally demonstrated in longitudinal studies of human abilities – those individuals who return for testing performed higher at the inception of the study than those who fail to return (e.g., Ghisletta, McArdle, & Lindenberger, 2006; Schaie, 2012). Selective attrition and population mortality are also likely to upwardly bias estimates of stability of individual differences in intelligence.

Nevertheless, even in positively selected samples, the less-than-perfect observed stability implies that there are at least some reliable individual differences in rates of change. When growth curve analyses or latent difference score analyses are performed on longitudinal cognitive data, reliable variances in the slopes of the growth curves are generally found (e.g., deFrias et al., 2007; Ghisletta et al., 2006; McArdle et al., 2002). Not all individuals are changing at the same rate – some decline faster than others, and some even show improvements. Schaie (2012) has argued that, although the modal pattern of individual change is one of relative stability in midlife, one can identify also individuals who reliably decline or who reliably improve, even on abilities related to cognitive mechanics. Data on six-year stability from the Victoria Longitudinal Study (VLS) showed reliable variances in latent difference scores (Hertzog et al., 2003) on multiple cognitive variables, including working memory, episodic memory, fluid intelligence, ideational fluency, verbal comprehension, and speed of processing. The variance in change was found despite corrected stability coefficients that were typically in the 0.8 to 0.9 range. As pointed out by deFrias and colleagues (2007), these individual differences in cognitive changes may also be more pronounced in old age than in middle age.

The existence of individual differences in change with regard to different human abilities raises an intriguing question. Are these changes related to each other? Rabbitt (1993) once framed the question as: Does it all go together when it goes? There is good evidence that changes across variables are positively correlated. Given the extended measurement batteries in studies like the Betula Longitudinal Study and the VLS, we probably know the most about associations in age-related changes in different aspects of memory. In the case of the VLS, analyses in two different six-year longitudinal samples show that individual differences in changes in working memory are correlated with changes in episodic memory (measured by free recall of word lists and narrative text content) and in a measure of semantic memory (fact recall). In addition, changes in working memory also correlate with changes in other abilities, including ideational fluency, inductive reasoning, and speed of processing (Hertzog et al., 2003; Hultsch et al., 1998). Betula study data indicates correlations among different aspects of episodic memory and processing speed (Lövdén et al., 2004). Hertzog and colleagues (2003) showed that one could fit a higher-order general factor of change to the latent change factors for multiple cognitive abilities. This latent variable was defined principally by working memory, but also had substantial loadings on most other variables, with the exception of changes in vocabulary.

One interesting feature of the VLS data was the strong association of changes in fact recall with changes in working memory. The fact-recall measure assessed cultural knowledge (e.g., "Who is the cartoon character who gets his strength from eating spinach?"). Cross-sectionally, the fact-recall measure behaves like a measure of crystallized intelligence, as one would expect (Hultsch et al., 1998). Longitudinally, changes in fact recall dissociate from verbal comprehension. Instead, longitudinal fact-recall changes correlated much more highly with changes in working memory and episodic memory than with changes in vocabulary scores. Such a pattern suggests individual differences in late-life changes in memory retrieval mechanisms required by both episodic and semantic memory tasks.

Typically, measures of inductive reasoning and working memory correlate strongly in cross-sectional data and in studies testing only young adults (e.g., Kane & Engle, 2002; Salthouse et al., 2008). Kyllonen and Chrystal (1990) once remarked that reasoning might not be, in fact, differentiable from working memory. Yet working-memory changes and reasoning changes are only moderately correlated in the VLS (Hertzog et al., 2003). Instead, changes in working memory are more highly correlated with changes in fact recall than with changes in reasoning. Collectively, these findings indicate that the ability components that drive age changes in these variables may not be the same influences that determine the factor structure of abilities in young adulthood.

Perhaps the most interesting aspect of the VLS change factors was that there was reliable change variance in almost all human abilities that was unique to each variable. Hertzog and colleagues (2003) modeled a higher-order factor of general cognitive change to account for the correlations of change among the different cognitive variables. This general change factor had moderate to strong relationships to changes in most of the cognitive variables. Thus, there is a coherence to the individual differences in rates of cognitive change in later life.

Tucker-Drob (2011) reported similar findings from the Virginia Cognitive Aging Study, using the Woodcock-Johnson – Revised tests of intellectual abilities. Figure 9.3 shows his results at the level of latent changes (ignoring intercept or start-point variance). The bottom part of the figure shows the hierarchical model for slopes (or individual differences in change), not intercepts (starting points). The pie chart at the top of the figure partitions the slope variance into general (shared) change across all variables, domain-specific shared change (e.g., among the three indicators for G*f*), and change specific to each ability test. Although 39 percent of the variance in intellectual change was shared in common (i.e., a common factor of individual differences in rates of change), fully 33 percent was specific to primary abilities, controlling on the general change factor. This pattern establishes that the similarity in cross-sectional gradients belies the underlying heterogeneity in change found in longitudinal data. We would not expect to see robust individual differences in change

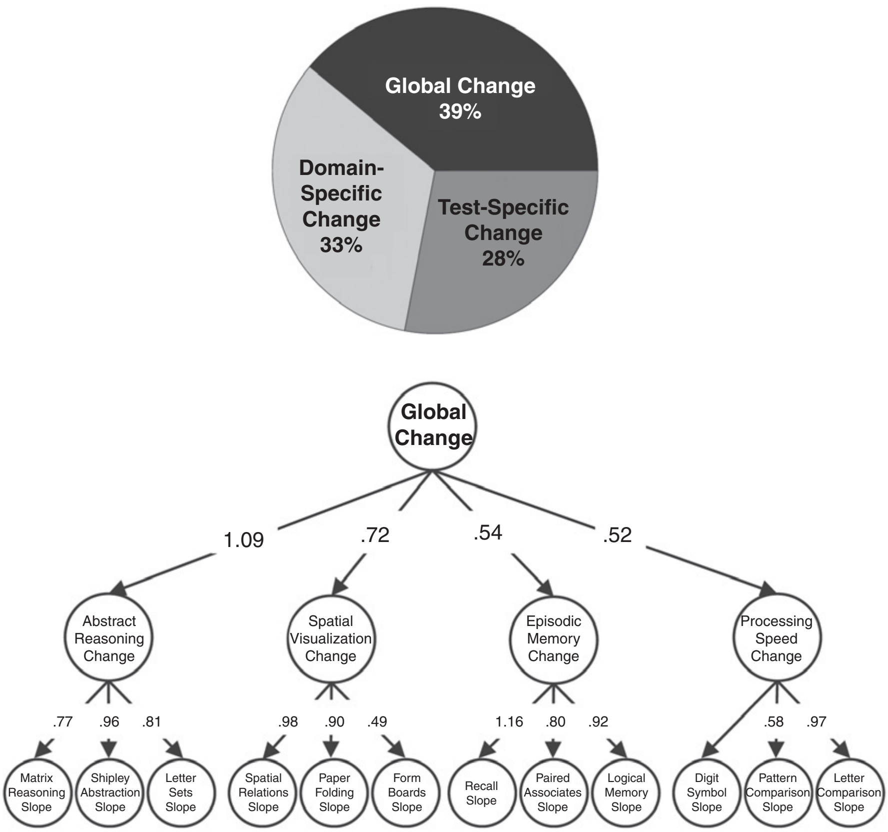

Figure 9.3 *Graphical representation of overlap of individual differences in cognitive change in adulthood, based on the standardized factors of slopes model from Tucker-Drob (2011).*

The top panel depicts a pie chart indicating proportion of change variance accounted for, on average, by different strata of latent variables of change in the estimated model: global change, domain-specific change (abstract reasoning change, spatial visualization change, episodic memory change, and processing speed change), and test-specific (e.g., matrix reasoning) change slopes. Results indicate measurable individual differences in rates of change at all three strata.

at the primary ability level, controlling on the general factor of change, if all intellectual abilities declined in tandem late in life (Rabbitt, 1993).

Individual differences in cognitive changes therefore diverge from the similarity of average age trends in fluid intelligence and other aspects of cognitive mechanics. The coherence of cognitive change – as manifested in moderate correlations of longitudinal changes across variables – obscures the fact that variables are changing independently, such that people will have different profiles of change across a set of cognitive variables. Although it is dangerous to draw conclusions about the number of underlying causes of change based on such patterns, it seems that there may be multiple, differentiable mechanisms influencing rates of age-related changes in cognitive abilities.

This conclusion can be contrasted to Salthouse and colleagues (2008), who implied that there may be relatively few underlying influences on cognitive change in adulthood. Why the discrepancy? The most obvious possibility is a disconnect between cross-sectional and longitudinal outcomes, with Salthouse relying almost exclusively on cross-sectional data as the basis for his inference. Certainly, there are potential issues with the validity of the longitudinal estimates of correlated change. For instance, Ferrer and colleagues (2005) noted that differential practice effects across variables could distort the estimated longitudinal change correlations. It is difficult to believe, however, that such effects could produce artifactual variable-specific change variance of the type reviewed here. For instance, the VLS uses rotating alternate forms to measure word recall, text recall, and fact recall with different items at each occasion of measurement, so any practice effects would be about learning how to memorize information, not a saving in memory for specific items tested earlier in the longitudinal study.

To my mind, the difference arises essentially because the question cannot be adequately addressed by statistical models of cross-sectional data (Hofer et al., 2006; Lindenberger et al., 2011). Cross-sectional analyses can only estimate, in effect, correlations among cross-sectional age curves by testing for whether cognitive variables have a partial correlation with age, controlling on other cognitive variables. This approach can reveal whether average age trends differ between variables (e.g., Horn, Donaldson, & Engstrom, 1981). Failing to detect different shapes of cross-sectional curves neither implies that the variables in question change in lockstep nor that their changes have the same underlying causes. To actually assess individual differences in change, one must repeatedly measure the same people (Baltes & Nesselroade, 1979).

In sum, there is a high degree of stability in human abilities across the adult life course, but at the same time there are individual differences in cognitive changes, particularly in old age. A critical question, then, is what determines these individual differences in cognitive trajectories.

Influences on Adult Cognitive Development

The individual differences in cognitive change just reviewed could in principle reflect a number of different influences. Cognitive psychologists tend to focus on processing mechanisms that are associated with changes in complex cognition. Information-processing resources like working memory, processing speed, and inhibitory aspects of attention are often cited as causes of age changes in intelligence (see Hertzog, 2008; Salthouse, 1996; Verhaeghen & Salthouse, 1997). Even so, the question remains as to what determines age-related changes in fundamental processing mechanisms that serve as resources for more complex human abilities.

One important influence is individual differences in genetically programmed biological aging – often termed senescence. In essence, the idea is that our biological aging clocks may be ticking in different metrics of time. Newer research derived from insights into the human genomic code indicates that genetic polymorphisms associated with neurotransmitters, neurotrophins, and related hormones influence adult cognitive development (e.g., Harris et al., 2006; Lindenberger et al., 2008). Behavioral-genetic studies indicate a considerable degree of heritability in cognitive change in late life (Reynolds, 2008). However, genetic predispositions interact with social and psychological mechanisms to produce cognitive phenotypes, so one should be circumspect about attributing observed age-related changes to genetic mechanisms.

Indeed, when we organize our data by chronological age, we cannot assume that we are measuring individual differences in rates of biological aging. The effects of age merely describe variation in cognition that is systematically correlated with how old people are. But there are many contextual variables that are also correlated with chronological age, including age-graded events like retirement, experience, and shrinkage of one's social network. Furthermore, nonnormative, negative life events are correlated with age, such as risks for contracting different kinds of chronic disease that can impact cognition, either directly through influences on the brain or indirectly through psychological effects of medications used to treat them (Birren, 1964). The longitudinal studies that generate the data in question may measure physical health, but typically cannot control for disease by only assessing disease-free older adults. There are very few disease-free individuals in an older population. The average older adult has three or more chronic health conditions, including conditions that affect cognitive performance: arthritis, vascular disease, Type II diabetes, reduced hormonal secretion, pulmonary or renal disease, and declining sensory and perceptual function (e.g., macular degeneration; see Spiro & Brady, 2008). There are also a host of brain pathologies that are correlated with age, and which may have impact on cognition before they are clinically detected, including different forms of dementia and Parkinson's disease. Lifestyles also change as people grow older, sometimes as a consequence of limitations produced by chronic disease, in other cases as a function of changing patterns of behavior that have psychological and social origins.

Certainly, structural features in the brain undergo changes that are correlated with cognition. For instance, Raz and colleagues (2008) analyzed a longitudinal sample that had been measured with structural magnetic resonance imaging to evaluate changes in gray-matter volume in the cerebral cortex. Individual differences in the structural changes in dorsolateral prefrontal cortex and hippocampal areas of the brain were correlated with changes in fluid intelligence.

Disease and brain pathology. Raz and colleagues' (2008) findings do not necessarily imply that neurobiological aging in the brain drives cognitive changes. The morphological changes in the brain can also be caused by disease, such as cardiovascular disease and dementia. Sliwinski and colleagues (2003) conducted a fascinating study in this regard, using data from the Bronx Longitudinal Study (Sliwinski & Buschke, 2004). The study involved a prospective design of the incidence of dementing illnesses in a non-demented control group collected as part of a larger study of Alzheimer's disease and related disorders. Individuals in this group were measured cognitively at regular intervals, but also assessed for dementia. Over time, some participants in the longitudinal study were clinically diagnosed as having dementia, and this allowed Sliwinski and colleagues to compare cognitive change in the preclinical phase with change in those individuals who did not convert to dementia. As might be expected, individuals who had not yet been diagnosed with dementia (but undoubtedly had contracted the disease) showed greater change in episodic memory during their preclinical phase, compared to individuals who did not later receive a dementia diagnosis. Even more interesting, however, was the fact that the aggregate control sample manifested individual differences in cognitive change, as well as correlations of changes across cognitive variables. However, the magnitude of these individual differences was reduced by controlling for later dementia diagnosis, as were the correlations of change among different variables. Furthermore, within the dementia group, organizing the data according to time of diagnosis rather than chronological age eliminated the individual differences in rates of cognitive change.

What does this pattern imply? It would appear that, in this sample, the presence of preclinical dementia was a major source of individual differences in cognitive change. Because people vary in the age at which the disease is contracted and later diagnosed, organizing the data by age (ignoring the information about the disease and its progression) produced larger individual differences in rates of change. Given that other prospective studies of Alzheimer's disease, vascular dementia, and other dementing illnesses indicate a fairly long preclinical period in which cognition may be affected (e.g., Bäckman & Small, 2007), it would appear that a major influence on individual differences in cognitive change in old age is the presence or absence of (possibly undetected) dementia. Furthermore, a number of studies have directly linked magnitudes of longitudinal changes in cognitive abilities to different kinds of disease, including cardiovascular and cerebrovascular disease, late-onset diabetes, and their precursors, or risk factors, such as obesity, hypertension, poor cholesterol profiles, and the like (Spiro & Brady, 2008).

Disease and terminal decline. A focus on disease effects on cognition raises an additional set of important questions about aging and intellectual development. To what extent are the average curves for cognitive abilities and age misleading, in the sense that they are not representative of the actual developmental trajectories of individuals? Aggregated statistics like means, even if generated from longitudinal data, are simply best guesses as to the level of function, on average, at a particular age. We connect the means of different ages with a line (or statistically fit a curve to the data), but this does not imply that the developmental pathway of individuals in the sample manifests the pattern of change implied by the shape of the aggregate mean curve.

Indeed, it has long been surmised that the population of adults might be quite heterogeneous in nature, with the major changes in psychological functioning, including cognition, occurring during a so-called period of terminal decline; i.e., a period of decline preceding death (e.g., Berg, 1996; Bosworth, Schaie, & Willis, 1999). Indeed, time to death may be a more important way of indexing cognitive loss in old age than chronological age (Singer et al., 2003). Some impressive data on this score comes from models of longitudinal data that jointly use time to death and age to organize the data (Ram et al., 2010). The modeling approach is fairly complex, requiring estimation of a change point (Hall et al., 2000), at which the slope of decline prior to the change point is lower than the slope immediately prior to death. Thorvaldsson et al. (2008) used this method to demonstrate accelerated cognitive decline occurring about seven years before death in the Swedish Goteborg Longitudinal Study data. Wilson and colleagues (2007) found evidence for a shorter period of terminal decline of about four years. Terminal decline was associated with the apolipoprotein E ε4 allele, a genetic polymorphism thought to be associated with risk for Alzheimer's disease (AD). Laukka, MacDonald, and Backman (2008) also concluded that a substantial proportion of the variance in terminal cognitive decline might be due to emergence of dementia, but there was evidence of decline in individuals who did not develop AD. Undoubtedly, future research will clarify the extent to which other disease factors play a role in terminal cognitive decline, including vascular disease and organ failure (e.g., renal dysfunction; see Buchman et al., 2009).

In light of the evidence for terminal decline effects, the possibility exists that the curvilinear age trends for cognitive function in late life are actually an artifact of aggregation over individuals with different functions. This idea was nicely illustrated by Baltes and Labouvie (1973), who showed that a combination of (1) a change-point function of stable level of cognition, followed by terminal decline, and (2) a variable onset of the terminal decline that was correlated with advancing age, could produce aggregate curvilinear functions that did not capture the functional form of individual change (see Figure 9.4). The aggregate function could be influenced by the increasing risk of terminal decline, with its curvature reflecting an averaging of persons in terminal decline with persons who are still stable.

Exercise and an engaged life style. A critical question regarding adult intellectual development is whether health-promoting behaviors such as exercise, nutrition, and an active lifestyle promote better developmental outcomes (Hertzog et al., 2008).

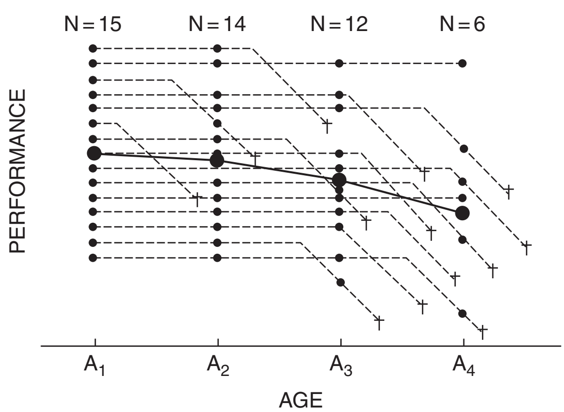

Figure 9.4 *Demonstration of how aggregating over persons conforming to a pattern of stability, followed by terminal decline, would produce a mean curvilinear change given (1) an age-related increase in the risk of terminal decline and (2) mortality-related attrition from the sample (from Baltes & Labouvie, 1973).*

Over the past two decades, compelling evidence has emerged that aerobic exercise in middle age and old age promotes enhanced cognitive function in older adults. Colcombe and Kramer's (2003) highly cited meta-analysis evaluated aerobic exercise intervention studies in older adults, compared the exercise groups' cognitive performance to performance in a group doing toning and stretching only (see also Kramer & Colcombe, 2018). Short-term aerobic exercise interventions generate improvements in tasks assessing executive functioning and controlled attention (domains highly correlated with fluid intelligence; Salthouse et al., 2008). The data are broadly consistent with cross-sectional studies suggesting an association of self-reported exercise with cognitive abilities (e.g., Eggermont et al., 2009), but the intervention effects help to argue for a causal influence of exercise on cognition. Unfortunately, there are at present no longitudinal studies that contrast longer-term adherence with exercise regimens and degree of cognitive change in adulthood.

Does engaging in intellectually stimulating activities also promote better cognitive outcomes? Salthouse (2006) expressed skepticism on this score, given that his cross-sectional data on self-reported activities have failed to observe interactions of age and activity (see Hertzog et al., 2008, for a critique of this argument). Certainly, simple cross-sectional correlations of activities and intelligence are insufficient grounds for arguing that activities help preserve cognitive functioning, because individuals with high intelligence tend to manifest higher levels of intellectual engagement in early adulthood (Ackerman & Heggestad, 1997). However, longitudinal evidence is needed, given the potential lack of sensitivity of cross-sectional data to change alluded to earlier. Longitudinal studies have often found relationships of self-reported intellectual engagement with cognition (e.g., Schooler, Mulatu, & Oates, 1999; Wilson et al., 2003; see Hertzog et al., 2008 for a review). As noted by Hultsch and colleagues (1999), longitudinal correlations of activities with cognitive change do not have an unambiguous causal interpretation, because the correlations could also reflect late-life cognitive changes leading to curtailed activity (MacKinnon et al., 2003).

There are fewer intervention studies with activities, but there is at least some indication that encouraging older adults to engage in stimulating activities may have cognitive benefits (Carlson et al., 2008; Stine-Morrow et al., 2007; Tranter & Koutstaal, 2008). For instance, some studies indicate that participation in a complex video game environment led to short-term improvements in attentional control and executive function (e.g., Basak et al., 2008). This outcome is consistent with intervention studies that target executive control (Hertzog et al., 2008), producing more transfer of training than is typically observed when training focuses on teaching specific processing strategies (e.g., Ball et al., 2002). The evidence favors an impact of activities on cognitive function, but there is still some disagreement and controversy on this point.

Functional Aspects of Adult Intelligence

Given that there are, on average, adult age changes in cognitive abilities, what are the practical consequences of these changes? Evidence is beginning to emerge that there are fewer practical implications for cognitive functioning in everyday life than some might have supposed.

For example, older workers, even those with intellectually demanding jobs, function well on the job even into old age (e.g., Ng & Feldman, 2008). Work by Colonia-Willner (1998) may suggest a reason for this maintenance; experience on the job (which correlates with age) brings with it increases in tacit knowledge (Cianciolo et al., 2006) about how to perform effectively on the job. Colonia-Willner studied bankers of different ages in Brazil. Although her cross-sectional sample showed typical age differences in fluid intelligence, expert ratings of tacit knowledge about hypothetical banking situations indicated age-related improvements in this domain.

Such effects can be observed in intellectually demanding game situations as well. Masunaga and Horn (2001) studied the relationship of fluid intelligence to performance on the Japanese game of Go, a cognitively demanding task with some resemblance to chess. Go performance was less correlated with standard measures of fluid intelligence and working memory than with measures of reasoning that directly represented reasoning about Go moves. In a similar vein, Charness and colleagues have demonstrated good memory retention for chess positions by older chess experts, relative to their impaired episodic memory for chess pieces placed in random positions on the chessboard (e.g., Charness, 1981). Hershey, Jacobs-Lawson, and Walsh (2003) reported sound simulated financial decision-making by older adults who had prior experience in investing or gained it through structured task experience. Performance in familiar environmental contexts is associated with beneficial effects of pragmatic knowledge about typical scripts and scenarios, common decisions and choice points, and intact access to effective strategies for performance that help older adults preserve effective cognitive functioning, even in the face of decline in fluid ability (Hertzog, 2008).

Older adults may also be effective at using strategies that enhance cognition in everyday life, such as through the use of external aids or behavioral routines that support timely remembering of what to do and when to do it. For instance, older adults are sometimes better at remembering to take medications than middle-aged and younger adults, despite age deficits in standard tests of reasoning and episodic memory (Park et al., 1999). In general, older adults do well in everyday prospective memory tasks relative to laboratory tasks (Phillips, Henry, & Martin, 2008), probably because of a more active use of strategies to promote remembering.

Conclusions

The study of adult cognitive and intellectual development is entering a vibrant new phase, one in which the advances in statistical methods for modeling individual differences are being integrated with designs and measures that permit a subtle understanding of individual differences in cognitive change. The next decades are likely to see an expanded understanding of how social and psychological forces interact with biological and genetic influences to shape individual trajectories of adult cognitive development, both at the level of brain structure and behavior.

References

Ackerman, P. L. (2000). Domain-specific knowledge as the "dark matter" of adult intelligence: Gf/Gc personality and interest correlates. *Journal of Gerontology: Psychological Sciences*, *55*, P69–P84.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.

Alwin, D. F. (2009). History, cohorts, and patterns of cognitive aging. In H. B. Bosworth & C. Hertzog (Eds.), *Aging and cognition: Research methodologies and empirical advances* (pp. 9–38). Washington: American Psychological Association.

Anstey, K. J., Hofer, S. M., & Luszcz, M. A. (2003). Cross-sectional and longitudinal patterns of dedifferentiation in late-life cognitive and sensory function: The effects of age, ability, attrition, and occasion of measurement. *Journal of Experimental Psychology: General*, *132*, 470–487.

Bäckman, L., & Small, B. J. (2007). Cognitive deficits in preclinical Alzheimer's disease and vascular dementia: Patterns of findings from the Kungsholmen project. *Physiology and Behavior*, *92*, 80–86.

Ball, K., Berch, D. B., Helmer, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., et al. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *Journal of the American Medical Association*, *288*, 2271–2281.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as a foundation for developmental theory. *American Psychologist*, *52*, 366–380.

Baltes, P. B., & Labouvie, G. V. (1973). Adult development of intellectual performance: Description, explanation, and modification. In C. Eisdorfer & M. P. Lawton (Eds.), *The psychology of adult development and aging* (pp. 157–219). Washington: American Psychological Association.

Baltes, P. B., & Nesselroade, J. R. (1970). Multivariate longitudinal and cross-sectional sequences for analyzing ontogenetic and generational change: A methodological note. *Developmental Psychology*, *2*, 163–168.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development*. New York: Academic Press.

Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology*, *50*, 471–507.

Basak, C., Boot, W. R., Voss, M. W., & Kramer, A. F. (2008). Can training in a real-time strategy videogame attenuate cognitive decline in older adults? *Psychology and Aging*, *23*, 765–777.

Beier, M., & Ackerman, P. L. (2005). Age, ability, and the role of prior knowledge on the acquisition of new domain knowledge: Promising results in a real-world learning environment. *Psychology and Aging*, *20*, 341–355.

Berg, S. (1996). Aging, behavior, and terminal decline. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (4th ed., pp. 323–337). San Diego, CA: Academic Press.

Birren, J. E. (1964). *The psychology of aging*. Englewood Cliffs, NJ: Prentice-Hall.

Bosworth, H. B., Schaie, K. W., & Willis, S. L. (1999). Cognitive and sociodemographic risk factors for mortality in the Seattle Longitudinal Study. *Journal of Gerontology: Psychological Sciences*, *54*, P273–P282.

Botwinick, J. (1977). Intellectual abilities. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 580–605). New York: Van Nostrand Reinhold.

Brickley, P. G., Keith, T. Z., & Wolfle, L. M. (1995). The three-stratum theory of cognitive abilities: Test of the structure of intellect across the adult life span. *Intelligence*, *20*, 309–328.

Buchman, A. S., Tanne, D., Boyle, P. A., Shah, R. C., Leurgans, S. E., & Bennett, D. A. (2009). Kidney function is associated with the rate of cognitive decline in the elderly. *Neurology*, *73*, 920–927.

Carlson, M. C., Saczynski, J. S., Rebok, G. W., McGill, S., Tielsch, J., Glass, T. A., et al. (2008). Exploring the effects of an everyday activity program on executive function and memory in older adults: Experience Corps. *The Gerontologist*, *48*, 793–801.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge, UK: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Charness, N. (1981). Aging and skilled problem solving. *Journal of Experimental Psychology: General*, *110*, 21–38.

Cianciolo, A. T., Matthew, C., Sternberg, R. J., & Wagner, R. K. (2006). Tacit knowledge, practical intelligence, and expertise. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 613–632). New York: Cambridge University Press.

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science*, *14*, 125–130.

Colonia-Willner, R. (1998). Practical intelligence at work: Relationships between aging and cognitive efficiency among managers in a bank environment. *Psychology and Aging*, *13*, 45–57.

Czaja, S., Charness, N., Fisk, A. D., Hertzog, C., Nair, S., Rogers, W. A., et al. (2006). Factors predicting the use of technology: Findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging*, *21*, 333–352.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology*, *86*, 130–147.

deFrias, C. M., Lövdén, M., Lindenberger, U., & Nilsson, L-G. (2007). Revisiting the de-differentiation hypothesis with longitudinal multi-cohort data. *Intelligence*, *35*, 381–392.

Eggermont, L. H. P., Milberg, W. P., Lipsitz, L. A., Scherder, E. J. A., & Leveille, S. G. (2009). Physical activity and executive function in aging: The MOBILIZE Boston study. *Journal of the American Geriatric Society*, *57*, 1750–1756.

Ferrer, E., Salthouse, T. A., McArdle, J. J., Stewart, W. F., & Schwartz, B. S. (2005). Multivariate modeling of age and retest in longitudinal studies of cognitive abilities. *Psychology and Aging*, *20*, 412–422.

Ferrer, E., Salthouse, T. A., Stewart, W. F., & Schwartz, B. S. (2004). Modeling age and retest processes in longitudinal studies of cognitive abilities. *Psychology and Aging*, *19*, 243–249.

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge, UK: Cambridge University Press.

Fox, M. C., Berry, J. M., & Freeman, S. P. (2014). Are vocabulary tests measurement invariant between age groups? An item response analysis of three popular tests. *Psychology and Aging*, *29*, 925–938.

Fox, M. C., & Mitchum, A. L. (2013). A knowledge-based theory of rising scores on "culture free" tests. *Journal of Experimental Psychology: General*, *142*, 979–1000.

Ghisletta, P., McArdle, J. J., & Lindenberger, U. (2006). Longitudinal cognition-survival relations in old and very old age: 13-year data from the Berlin Aging Study. *European Psychologist*, *11*, 204–223.

Hall, C. B., Lipton, R. B., Sliwinski, M., & Stewart, W. F. (2000). A change point model for estimating the onset of cognitive decline in preclinical Alzheimer's disease. *Statistics in Medicine*, *19*, 1555–1566.

Hambrick, D. Z., Meinz, E. J., & Salthouse, T. A. (1999). Predictors of crossword puzzle proficiency and moderators of age-cognition relations. *Journal of Experimental Psychology: General*, *128*, 131–164.

Harris, S. E., Fox, H., Wright, A. F., Hayward, C., Starr, J. M., Whalley, L. J., Deary, I. J. (2006). The brain-derived neurotrophic factor Val66Met polymorphism is associated with age-related change in reasoning skills. *Molecular Psychiatry*, *11*, 505–513.

Hershey, D. A., Jacobs-Lawson, J. M., & Walsh, D. A. (2003). Influences of age and training on script development. *Aging, Neuropsychology, and Cognition*, *10*, 1–19.

Hertzog, C. (1989). The influence of cognitive slowing on age differences in intelligence. *Developmental Psychology*, *25*, 636–651.

Hertzog, C. (2008). Theoretical approaches to the study of cognitive aging: An individual-differences perspective. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging: Interdisciplinary perspectives* (pp. 34–49). Thousand Oaks, CA: Sage.

Hertzog, C. (2009). Use it or lose it: An old hypothesis, new evidence, and an ongoing controversy. In H. Bosworth & C. Hertzog (Eds.), *Cognition and aging: Research methodologies and empirical advances* (pp. 161–179). Washington: American Psychological Association.

Hertzog, C., & Bleckley, M. K. (2001). Age differences in the structure of intelligence: Influences of information processing speed. *Intelligence*, *29*, 191–217.

Hertzog, C., Dixon, R. A., Hultsch, D. F., & MacDonald, S. W. S. (2003). Latent change models of adult cognition: Are changes in processing speed and working memory associated with changes in episodic memory? *Psychology and Aging*, *18*, 755–769.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2008). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest*, *9*, 1–65.

Hertzog, C., & Schaie, K. W. (1986). Stability and change in adult intelligence: 1. Analysis of longitudinal covariance structures. *Psychology and Aging*, *1*, 159–171.

Hofer, S. M., Flaherty, B. P., & Hoffman, L. (2006). Cross-sectional analysis of time-dependent data: Mean-induced association in age-heterogeneous samples and an alternative method based on sequential narrow age-cohort samples. *Multivariate Behavioral Research*, *41*, 165–187.

Horn, J. L. (1985). Remodeling old models of intelligence: Gf–Gc theory. In B. B. Wolman (Ed.), *Handbook of intelligence* (pp. 267–300). New York: Wiley.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 107–129.

Horn, J. L., Donaldson, G., & Engstrom, R. (1981). Apprehension, memory, and fluid intelligence decline in adulthood. *Research on Aging*, *3*, 33–84.

Horn, J. L., & Hofer, S. M. (1992). Major abilities and development in the adult period. In R. J. Sternberg & C. A. Berg (Eds.), *Intellectual development* (pp. 44–99). New York: Cambridge University Press.

Hultsch, D. F., Hertzog, C., Dixon, R. A., & Small, B. J. (1998). *Memory change in the aged*. New York: Cambridge University Press.

Hultsch, D. F., Small, B. J., Hertzog, C., & Dixon, R. A. (1999). Use it or lose it: Engaged lifestyle as a buffer of cognitive decline in aging. *Psychology and Aging*, *14*, 245–263.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. *Psychonomic Bulletin and Review*, *9*, 637–671.

Kramer, A. F., & Colcombe, S. (2018). Fitness effects on the cognitive function of older adults: A meta-analytic study – revisited. *Perspectives on Psychological Science*, *13*, 213–217.

Kyllonen, P. C., & Chrystal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, *14*, 389–433.

Lane, C. J., & Zelinski, E. M. (2003). Longitudinal hierarchical linear models of the Memory Functioning Questionnaire. *Psychology and Aging*, *18*, 38–53.

Laukka, E. J., MacDonald, S. M. S., & and Bäckman, L. (2008). Terminal-decline effects for select cognitive tasks after controlling for preclinical dementia. *American Journal of Geriatric Psychiatry*, 16, 355–365.

Lindenberger, U., Nagel, I. E., Chicherio, C., Li, S-C., Heekeren, H. R., & Bäckman, L. (2008). Age-related decline in brain resources modulates genetic effects on cognitive functioning. *Frontiers in Neuroscience*, *2*, 234–244.

Lindenberger, U., von Oertzen, T., Ghisletta, P., & Hertzog, C. (2011). Cross-sectional age variance extraction: What's change got to do with it? *Psychology and Aging*, 26, 34–47.

Lövdén, M., Rönnlund, M., Wahlin, A., Bäckman, L., Nyberg, L., & Goran-Nilsson, L. (2004). The extent of stability and change in episodic and semantic memory in old age: Demographic predictors of stability and change. *Journal of Gerontology*: *Psychological Sciences*, *59B*, P130–P134.

MacKinnon, A., Christensen, H., Hofer, S. M., Korten, A. E., & Jorm, A. F. (2003). Use it and still lose it? The association between activity and cognitive performance established using latent growth techniques in a community sample. *Aging Neuropsychology and Cognition*, *10*, 215–222.

Masunaga, H., & Horn, J. L. (2001). Expertise and age-related changes in components of intelligence. *Psychology and Aging*, *16*, 293–311.

McArdle, J. J., & Bell, R. Q. (2001). An introduction to latent growth models for developmental data analysis. In T. D. Little & K. U. Schabel (Eds.), Modeling longitudinal and multi-level data: Practical issues, applied approaches, and specific examples (pp. 69–81). Mahwah, NJ: Lawrence Erlbaum Associates.

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology*, *38*, 115–142.

Meredith, W., & Horn, J. L. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 203–240). Washington: American Psychological Association.

Ng, T. W. H., & Feldman, D. C. (2008). The relationship of age to ten dimensions of job performance. *Journal of Applied Psychology*, *93*, 392–423.

Nilsson, L-G., Sternäng, O., Rönnlund, M., & Nyberg, L. (2009). Challenging the notion of an early onset of cognitive decline. *Neurobiology of Aging*, *30*, 521–524.

Park, D. C., Hertzog, C., Leventhal, H., Morrell, R. W., Leventhal, E., Birchmore, D., et al. (1999). Medication adherence in rheumatoid arthritis patients: Older is wiser. *Journal of the American Geriatrics Society*, *47*, 172–183.

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., &Gaines, C. L. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging*, *11*, 621–637.

Phillips, L. H., Henry, J. D., & Martin, M. (2008). Adult aging and prospective memory: The importance of ecological validity. In M. Kliegel, M. A. McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, neuroscience, developmental, and applied perspectives* (pp. 161–185). New York: Taylor & Francis.

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science*, *10*, 282–315.

Rabbitt, P. (1993). Does it all go together when it goes? The nineteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, *46A*, 385–434.

Rabbitt, P., Diggle, P., Holland, F., & McInnes, L. (2004). Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. *Journal of Gerontology: Psychological Sciences and Social Sciences*, *59B*, P84–P97.

Ram, N., Gerstorf, D., Fauth, E., Zarit, S., & Malmberg, B. (2010). Aging, disablement, and dying: Using time-as-process and time-as-resources metrics to chart late-life change. *Research on Human Development*, *7*, 27–44.

Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*, 1–48.

Raz, N., Lindenberger, U., Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Acker, J. M. (2008). Neuroanatomical correlates of fluid intelligence in healthy adults and persons with vascular risk factors. *Cerebral Cortex*, *18*, 718–726.

Reynolds, C. A. (2008). Genetic and environmental influences on cognitive change. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging: Interdisciplinary perspectives* (pp. 557–574). Thousand Oaks, CA: Sage.

Rönnlund, M., & Nilsson, L-G. (2008). The magnitude, generality, and determinants of Flynn effects on forms of declarative memory and visuospatial ability: Time-sequential analyses of data from a Swedish cohort study. *Intelligence*, *36*, 192–209.

Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L-G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Data from a population-based study. *Psychology and Aging*, *20*, 3–18.

Salthouse, T. A. (1982). *Adult cognition: An experimental psychology of human aging*. New York: Springer-Verlag.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Salthouse, T. A. (2006). Mental exercise and mental aging: Evaluating the validity of the "use it or lose it" hypothesis. *Perspectives on Psychological Science*, *1*, 68–87.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*, 507–514.

Salthouse, T. A., Pink, J. E., & Tucker-Drob, E. M. (2008). Contextual analysis of fluid intelligence. *Intelligence*, *36*, 464–486.

Schaie, K. W. (1977). Quasi-experimental designs in the psychology of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 39–58). New York: Van Nostrand Reinhold.

Schaie, K. W. (1989). Perceptual speed in adulthood: Cross-sectional and longitudinal studies. *Psychology and Aging*, *4*, 443–453.

Schaie, K. W. (2009). “When does age-related cognitive decline begin?”: Salthouse again reifies the “cross-sectional fallacy.” *Neurobiology of Aging*, *30*, 528–529.

Schaie, K. W. (2012). *Developmental influences on adult intelligence: The Seattle Longitudinal Study* (2nd ed.). New York: Oxford University Press.

Schaie, K. W., Maitland, S. B., Willis, S. L, & Intrieri, R. C. (1998). Longitudinal invariance of adult psychometric ability factor structures across 7 years. *Psychology and Aging*, *13*, 8–20.

Schooler, C., Mulatu, M. S., & Oates, G. (1999). The continuing effects of substantively complex work on the intellectual functioning of older workers. *Psychology and Aging*, *14*, 483–506.

Shadish, W., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., & Baltes, P. B. (2003). The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychology and Aging*, *18*, 318–331.

Sliwinski, M., & Buschke, H. (2004). Modeling intraindividual cognitive change in aging adults: Results from the Einstein Aging Studies. *Aging, Neuropsychology and Cognition*, 11, 196–211.

Sliwinski, M. J., Hofer, S. M., Hall, C., Bushke, H., & Lipton, R. B. (2003). Modeling memory decline in older adults: The importance of preclinical dementia, attrition and chronological age. *Psychology and Aging*, *18*, 658–671.

Sliwinski, M. J., Hoffman, L., & Hofer, S. M. (2010). Evaluating convergence of within-person change and between-person differences in age-heterogeneous longitudinal studies. *Research on Human Development*, *7*, 45–60.

Spiro, A. III, & Brady, C. B. (2008).Integrating health into cognitive aging research and theory: Quo vadis? In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging: Interdisciplinary perspectives* (pp. 260–283). Thousand Oaks, CA: Sage.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, UK: Cambridge University Press.

Stine-Morrow, A. L., Parisi, J. M., Morrow, D. G., Greene, J., & Park, D. C. (2007). The senior odyssey project: A model of intellectual and social engagement. *Journal of Gerontology: Psychological Sciences*, *62B*, P62–P69.

Thorvaldsson, V., Hofer, S. M., Berg, S., Skoog, I., Sacuiu, S., & Johansson, B. (2008). Onset of terminal decline in cognitive abilities in individuals without dementia. *Neurology*, *71*, 882–887.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Tranter, L. J., & Koutstaal, W. (2008). Age and flexible thinking: An experimental demonstration of the beneficial effects of increased cognitively stimulating activity on fluid intelligence in healthy older adults. *Aging, Neuropsychology, and Cognition*, *15*, 184–207.

Tucker-Drob, E. M. (2011). Global and domain-specific changes in cognition throughout adulthood. *Developmental Psychology*, *47*, 331–343.

Verhaeghen, P., & Salthouse, T. A.(1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and non-linear age effects and structural models. *Psychological Bulletin*, *122*, 231–249.

Wechsler, D. (1939). *Measurement of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Wilson, R. S., Beck, T. L., Bienias, J. L., & Bennett, D. A. (2007). Terminal cognitive decline: Accelerated loss of cognition in the last years of life. *Psychosomatic Medicine*, *69*, 131–137.

Wilson, R. S., Bennett, D. A., Bienias, J. L., Mendes de Leon, C. F., Morris, M. C., & Evans, D. A. (2003). Cognitive activity and cognitive decline in a biracial community population. *Neurology*, *61*, 812–816.

Zelinski, E. M., & Kennison, R. F. (2007). Not your father's test scores: Cohort reduces psychometric aging effects. *Psychology and Aging*, *22*, 546–557.

Zelinski, E. M., Kennison, R. F., Watts, A., & Lewis, K. L. (2009). Convergence between cross-sectional and longitudinal studies: Cohort matters. In H. B. Bosworth & C. Hertzog (Eds.), *Aging and cognition: Research methodologies and empirical advances* (pp. 101–118). Washington: American Psychological Association.

10 Developing Intelligence through Instruction

Raymond S. Nickerson

In the earlier version of this chapter (Nickerson, 2011), I argued that intelligence is a vexed concept and surmised that it was likely to remain so. I believe that it has indeed remained so. However, the working definition that I took then seems still appropriate, so I shall use it again: the ability to learn, to reason well, to solve novel problems, and to deal effectively with the challenges – often unpredictable – that confront one in daily life. Synonyms, or closely related terms, that one finds in the literature include mental capacity, brainpower, acumen, reasoning ability, astuteness, cleverness, expertise, skill, wisdom, insightfulness, cleverness, discernment, good judgment, problem-solving ability, and decision-making ability, among others. Many of these descriptors are generally recognized as traits or capabilities that have to be developed; they are compatible with the conception of intelligence, expressed by Brody (2014), "as a latent trait that is partially determined by influences early in life that remains relatively invariant over the life-span that has important real-world influences on educational achievement" (p. 137).

Many, if not most, theorists make a distinction, first made and elaborated by Cattell (1943, 1963, 1971), between *fluid* intelligence (G*f*) and *crystallized* intelligence (G*c*). Fluid intelligence connotes the ability to engage in abstract reasoning and to solve problems one has not encountered before; crystallized intelligence refers to domain-specific knowledge gained through experience. Although the latter term has been used extensively since it was introduced by Cattell, precisely what is meant by it is not always clear (Keith & Reynolds, 2010). Schipolowski, Wilhelm, and Schroeders (2014) contend that there is no consensus regarding the nature of crystallized intelligence and that it has been changing over time. "Originally conceptualized as capturing acquired skills and declarative knowledge in different content domains, more recent definitions and typical indicators focus on verbal ability" (p. 156). Schipolowski and colleagues hold that while verbal ability and knowledge are closely related constructs, they are "empirically distinguishable facets of crystalized intelligence" (p. 156).

For the purposes of this chapter, the distinction between fluid and crystallized intelligence is not emphasized. Our concern is with the question of whether intelligence, as defined above, and generally quantified by IQ, can be increased by instruction.

An Important Distinction

One frequently encounters in the literature about intelligence the term *teaching intelligence*, when a more accurate term would be *teaching about intelligence*. The intent generally is to convey something about how to inform or educate some group – college students, government agents, general public – regarding what intelligence is, how it is measured, how it affects performance of various tasks, the extent to which it is genetically determined, whether it changes with age, and so on. Relatively seldom do authors intend to convey by *teaching intelligence* the idea of teaching with the goal of increasing the amount of intelligence one has.

The distinction between teaching *about* intelligence and teaching *intelligence* is especially germane for present purposes inasmuch as it helps to sharpen the major focus of the chapter, which is not the question of how to teach about intelligence, but that of whether intelligence can be increased by teaching, or by other interventionist means. I note in passing, however, the claim that interest in teaching about intelligence has been declining in recent years. More specifically, the claim is that teaching about intelligence has less of a presence in college curricula in the United States today than it once did. This is not universally seen as a good thing.

Detterman (2014a, 2014b) argues that there are far too few college courses on intelligence being offered, and far too little understanding by the general public of what intelligence is. Hunt (2014) notes that several of the most prestigious US universities offer no undergraduate courses on intelligence. He makes some suggestions as to why this is the case. He recounts the contentious and highly politicized history of debates about the proper place (if any) of the teaching about intelligence in universities:

> The result was that faculty and administrators who did not themselves study intelligence came to believe that classes on intelligence were at worst an echo of the spurious racism/eugenics arguments of the 1930s and at best a can of worms that should not be opened to avoid trouble on campus. The can of worms argument was strengthened by the fact that, due to cognitive segregation, the personal experiences of most faculty and administrators led them to believe that intelligence was not a very important topic anyway. (p. 158)

Hunt goes on to describe what he recommends for an effective introductory course on intelligence for undergraduates. As suitable texts, he recommends his own (Hunt, 2011) and one by Mackintosh (2011).

Deary (2014) gives a first-person account of teaching (about) intelligence at various levels of schooling and to the general public. The account points to literature that should be helpful to instructors who cover aspects of teaching about intelligence at any level. Deary's emphasis is on individual differences in intelligence. His perspective and teaching techniques are spelled out in numerous books and articles, notably Deary (2000, 2001). Deary has also produced extensive reviews of research on intelligence (Deary, 2012a, 2012b). Other first-person accounts of teaching about intelligence include those of Detterman (2014b), Haier (2014), Hunt (2014), Mackintosh (2014), and Sternberg (2014), all conveniently in the same issue of *Intelligence*.

Should Intelligence Be Taught?

We see that teaching *about* intelligence – whether it should be done, and if so, at what level in the curriculum and in how much detail – is a somewhat delicate subject, regarding which there are a variety of views. But now suppose it were possible to teach intelligence – not *about* intelligence, but *intelligence*. Imagine that people's intelligence, as conventionally measured, could be increased through instruction or other interventions, to a practically significant degree. How should we feel about that?

Haier (2014) refers to the question of whether we can increase intelligence as a key question in the public's mind. He also asks whether we *should* raise intelligence, if we could. "If we could raise intelligence, should we? Consider the current controversy over the use of drugs by students to get better test scores. What about the future, when drugs to enhance memory, attention, and learning may be vastly better?" (p. 155).

The question lacks gravity only because the goals of past and current efforts to increase intelligence are so modest; If an effort to increase intelligence by a few points could be shown to be reliably effective, it would be newsworthy. In other words, the goals are so low, attaining them would constitute no great threat to the status quo. But suppose it were possible to raise everyone's IQ by 30 points, say, perhaps by spiking the drinking water with IQ juice. If we could do it, should we? Would the world become a better place in which to live? Would people get along better with each other? Would wars become less frequent? Would violence and crime diminish? Such questions are unlikely to admit of simple "yes" or "no" answers, and I suspect that much depends on the extent to which any increases in intelligence are accompanied by corresponding increases in moral, ethical, and commonsense sensibilities. We don't want to produce more capable criminals (Boccio, Beaver, & Schwartz, 2018) or "clever sillies" (Charlton, 2009). But how can we rule out the possibility of doing so? The question strikes me as a very important one, and one that has not received the attention it deserves. There is ample evidence that high intelligence alone does not make one immune to myside bias (Stanovich & West, 2008), false consensus effect (Ross, Greene, & House, 1977), overconfidence (Krueger, 2000), and other examples of faulty reasoning. Nor does it guarantee socially acceptable behavior.

Imagine the more modest possibility that intelligence could be raised through the use of special training or educational programs that are readily available to everyone. There already exist many programs, available through the Internet and elsewhere, that promise improved cognitive function if faithfully followed. Such promises are seldom backed up by compelling evidence of the effectiveness of the offered products; but if it could be proved beyond doubt that such products actually work, let us say in raising IQ by 10 or 15 points, what then?

Intelligence and National Well-Being

Numerous studies have shown a positive relationship – not necessarily cause-effect – between a country's national average IQ or other indicants of intelligence and its productivity or pattern of economic growth or technological development, among other indicators of national well-being (Burhan et al., 2014; Jones & Schneider, 2006; Lynn, 2012; Lynn & Vanhanen, 2012; Rindermann & Becker, 2018). Kanyama (2014) presents data suggesting that the higher the average intelligence in a society, the higher the quality of the society's institutions as reflected in such variables as "control of corruption, government efficiency, regulatory quality and rule of law" (p. 44). The importance that is sometimes given to a nation's intelligence, as usually represented by its average IQ, is reflected in reference to it as "a measure of human capital" (Meisenberg & Lynn, 2011).

In a study of the relationship between the average IQ of the residents of the United States, and a variety of indicants of the economic performance of the states, Yang and Lester (2016) found that "States whose residents had higher estimated intelligence, based on standardized tests given to students in the states, had better economic performance, with higher per capita income, stronger growth in gross state product per capita, lower unemployment rates, lower foreclosure rates during the recent economic crisis, and lower credit card debt" (p. 33).

In sum, there is considerable evidence of a positive relationship between national or regional well-being and national or regional average IQ. To be sure, most, if not all, of the studies showing this relationship have been correlational, so they do not establish a cause-effect relationship. Moreover, Daniel (2016), who considers national differences in intelligence as far back as 1500, argues that they cannot be a root cause of inequalities among nations. However, the possibility is there and it should not be overlooked.

Intelligence and Individual Well-Being

There is evidence that people with higher intelligence tend to be mentally and physically healthier (Davies et al., 2018; Gale et al., 2012; Wrar et al., 2015; Wrulich et al., 2014) and to live longer (Calvin et al., 2017), than do people with less intelligence, although which is cause and which is effect is generally not clear. Numerous studies have shown a positive relationship between the preservation of cognitive function and lifestyle over the life span. The incidence of various forms of dementia varies inversely with people's level of education and their habitual engagement in cognitively challenging activities (Hertzog et al., 2009; Hultsch et al., 1999; Ott et al., 1999; Scarmeas et al., 2001; Verghese et al., 2009).

The relationship, again not necessarily cause-effect, between intelligence and personal wealth, power, and social status is seen in the fact that higher-IQ people are greatly overrepresented among the world's billionaires and people in positions of power. Wai (2013) found cognitively elite people in the United States – senators, federal judges, CEOs of major corporations – to be greatly overrepresented among social/financial elites. Wai and Nisen (2013) and Wai (2014) provide details regarding the situation in various countries. Wai (2014) summarizes the findings thus: "Today highly educated and intelligent individuals are overrepresented among the sliver of people who control a disproportionate share of the world's money and power. This shows the importance of education and cognitive ability in being able to attain a position in the global elite. The clustering of brains, wealth, and power may have important implications" (p. 67). Surely there can be little doubt of the importance of the implications of these findings; sorting out the cause-effect relationships involving intelligence and other personal characteristics and assets deserves to be a continuing focus of research.

That IQ is a predictor of criminal behavior – people with lower IQs being more likely than people with higher IQs of being convicted of a crime – is well known (Beaver et al., 2013; Neisser et al., 1996). Again, the data show a correlation – in this case a negative one – but they do not demonstrate a cause-effect relationship. However, the possibility that increasing intelligence by educational means could decrease criminal behavior seems a worthy focus of research.

Intelligence and Other Traits and Abilities

Intelligence relates to numerous desirable human traits and abilities – rationality, expertise, creativity, innovation – to greater or lesser extents. The exact nature of the relationship in specific instances has been of interest to psychologists and economists for a long time. Citing Kaufman and Plucker (2011) with respect to the relationship between intelligence and creativity, Jauk and colleagues (2013) describe the situation this way: "Although empirical creativity research can meanwhile look back on a scientific tradition of over 60 years of investigation, it is still unclear how the concepts of creativity and intelligence relate to each other" (p. 212).

Some researchers emphasize the low correlations that are often found between intelligence and creativity (Kim, 2005; Wallace & Kogan, 1965) while others argue that the connection between these assets is stronger than generally believed (Cho et al., 2010; Nusbaum & Silvia, 2011). Squalli and Wilson (2014) argue that the question of the relationship between intelligence and creativity is under-researched. For present purposes, a question of interest is whether creativity can be increased through training. This too appears to be a question that is under-researched. However, creativity is generally considered to be a valuable asset, and the evidence appears to be that the relationship between intelligence and creativity is positive, if not strong; and this is a good reason for teaching intelligence if it can be taught.

Another reason for teaching intelligence is the possibility of increasing the retention of cognitive capacity over the life span. Mean IQ scores tend to change systematically with age, rising from adolescence until the mid-twenties and then falling regularly, perhaps by as much as 25 to 30 percent over the next fifty years (Wechsler, 1981).

There are many questions about the relationship between intelligence and aging. Of special interest here is the possibility of preventing, or at least attenuating, the decline in cognitive function that typically occurs in advanced age of otherwise healthy adults. Is "use it or lose it" more than a catchy adage? Are there effective strategies for maintaining intelligence into advanced age – for "stretching" gray matter, as it were? Does exercising the brain by playing chess, doing crossword puzzles, solving Sudokus or KenKens, playing Scrabble or other word games that force one to think help keep the brain in shape? Are there effective programs the purpose of which is to forestall cognitive decline with age, or can such programs be designed? Such questions become more and more important as the expected life span continues to increase.

Another focus of interest among researchers who study aging is evidence regarding the adult and aging brain being able to continue development to a greater degree than was long believed. Of particular interest for present purposes is the finding that neurogenesis appears to be facilitated by mental activity, which suggests the importance of lifestyle factors in maintaining brain function (Jak, 2011; Valenzuela, Breakspear, & Sackdev, 2007; Valenzuela et al., 2008).

Worldwide average life expectancy has been increasing regularly for the past few centuries and there is every reason to believe that, barring some global catastrophe of epic proportions, it will continue to do so at least for the foreseeable future. That being the case, the possibility of finding ways to preserve intelligence into old age should have a high priority among researchers, doctors, government officials, and the general public alike. Letting the current relationship between intelligence and age persist would mean dramatically enlarging the fraction of the population that is cognitively impaired as life expectancy continues to increase.

In sum, there are many reasons for teaching intelligence, if it can be taught. It should not be taught in a vacuum, but in the context of education that emphasizes the importance of ethics, social skills, and what it means to be a responsible citizen and human being.

Can Intelligence Be Taught?

Having argued that intelligence *should* be taught, if it *can* be, we come now to the question of whether intelligence can be taught – not taught *about*, but *taught*. Can intelligence be enhanced intentionally by education or training?

That the answer to this question has considerable practical significance is seen in the fact, as pointed out by Hayes, Petrov, and Sederberg (2015) that “Millions of customers buy ‘brain building’ games and subscribe to ‘mental gyms’ on-line where they perform various ‘cognitive workouts’ in the hope of raising their IQ . . . Hundreds of millions of dollars are being invested in educational (e.g., Cogmed, www.cogmed.com), military, and commercial programs (e.g., Lumosity, www.lumosity.com) on the assumption that intelligence can be improved through training” (p. 1).

Reasons for Believing that Intelligence Is Malleable

In Nickerson (2011), I suggested several reasons for believing that intelligence is malleable. Here I shall supplement that discussion with some additional and more recent relevant references.

• *Experiences can affect the central nervous system.* Young brains have greater plasticity than adult brains, but the latter also have generative – and regenerative – ability; the extent to which and the conditions under which new brain tissue and connections can be produced are active areas of research (Bouchard & Villeda, 2014; Gage, 2003; Greenwood, 2007; Nottebohm, 2002; Pardon & Bondi, 2011; Park & Reuter-Lorenz, 2009). Of particular interest for present purposes is the finding that neurogenesis appears to be facilitated by mental activity, which suggests the importance of lifestyle factors in maintaining brain function (Bouchard & Villeda, 2014; Pardon & Bondi, 2011).

• *The IQs of individuals may change over time.* Although IQ test scores of an individual obtained at one time correlate highly with those obtained from the same individual at other times (Bradway, Thompson, & Cravens, 1958; McCall, Appelbaum, & Hogarty, 1973), the correlation is far from perfect, and investigators have documented many cases of large increases or decreases in measured IQ (Anastasi, 1988; Honzik, Macfarlane, & Allen, 1948; Schneider, Niklas, & Schmiedeler, 2014).

• *Beliefs about intelligence can affect performance*. Beliefs are important determinants of behavior, including performance of cognitively demanding tasks (Andrews & Debus, 1978; Baron, 1991; Deci & Ryan, 1985; Haimovitz, Wormington, & Corpus, 2011). People who believe that intelligence is malleable are more likely to attempt to improve their problem-solving capabilities than are those who believe it to be innate and fixed (Dweck, 1999; Heyman & Dweck, 1998). Beliefs about the nature of intelligence – whether it is immutable – can be changed through instruction and in ways that can translate into improved performance (Hong et al., 1999).

• *Motivation, practice, and persistence matter.* The importance of motivation as a determinant of performance on cognitively demanding tasks has long been known (Botvinik & Braver, 2015). It is possible to have the ability to act in a certain desirable way and lack the disposition to do so. Data obtained by Klemp and McClelland (1986) from interviews with average and outstanding managers regarding their own job performance led these investigators to the conclusion that outstanding managers differ from average managers more with respect to disposition than with respect to capacity. Arguing that rationality is "less a matter of capability than of a *disposition* to shape one's beliefs by evidence and to strive to maintain consistency among those beliefs" (emphasis in original), Stanovich (1994) coined the term *dysrationalia* to mean "the inability to think and behave rationally, despite adequate intelligence" (p. 11).

Duckworth and Seligman (2005) have reported data that suggest that indicators of motivation may do at least as well as IQ in predicting students' course grades. Students from East Asia typically outperform American students in educational achievement (Byun & Park, 2012; Geary, 1996; Stevenson, Chen, & Lee, 1993; Stevenson, Lee, & Stigler, 1986). Factors that have been identified as probably contributory to these differences in performance include motivation, beliefs about the dependence of success on effort, and the relatively high value that Asian parents place on academic achievement (Caplan, Choy, & Whitmore, 1992; Chen & Stevenson, 1995; Tsang, 1988).

The measurement of intelligence generally is done by having people perform certain cognitively demanding tasks. The extent to which practice can enhance the performance of these or similar tasks is a question for which different experts give different answers. In a widely cited review of the role of practice in the development of expertise, Ericsson and colleagues (1993; see also Ericsson, 2007; Ericsson, Prietula, & Cokely, 2007) argue that (discounting such physical factors as body size) the difference between the performance of an expert and that of a novice is largely due to the different amounts of time they have spent practicing the task. Many researchers have accepted the idea that "practice makes perfect," and have applied it to essentially all types of skills. Some argue that practice alone does not account fully for the differences between expert and novice performance (Ackerman, 2014; Anderson, 2000; Gardner, 1995; Hambrick et al., 2014), but the evidence that it has some beneficial effect is strong.

• *Working memory appears to be malleable*. Working memory capacity has been considered by some researchers to be a major determinant of intelligence. Tourva, Spanoudis, and Demetriou (2016), for example, present data that they interpret to mean that working memory is the main cognitive function underlying general (fluid and crystallized) intelligence in children and adolescents. So it is not surprising that efforts have been made to increase intelligence by enhancing people's ability to perform memory-dependent tasks. Results from several studies suggest that working memory capacity can be increased by practice (Thorell et al., 2008; Verhaeghen, Cerella, & Basak, 2004). Whether equating improved performance on the working memory tasks studied with increased working memory capacity is the best explanation of the results has been questioned, but the point that training can produce improvements in performance seems not to be contested.

• *Education correlates positively with IQ* (Brinch & Galloway, 2012; Ceci, 1991). On the basis of a meta-analysis involving more than 600,000 students, Ritchie and Tucker-Drob (2018) concluded that intelligence test scores are positively correlated with the duration of education. They note that the correlation could be interpreted in two ways: "Students with greater propensity for intelligence go on to complete more education, or a longer education increases intelligence" (p. 2). More generally, Ritchie and Tucker-Drob not only recognize that the association between education and intelligence could be a consequence of a selection process, they also note that evidence makes it clear that such selection processes exist.

The existence of such processes complicates considerably the goal of determining whether education increases intelligence. (The possibility that education increases intelligence *and* that more intelligent people seek more education are *both* true should not be overlooked; indeed, such a positive feedback situation seems intuitively likely.) However, presenting a variety of statistical analyses, on numerous studies designed to control for the endogeneity problem – "confounds resulting from selection processes, where individuals with a propensity toward higher intelligence tend to complete more years of education" – Ritchie and Tucker-Drob argue that the evidence shows that intelligence is increased by between 1 and 5 IQ points for every year of education. An analysis of the data combined from three different study designs yielded an average of 3.4 IQ points for every year of education; however, Ritchie and Tucker-Drob use the less specific, but safer, estimate of 1 to 5 points in their concluding observation that "the results support the hypothesis that education has a causal effect on intelligence test scores" (p. 16).

Ritchie and Tucker-Drob did not include in their sample children under the age of six, so to the extent that changes in intelligence occur in that cohort, perhaps as a consequence of preschool attendance, they are not reflected in their data. Nor do they distinguish among educational programs of different quality, a factor that could turn out to have a modulating effect on the relationship between years of schooling and intelligence (Allensworth et al., 2017).

• *Average intelligence has been increasing.* Average scores on standardized intelligence tests have been increasing regularly around the world at the rate of about a point approximately every three years (Neisser, 1997, 1998). This is generally known as the "Flynn effect," named for James Flynn, who published widely cited articles about it (Flynn, 1984, 1987, 2007). Numerous studies of the Flynn effect have yielded a great deal of data (Williams, 2013). The variety of results that have been obtained makes them difficult to summarize, but the basic finding of increasing IQ (G*f* and G*c*) scores over time has been found again and again in different countries and in a variety of different groups. Several theoretical explanations of the effect have been offered, but none has dominated all the others. For the purposes of this chapter, the important point was made by Nisbett (2009) who followed a discussion of the evidence that average intelligence has been increasing with the observation that "the fact that gains have occurred over time in skills that society cares about – both for everyday life and for advanced work in science, industry, and other professions – establishes that people can become smarter in very real and important ways" (p. 56).

We see that there are many reasons for believing that intelligence is malleable. Precisely *how* malleable it is remains undetermined, but evidence supports the conclusion that it is malleable enough to make a practical difference, and to warrant continuing efforts to find effective ways to increase it.

Early-Life Experiences as Determinants of Adult IQ

Considerable attention has been (and is being) given to the question of how early-life experiences – especially family context – affect later-life capabilities (Barreto et al., 2017; Hammond et al., 2012). School attendance can positively affect adult IQ (Brinch & Galloway, 2012; Ceci, 1991; Ceci & Williams, 1997; Clouston et al., 2012), as can parental training of children with respect to mental and motor skills (McCall et al., 1973). Adoption of a child into a supportive family can have a similar effect (Deary, 2013). As Nisbett (2009) puts it, "being raised under conditions highly favorable to intelligence has a huge effect on IQ" (p. 32). (It would be good to know how IQ changes during the first few years of life, but although there is evidence that individual differences in intelligence can be assessed in infants [Colombo, 1993; Fagan, 2011], I am unaware of successful efforts to attach IQ scores to infants.) My sense is that there is general agreement among researchers that early life experiences can have profound effects on adult IQ, but that precisely what those experiences are is still not clear.

Genetics and Environment

Much of what is known about genetics as a determinant of intelligence has been obtained from studies of twins and other siblings. These studies have shown that the IQs of identical twins generally are closer than are those of non-twin siblings or those of fraternal twins, and those of identical twins reared apart are nearly as similar as those of identical twins reared together (Plomin & DeFries, 1998; Plomin et al., 1994).

It is still the case that the relationship between heredity and environment as causes of intelligence as represented by IQ is not known, although different researchers have different proposals on the matter. Probably all who have studied the subject agree that both are significant determinants and that how they relate and interact in specific individuals depends on a host of other factors. Nisbett (2009) estimates that, in the aggregate, the maximum contribution of genetics is probably about 50 percent, and that the remaining variation is largely due to environmental factors. Citing Neisser and colleagues (1996), Hayes and colleagues (2015) give between 50 and 75 as the percentage of "the variance of intelligence test scores in healthy adults [that] is linked to genetic variations" (p. 2).

Sternberg (2012) estimates that inheritance may account for between 40 and 80 percent of intelligence, but notes that "heritability varies as a function of socio-economic status and other factors" (p. 19). He also cautions that such figures are easily misinterpreted. "Both heritability and environmentality [the complement to heritability] are applicable only to populations, not to individuals. There is no way of estimating heritability for a particular individual, nor is the concept of heritability even meaningful for individuals" (p. 22). He goes on to say:

> Heritability and environmentality add up to 1. Thus, if IQ has a heritability of .50 within a certain population, then 50% of the variation in scores on the attribute within that population is due (in theory) to genetic influences. This statement is completely different from the statement that 50% of the attribute is inherited. Similarly, if a trait has a heritability equaling .70, it does not mean that the trait is 70% genetic for any individual, but rather that 70% of the variation across individuals is genetic. (p. 22)

Further, Sternberg notes that genetic effects occur in environments, which may modulate those effects; in other words, it does not follow from the fact that a trait is hereditary that it cannot be modified. "Because the value of a given heritability statistic is relevant only under existing circumstances, the statistic does not and cannot address the modifiability of a trait. A trait could have a high level of heritability and nevertheless be highly modifiable" (p. 23).

It is easy to find different estimates of the relative importance of genetics and environmental factors as determinants of intelligence. For present purposes, the important point is that the evidence that one's intelligence is not entirely determined by genetics justifies the question of whether, or the extent to which, it might be increased by nongenetic factors, such as instruction. (The possibility of changing it genetically is sure to be a topic of increasing interest in the future, not only to scientists, but to ethicists, lawmakers, and the general public as well. A discussion of the possibility is beyond the scope of this chapter.)

Detterman (2014a) summarizes the articles published in a special issue of *Intelligence* on the development of expertise this way: “In sum, the papers included here present data or logical arguments that heritable ability, and particularly intelligence, is a necessary component to acquiring expertise in nearly any domain. The evidence presented comes from diverse sources and seems nearly overwhelming in both its depth and breadth” (p. 3). This assessment identifies intelligence as a necessary component to the acquisition of expertise. It seems appropriate, too, to consider the possibility that increased intelligence could be a consequence of efforts to acquire expertise.

Organized Attempts to Increase Intelligence

Given the considerable evidence that intelligence is malleable, and that it is so pretty much throughout the entire life span, it is only natural to expect there to be organized efforts to increase intelligence – or, if one prefers, to improve people’s performance on cognitively demanding tasks. And there have been many such efforts. Here I will briefly describe three of them in which instruction has played a leading role.

Head Start

The largest, most durable and probably best-known project aimed at increasing intelligence is Head Start (Payne et al., 1973). Established by the US government in 1965, and still functioning, the purpose of this program is to promote school readiness in disadvantaged preschoolers – mostly three- and four-year-olds – by helping them develop early reading and mathematics skills that will contribute to their later success in school. In 1995, the program was extended, with the establishment of Early Head Start, to include children from birth to age three. The program is administered by the Office of Head Start, within the Administration for Children and Families, US Department of Health and Human Services.

Head Start functions as an umbrella entity under which numerous local projects exist – mostly in preschool classrooms – throughout the United States. Parental involvement is strongly encouraged. Funding increased from approximately $200 million for its first full year (1966) to approximately $9.5 billion for the fiscal year 2008. According to the National Head Start Fact Sheet, total expenditures for Head Start and two closely related programs – Early Head Start and Early Head Start–Child Care – were still approximately $9.5 billion in the fiscal year 2017. As of the end of the fiscal year 2017, the program claimed a total enrollment of approximately 900,000 and functioned in approximately 57,000 classrooms.

Since the beginning there have been issues concerning objectives (what the precise goals of the project should be) and evaluation (how success or failure should be assessed). Early in the project’s history, a panel of experts tasked with defining social competency identified twenty-nine components that could serve as goals for the project (Anderson & Messick, 1974). There appears to have been general agreement that assessment should not focus, at least not exclusively, on effects of the program on IQ scores (Lewis, 1973; Sigel, 1973).

Published assessments of the effectiveness of Head Start are mixed, ranging from severely critical, through middle of the road (pointing out what the assessors see as the program's strengths and weaknesses), to strongly positive. Among the more thought-provoking outcomes of assessment efforts is the finding that, although substantial gains in performance are realized while the children are participating in the program, the gains appear to diminish, if not disappear, after participation in the program is over and the children have entered school (Lazar & Darlington, 1982; McKay et al., 1985; Ramey, Bryant, & Suarez, 1985), The post-participation fading of the positive effects has been blamed by some on the low quality of the schools that most Head Start participants enter (Lee & Loeb, 1994). There is also the claim that the fading of the effects is a myth (Barnett, 2002).

The Carolina Abecedarian Project

The Carolina Abecedarian Project was established in 1972 to address the needs of preschoolers and schoolchildren considered to be at risk for delayed development and school failure through the first three years of elementary school. Participants were from low-income families, mostly African American (98%) and single female parent (85%). Parents' average age was twenty and their average IQ 85. The preschool program was a daycare service that provided, for children from six weeks of age until entry to kindergarten, nutritional supplements, pediatric care, social work services and, of special interest in the present context, an environment intended to enhance cognitive and linguistic development. For children three years old and older, this environment included structured curricula designed to become increasingly similar to what a child would experience on entering public school. The program for school-age children provided a resource teacher for each child, who served as an intermediary between the classroom teachers and parents, facilitating communication both ways and engaging parents in home activities with children to support and complement what was being taught in the classroom. Resource teachers made frequent visits both to their students' schools and homes.

Evaluation of the program involved a controlled study in which participants were assigned to intervention and control groups. Performance data on a variety of intelligence and abilities tests were collected at various times during the intervention and at regular intervals for several years later (from former participants at ages ranging from 8 to 21 years). Results of evaluation studies are documented in a series of publications (Burchinal, Lee, & Ramey, 1989; Horacek et al., 1987; Martin, Ramey, & Ramey, 1990; Ramey & Campbell, 1984, 1994). Longer-term results are reported by Campbell and Ramey (1994, 1995), Campbell and colleagues (2002), and Clarke and Campbell (1998). In brief, scores on assessment tests were higher for children in the intervention group than for those in the control group over the entire span of the assessment period; academic achievement of the children in the intervention group was also enhanced. Evidence of positive effects on the subsequent education and employment of parents of participating children was also obtained.

Wikipedia provides a list of quantitative positive results from the study, including results observed at various times – e.g., three, twenty-one, and thirty years – after participation in the program. The list is impressive; however, criticisms that have been made of the evaluation are also noted, some of which challenge conclusions drawn from the data.

Project Intelligence

Project Intelligence is the label that was given to a project undertaken in Venezuela in the early 1980s. The idea for the project originated with Luis Alberto Machado, then Venezuelan Minister of State for the Development of Human Intelligence, a post created at his suggestion to make possible the establishment of a variety of innovative projects aimed at improving the educational opportunities and accomplishments of Venezuelan youth. Machado was a firm believer that intelligence is determined, to a large extent, by experience, especially by events in early childhood. A visionary and activist, he had aggressively promoted the idea that the state has an obligation to see that every child has the opportunity to develop their potential intelligence to the fullest, and had expressed his views and vision in several publications, notably *The Right to Be Intelligent*, which appeared in 1980, shortly after creation of the ministerial post that he occupied.

Project Intelligence was undertaken, at Minister Machado's request, as a collaboration among researchers at Harvard University, Bolt Beranek and Newman Inc. (BBN), and teachers in Venezuela. It is described in several publications (Adams, 1989; Chance, 1986; Nickerson, 1987, 1994; Nickerson, Perkins, & Smith, 1985; Perkins, 1995) and most completely in the project's final report submitted to the government of Venezuela (Harvard University, 1983) and in Herrnstein and colleagues (1986).

The project's objectives were to develop and evaluate materials and methods for teaching cognitive skills in seventh-grade classrooms in Venezuela. A one-year course intended to engage students in discussion and thought-provoking classroom activities was designed and implemented in several Venezuelan schools. Course materials and activities focused on specific capabilities such as observation and classification, critical and careful use of language, reasoning, problem-solving, inventive thinking, and decision-making. Development of the materials was a collaborative effort among members of the Harvard/BBN team in consultation with several experienced Venezuelan teachers who were to prepare a larger group of Venezuelan teachers to use the materials in a planned year-long evaluation.

The evaluation used matched experimental and control groups in six public schools in Barquisimeto, Venezuela – twenty-four classes, four from each school, the four classes from three of the schools serving as the experimental classes and the four from the other three serving as controls. Each class had approximately thirty to forty students. Control classes were matched, insofar as was possible, with experimental classes. The experimental classes, which were taught by regular Venezuelan middle school teachers who had volunteered to participate in the project, met for about forty-five minutes a day, four days a week. Tests that were used for evaluation purposes were the Otis-Lennon School Ability Test (Olsat) (Otis & Lennon, 1977), the Cattell Culture-Fair Intelligence Test (Cattell & Cattell, 1961), and a group of general abilities tests (GAT) (Manuel, 1962). In addition, about 500 special test items were constructed to assess competence with respect to the specific skills the course was intended to enhance.

The standardized general abilities tests and the target abilities tests were administered to experimental and control groups before and after the teaching of the course. Both groups improved their scores on both types of test over the period of the course. The effectiveness of the course was judged by comparing the magnitudes of the gains realized by the two groups. Details of test administration and test results are reported in Herrnstein and colleagues (1986) and Swets and colleagues, 1988). Gains on both types of test were significantly greater for the experimental students than for the controls. The gains realized by the students in the experimental classes were 121 percent, 146 percent, 168 percent, and 217 percent of those realized by the controls on the Cattell, the Olsat, the GAT and the target abilities battery, respectively. Further analyses showed the magnitude of the gains to have been relatively independent of the initial ability levels of the students as indicated by pretest scores. Unfortunately, data regarding long-term effects of the intervention are not available. Presumably whether gains realized in any limited-time project of this sort are maintained and amplified following completion of the project will depend greatly on the extent to which subsequent educational experiences build on them.

An English adaptation of parts of the Project Intelligence course was published in 1986 by Mastery Education Corporation under the title *Odyssey: A Curriculum for Thinking*. I am unaware of any data regarding sales of the adaptation or of the extent or effectiveness of its use.

Others

There have been many other organized efforts to improve cognitive performance. Some of these are described in Nickerson and colleagues (1985), including the Instrumental Enrichment Program (Feuerstein et al., 1980), the Structure of Intellect Program (Meeker, 1969), Science – A Process Approach (Gagne, 1967; Klausmeier, 1980), Thinkabout (Sanders & Sonnad, 1982), Basics (Ehrenberg & Ehrenberg, 1982), Patterns of Problem Solving (Rubenstein, 1975), Schoenfeld's (1985) approach to teaching mathematical problem-solving, and the Productive Thinking Program (Covington et al., 1974), among several others. Some of these programs, and others, are also described in Nickerson (1988/1989, 1994) and in Perkins (1995).

Among the programs that have been developed are several designed to provide remedial help for college students to develop the cognitive (or metacognitive, self-management) skills needed to do well with conventional college work. Examples described in Nickerson and colleagues (1985) are ADAPT (Accent on the Development of Abstract Processes of Thought), DOORS (Development of Operational Skills), COMPAS (Consortium for Operating and Managing Programs for the Advancement of Skills), SOAR (Stress on Analytical Reasoning) and DORIS (Development of Reasoning in Science). The programs mentioned, among others, are responses to the need for remedial training for many students entering college that has been well documented in numerous reports (e.g., *A Nation at Risk* [National Commission on Excellence in Education, 1983] and others). Unfortunately, evaluative data regarding the effectiveness of the various efforts to address this problem are less plentiful and conclusive than one would like.

Results of Efforts to Teach Intelligence

Hayes and colleagues (2015) argue that of three important questions – (1) What is intelligence? (2) How is it measured? And (3) How can *gains* in intelligence be measured? – the first two have been debated and researched for over a century, while the third has not received the attention it deserves. They contend that many of the experimental results that have been taken as evidence that intelligence can be (has been) taught do not really support that conclusion. The problem, in their view, is that the test-retest method that is typically used to determine whether intelligence has increased is seriously flawed. "The overwhelming majority of studies use test-retest score gains to measure G*f* gains. . . . This practice is based on the misleading intuition that if a test such as Raven's APM is a valid measure of G*f*, then a gain in the score on this test is a valid measure of G*f* gain" (p. 2). Hayes and colleagues argue that this inference cannot be made because test (and retest) scores are affected not only by intelligence but by other factors, such as visuospatial ability, motivation, and test-taking strategy, as well, and those factors can improve even if intelligence does not. Hayes and colleagues present data that they interpret to show that test-score gains that are used to measure the effects of cognitive training may reflect refinement of test-taking strategy rather than gains in intelligence.

Referring to the distinction between *skills* and *abilities* (Anderson, 2000), Hayes and colleagues contend that, although it is hard to acquire specific skills, it is very much harder to improve general abilities. Following a review of efforts to do the latter, they conclude, regarding whether intelligence can be improved with training:

> The issues are complex and much of the current disagreement stems from incompatible interpretations of the vague and ambiguous term "fluid intelligence." One important piece of this large puzzle is the ability to flexibly deploy a judicious variety of cognitive strategies and to adaptively learn their utilities for various tasks. If this ability is taken to be part and parcel of G*f* then the answer to the opening question [whether intelligence can be taught] may well be yes. If, however, G*f* is interpreted in narrow neurobiological terms (e.g., Duncan et al., 2000; Gray & Thompson, 2004) then the answer remains elusive. So far we have seen no conclusive evidence that the brain can be trained like a muscle. (p. 11)

Daugherty and colleagues (2018) gave healthy adults several months of training on one of several objectives – physical fitness, cognitive ability, mindfulness, and combinations thereof. Results were mixed. Some training combinations showed improvement on some tests of intelligence; physical training by itself produced no improvement in cognition.

Jaeggi and colleagues have reported positive results on fluid intelligence of training on short- and long-term memory tasks (Jaeggi et al., 2008; Jaeggi et al., 2011). However, attempts to replicate results indicating gains in intelligence from training of long- or short-term memory have not always been successful (Chooi & Thompson, 2012; Greenwood & Parasuraman, 2015; Harrison et al., 2013).

Barnett (1995) analyzed results from thirty-six early-intervention programs. He found no consistent pattern of effects on either IQ or achievement tests beyond the end of elementary school. However, more recently, Barnett (2011) reviewed four early-intervention programs – Head Start, Early Head Start, Perry Preschool and Abecedarian – and found positive effects in all four cases, and that effect sizes declined over time, but at different rates and to different degrees: "HS and EHS effect sizes start small and disappear shortly after school entry, whereas Perry and Abecedarian effects are relatively large and long-lasting" (p. 976). Barnett considered several possible explanations for these results and concluded that early educational intervention can have substantive effects on cognition, among other desirable results. He noted that long-term effects may be smaller than initial effects, but may still be substantial, and argued that although not every early-intervention program has been successful, the potential return on investment of such programs justifies their cost.

Clearly, the answer to the question of whether intelligence can be taught must depend, to no small degree, on how one defines *intelligence.* I have argued elsewhere (Nickerson, 2004) that if one takes a broad view of what constitutes intelligence, as several theorists do (Gardner, 2011; Perkins, 1995; Sternberg & Wagner, 1986), increasing it substantially in one or more respects is generally seen to be feasible (Gardner et al., 1994; Sternberg, 1986; Swartz, 1991).

Deary (2013) describes the situation regarding the question of whether intelligence can be taught thus:

> There is still unresolved researching and discussion of the possible social boosters of intelligence. For example, adoption from a deprived to a more affluent setting is reported to be associated with an intelligence advantage. There is still debate about the effectiveness of intensive intervention programmes early in life, and whether any cognitive advantages last or whether advantage accrues to social rather than cognitive skills. (p. 676)

Crane and Barg (2003) note that the primary objective of the early-intervention programs that were developed in the 1960s "was to raise the intellectual achievement of disadvantaged children" (p. 2). They contend that while "there is little doubt that these programs can increase test scores in the near term, . . . there is a definite tendency for these gains to fade out over time" (p. 2). As to whether it is possible that such gains can be made permanent, they see the evidence as mixed. They argue, however, that there is increasing evidence that early-intervention programs can produce lasting changes in social behavior, as distinct from increased cognitive capacity.

Crane and Barg (2003) discuss several major projects/programs – Head Start, the Abecedarian Project, the Milwaukee Project (Garber, 1988), the Chicago Child-Parent Centers project (Reynolds, 1998), the Elmira Nurse Home Visiting Program (Olds et al., 1998), and the Syracuse University Family Development Research

Program (Lally, Mangione, & Honig, 1998) – that they believe have been shown to reduce the incidence of social problems from participants when they reach adolescence and adulthood. They make the case that even if early-intervention programs such as those considered are unsuccessful in producing large and lasting increases in intellectual ability, they may be worth their cost because of their beneficial effects on social problems.

Citing several studies by him and his colleagues (e.g., Detterman & Sternberg, 1982; Jaeggi et al., 2008; Sternberg & Grigorenko, 2007; Sternberg, Jarvin, & Grigorenko, 2011), Sternberg (2014) draws the conclusion that "Intelligence is modifiable in some degree, although there is no consensus on just what this degree is" (p. 179).

Producing compelling demonstrations of successful efforts to increase intelligence, and sustain the increase, has proved to be frustratingly difficult. Citing Campbell and Burchinal (2008) and Campbell and Ramey (1994), who note the disappointing results of efforts to increase IQ through early interventions, Brody (2014) contends that "Although there is evidence that intelligence may be increased as a result of environmental interventions, there is relatively little evidence of large, enduring changes attributable to manipulations designed to increase intelligence" (p. 137). Generally, increases, when they have been attained at all, have been modest at best (Melby-Lervåg̊ & Hulme, 2013; Papageorgiou et al., 2016; Shipstead, Redick, & Engle, 2012).

Increasing IQ versus Increasing Cognitive Ability

One might argue that a better (more tractable) question than "Can intelligence be increased?" is "Can people be taught to behave (perform cognitively demanding tasks) more effectively (more intelligently)?" There can be little doubt, in my view, that increasing one's ability to meet life's many challenges effectively is considerably more important than that of increasing one's IQ – assuming both are possible. Unfortunately it is much easier to measure success or failure with respect to the goal of increasing IQ, as measured by some standardized test, than with respect to increasing one's ability to deal effectively with the cognitive problems that life presents.

IQ testing remains the most widely used measure of cognitive ability, for a variety of reasons: IQ tests are easy to administer, they yield a number that is easy to understand (or so it appears), they are quite stable over time (Gottfredson, 1997; Schwartzman et al., 1987; Yu et al., 2018), and they have proved to be relatively predictive of mental competence in school or the workplace (Deary et al., 2007; Roth et al., 2015). "A single index of intelligence obtained at age 11 is more predictive of a variety of outcomes than any other single thing we can know about a person" (Brody, 2014, p. 139). Perhaps more than any other property, the predictive function is what gives IQ testing its practical value (Carroll, 1993; Jensen, 1998).

However, despite the relatively straightforward meaning of IQ, the question of how to determine whether it has been increased by one or another intervention is the subject of much debate. Hayes and colleagues (2015) ask whether measured improvement in performance on IQ, or IQ-like, tests demonstrates that intelligence has been increased. They distinguish between *strategy refinement* and *intelligence gains*. One might argue that, for practical purposes, this may be a distinction of little value, because improving one's ability to use problem-solving strategies effectively in real-life (out-of-the-laboratory) situations is of greater interest than raising one's IQ test score.

Head Start, The Carolina Abecedarian Project, and Project Intelligence, all discussed above, were different from each other with respect to goals, methodology, number of people involved, duration, cost, and evaluation, among other respects. All were ambitious. Only Head Start, by far the largest project, continues. Evaluations of long-term effects of Head Start and the Abecedarian Project have been mixed, ranging from strongly positive to strongly negative in both cases. Immediate effects of Project Intelligence were encouraging, but assessment of long-term effects was not feasible.

Nisbett (2009) includes the Perry Preschool Program (Schweinhart et al., 2005), the Milwaukee Project (Garber, 1988), and some replications of the Abecedarian Project (Campbell & Burchinal, 2008) among the early-intervention programs that have produced substantial improvements in school grades and other indicators of academic achievement and, in some cases, post-school success, without yielding noteworthy lasting increases in IQ. He concludes that "early childhood intervention for disadvantaged and minority children works – when it is strenuous and well conducted" (p. 130). Among the achievement gains that have been observed are "lower percentage of children assigned to special education, less grade repetition, higher achievement on standardized tests, better rates of high school completion and college attendance, less delinquency, higher incomes, and less dependence on welfare" (p. 130).

A Perspective

In view of the mixed results of significant past efforts to increase intelligence, the question naturally arises as to whether the efforts are worth their costs: Should efforts to increase intelligence continue to be made? In attempting to answer this question, there are two ways to be wrong: (1) assuming intelligence is not changeable and it is; (2) assuming intelligence is changeable and it is not. The question now becomes, which of these two ways to be wrong is the more regrettable. It seems to me that it is the first one. If we assume that intelligence is not changeable and it is, we will have missed an opportunity to have effected a major educational accomplishment; while if we assume that intelligence is changeable and it is not, at worst we will have wasted a bit of time in chasing a fantasy.

What Can Be Taught to Increase One's Ability to Perform Cognitively Demanding Tasks?

Imagine that it were possible by instruction either (1) to raise one's IQ score or (2) to enhance one's ability to learn, to reason well, to solve novel problems, and to deal effectively with the challenges of daily life, but not to do both. Surely there can be no question about the preference for the second objective over the first. In fact, nothing we know rules out the possibility of doing both, and it seems likely that it would be difficult to do one without also doing the other to some degree.

Assuming one wants to enhance the cognitive performance of people, and one is not concerned with whether in doing so one also increases their IQ scores, what might one do? I believe the evidence indicates that there is much that can be taught that can be effective in realizing that goal. Among the possibilities are the following, most of which I have discussed elsewhere (Nickerson, 1988/1989, 1994, 2004). This list is essentially as it appeared in the earlier version of this chapter. I believe it to be as cogent now as it was eight years ago.

• *Knowledge*. The importance of domain-specific knowledge to effective problem-solving in specific domains has been emphasized by many researchers (Hunter, 1986; Larkin et al., 1980). Knowledge about cognition, and especially about how human reasoning commonly goes astray (e.g., confirmation bias, myside bias, gambler's fallacy, rationalizing versus reasoning, effects of preferences on beliefs, overconfidence in one's own judgments, weighting irrelevancies in argument evaluation, and so on) has also been stressed (Evans, 1989; Nickerson, 1998; Piattelli-Palmarini, 1994; Stanovich, 1999).

• *Logic* (*both formal and – perhaps more importantly – informal)*. The teaching of formal logic as a means of enhancing cognitive performance is not promoted by most psychologists and educators. Some argue that it has little to do with the way people actually think (Cheng & Holyoak, 1985; Evans, 1989); arguments need not be logically valid to be persuasive (Nickerson, Butler, & Barch, in press). Despite of all this, I lean toward believing that neglecting it is a bad idea; and there is some empirical evidence to support this view (Dickstein, 1975; Rips & Conrad, 1983). Familiarity with informal logic – with techniques commonly used to persuade and/or win arguments – strikes me as an important requirement for intelligent living in modern society.

• *Statistics*. Much of the problem-solving and decision-making that people do in their daily lives is done under conditions of uncertainty. Judging the likelihoods of possible events, assessing the risks associated with specific courses of action, estimating costs and benefits of possible consequences of decisions are things we all do frequently, either explicitly or implicitly. Dealing with situations that require probabilistic or statistical thinking is improved by training in probability or statistics (Fong, Krantz, & Nisbett, 1986; Kosonen & Winne, 1995).

• *Specific cognitive skills*. Increasingly in recent years researchers have been exploring the effectiveness of efforts to train people – especially elderly people – on specific cognitive skills. Target skills include methods to improve attention control, memory (mnemonic systems), visual search, reasoning, and performance on other tasks of the types that are found on tests of intelligence. The results of such efforts have been mixed – and transfer of positive results to tasks other than those on which training is focused has been limited – but, on balance, the results have been sufficiently promising to motivate further research (Jaeggi et al., 2011). Hertzog and colleagues (2009) point out that most training studies in this area are of very short duration relative to the time it typically takes in the normal course of life to acquire or hone cognitive skills; it remains to be seen what can be accomplished with much longer training regimens.

• *Strategies/heuristics*. Strategies for learning are teachable (Jones et al., 1987; Paris, Lipson, & Wixson, 1983), as are strategies for problem-solving (Bransford & Stein, 1984; Wickelgren, 1974), and for decision-making (Beyth-Marom et al., 1991). Some strategies are general, not specific to subject matter or problem type; these include breaking the problem down into manageable bites, finding a similar (but easier or more familiar) problem, finding a helpful way of representing the problem (a figure, a table, a flowchart), working backward (from where one wants to be – at the solution – to where one is), considering extreme cases, and so on. Specific disciplines and problem domains have heuristics and "tricks of the trade" that are teachable and useful for people that work in those areas. Domain-specific heuristics are typically more effective than the more general ones for problems in the relevant domains, but are less likely to be useful across domains.

• *Self-management and other metacognitive skills and knowledge.* The effectiveness of self-monitoring and self-management skills and knowledge is well documented (Batha & Carroll, 2007; Flavell, 1981; Weinert, 1987). Among other important aspects of metacognition are knowledge of one's own strengths and weaknesses and acceptance of responsibility for one's own learning.

• *Habits of thought – thoughtful habits*. Often poor performance on cognitively challenging tasks is due to inattentiveness, carelessness, or failure to check one's work. Hasty and careless reading of instructions can result in misunderstanding of the problem(s) one is trying to solve. Mechanical application of problem-solving procedures or failure to check the results of one's work can yield nonsensical "solutions." I am not aware of data-based estimates of the percentage of errors that are made on ability or achievement tests that are due to carelessness and that could be avoided by reflection, but I suspect that it is not negligible.

• *Attitudes and beliefs conducive to learning and thinking.* Fostering an attitude of carefulness and reflectiveness regarding one's work has been promoted as an eminently worthwhile goal (Ennis, 1986; Resnick, 1987). Other attitudes the importance of which has been stressed include inquisitiveness (Dillon, 1988; Millar, 1992) and fair-mindedness (Baron, 1988). Beliefs about whether one has any control over the retention of skills, or the learning of new ones, during one's later years can help determine how well one does in this regard (Bandura, 1997; Seeman et al., 1996). An important caveat comes from Nisbett (2009) regarding the use of praise as a motivator when working with children: "It is probably a bad idea to praise children for being intelligent. Instead praise hard work, which is under their direct control" (p. 188). This strikes me as an example of what should be a general principle: Children (people of any age) should be praised or complimented for what they can control (effort, honesty, perseverance, good will) and not for what they cannot (inherited talent or wealth, good looks, height, ethnicity). Citing experimental results obtained by Mueller and Dweck (1998), Nisbett argues that praising children for being intelligent makes them focus on trying to show how smart they are, and that to do this they are likely to avoid tasks that could prove to be too difficult and make them appear not so intelligent.

• *Other*. This list of things that can be taught in the interest of enhancing cognitive performance could easily be extended to include *principles of good reasoning,* *outlooks that motivate effort* (seeing the world as an incredibly interesting place and learning as not only important for practical reasons but intrinsically rewarding), *counterfactual thinking* (the usefulness of imagining alternative possibilities), *perspective taking* (looking at things from different points of view), and numerous other principles, practices, and perspectives that are conducive to a thoughtful approach to problems and life more generally.

Food for Thought

Information technology has dramatically changed the ways in which people access knowledge, communicate with each other, and perform their daily tasks; and it promises to provide even more powerful tools in the future. Does the technology that already is widely accessible act as an equalizer, decreasing the magnitude of differences in the ability of less- and more-intelligent people to perform cognitively demanding tasks (Hansen, Heckman, & Mullen, 2004), or does it act as an amplifier, increasing the magnitude of the differences (Stanovich, 1986)? That both possibilities are recognized in the literature should motivate more thought and discussion regarding what the goals of increasing intelligence should be.

As information technology develops increasingly versatile tools to facilitate performance of more and more of the cognitively demanding tasks that are performed by people – physicians, lawyers, mathematicians, scientists, truck drivers . . . – what it means to be intelligent, or to act intelligently, seems likely to have to take account of the effectiveness with which one can use, or interact with, these tools. One can imagine the possibility of an IQ test to quantify the intelligence of a person-tools combination.

What will be the effect of learning one day, as I expect we will in the not-distant future, that computers readily score (much) higher than any intelligent human on any test that is generally used to measure IQ? How will this eventuality affect our ideas about what intelligence is and how it should be measured and taught?

Summing Up

Whether intelligence can be taught is a considerably more difficult question to answer than one might assume. The question has motivated much research that has yielded many findings, including, in some cases, findings that support conflicting conclusions. I think it is safe to say that most researchers who have studied the question agree that the performance of specific tasks of the sort that typically appear on tests of intelligence can be taught effectively, which is to say that people's ability to perform such tasks can be improved, at least by modest amounts, by training. Citing numerous reviews of work on this subject, Salthouse (2015) concludes that "although reviewers differ in their estimates of the magnitude of the intervention effects, there is a consensus that cognitive interventions can be effective in increasing the level of performance in the trained tasks" (p. 86). Evidence that what is learned in such interventions improves performance on cognitively demanding tasks more generally is mixed. In sum, the extent to which intelligence, however defined, can be increased through instruction, and in such a way that the benefit lasts, remains a challenge to research. However, the relevant literature provides reasons for optimism that the quest is not quixotic.

References

Ackerman, P. L. (2014). Nonsense, common sense, and science of expert performance: Talent and individual differences. *Intelligence*, *45*, 6–17.

Adams, M. J. (1989). Thinking skills curricula: Their promise and progress. *Educational Psychologist*, *24*, 25–77.

Allensworth, E. M., Moore, P. T., Sartain, L., & de la Torre, M. (2017). The educational benefits of attending higher performing schools: Evidence from Chicago high schools. *Educational Evaluation and Policy Analysis*, *39*, 175–197.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.). New York: John Wiley & Sons.

Anderson, S. B., & Messick, S. (1974). Social competency in young children. *Developmental Psychology*, *10*, 282–293.

Andrews, G. R., & Debus, R. I. (1978). Persistence and the causal perception of failure: Modifying cognitive attributions. *Journal of Educational Psychology*, *70*, 154–166.

Bandura, A. (1997). *Self efficacy: The exercise of control*. New York: Freeman.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *Future of Children*, *5*(3), 25–50.

Barnett, W. S. (2002). *The battle over Head Start: What the research shows*. New Brunswick, NJ: National Institute for Early Education Research.

Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science*, *333*, 975–978.

Baron, J. (1988). *Thinking and deciding*. New York: Cambridge University Press.

Baron, J. (1991). Beliefs about thinking. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 169–186). Hillsdale, NJ: Erlbaum.

Barreto, F. B., de Miguel, M. S., Ibarluzea, J., Andiarena, A., & Arranz, E. (2017). Family context and cognitive development in early childhood: A longitudinal study. *Intelligence*, *65*, 11–22.

Batha, K., & Carroll, M. (2007). Metacognitive training aids decision making. *Australian Journal of Psychology*, *59*, 64–69.

Beaver, K. M., Schwartz, J. A., Nedelec, J. L., Connolly, E. J., Boutwell, B. B., & Barnes, J. C. (2013). Intelligence is associated with criminal justice processing: Arrest through incarceration. *Intelligence*, *41*, 277–288.

Beyth-Marom, R., Fischhoff, B., Quadrel, M. J., & Furby, L. (1991). Teaching adolescents decision making: A critical review. In J. Baron & R. V. Brown (Eds.), *Teaching decision making to adolescents* (pp. 19–59). Hillsdale, NJ: Erlbaum.

Boccio, C. M., Beaver, K. M., & Schwartz, J. A. (2018). The role of verbal intelligence in becoming a successful criminal: Results from a longitudinal sample. *Intelligence*, *66*, 24–31.

Botvinik, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, *66*, 83–113.

Bouchard, J., & Villeda, S. A. (2014). Aging and brain rejuvenation as systemic events. *Journal of Neurochemistry*, *132*, 5–19.

Bradway, K. P., Thompson, C. W., & Cravens, R. B. (1958). Preschool IQs after twenty-five years. *Journal of Educational Psychology*, *49*, 278–281.

Bransford, J. D., & Stein, B. S. (1984). *The ideal problem solver: A guide for improving thinking, learning, and creativity*. New York: Freeman.

Brinch, C. N., Galloway, T. A. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences*, *109*, 425–430.

Brody, N. (2014). A plea for the teaching of intelligence: Personal reflections. *Intelligence*, *42*, 136–141.

Burchinal, M., Lee, M., & Ramey, C. T. (1989). Type of daycare and preschool intellectual development in disadvantaged children. *Child Development*, *60*, 128–137.

Burhan, N. A. S., Mohamad, M. R., Kurniawan, Y., & Sidek, A. H. (2014). National intelligence, basic human needs, and their effect on economic growth. *Intelligence*, *44*, 103–111.

Byun, S-y., & Park, H. (2012). The academic success of East Asian American youth: The role of shadow education. *Sociology of Education*, 85(1), 40–60.

Calvin, C. M., Batty, G. D., Der, G., Brett, C. E., Taylor, A., Pattie, A., et al. (2017). Childhood intelligence in relation to major causes of death in 68 year follow-up: Prospective population study. *British Medical Journal*, 357. https://doi.org/10.1136/bmj.j2708

Campbell, F. A., & Burchinal, M. R. (2008). Early childhood interventions: The Abecedarian Project. In P. C. Kyllonen, R. D. Roberts, & L. Stankov (Eds.), *Extending intelligence: Enhancement and new constructs* (pp. 61–84). New York: Erlbaum.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development*, *65*, 684–698.

Campbell, F. A., & Ramey, C. T. (1995). Cognitive and school outcomes for high-risk African-American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal*, *32*, 743–772.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, *6*, 42–57.

Caplan, N., Choy, M. H., & Whitmore, J. K. (1992). Indochinese refugee families and academic achievement. *Scientific American*, *266*(2), 36–42.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*, 153–193.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1–22.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Cattell, R. B., & Cattell, A. K. S. (1961). *Culture Fair Intelligence Test (Scale 2, Forms A & B)*. Champaign, IL: Institute for Personality and Ability Testing.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, *27*, 703–722.

Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence and income. *American Psychologist*, *52*, 1051–1058.

Chance, P. (1986). *Thinking in the classroom*. New York: Teachers College Press.

Charlton, B. (2009). Clever sillies: Why high IQ people tend to be deficient in common sense. *Medical Hypotheses*, *73*, 867–870.

Chen, C., & Stevenson, H. W. (1995). Motivation and mathematics achievement: A comparative study of Asian-American, Caucasian-American and East Asian high school students. *Child Development*, *66*, 1215–1234.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

Cho, S. H., te Nijenhuis, J., van Vianen, A. E., Kim, H.-B., & Lee, K. H. (2010). The relationship between diverse components of intelligence and creativity. *Journal of Creative Behavior*, *44*, 125–137.

Chooi, W. T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence*, *40*, 531–542.

Clarke, S. H., & Campbell, F. A. (1998). Can intervention early prevent crime later? The Abecedarian Project compared with other programs. *Early Childhood Research Quarterly*, *13*, 31–343.

Clouston, S. A., Kuh, D., Herd, P., Elliott, J., Richards, M., & Hofer, S. M. (2012). Benefits of educational attainment on adult fluid cognition: International evidence from three birth cohorts. *International Journal of Epidemiology*, *41*, 1729–1736.

Colombo, J. (1993). *Infant cognition: Predicting later intellectual functioning*. Newberry Park, CA: Sage.

Covington, M. V., Crutchfield, R. S., Davies, L., & Olton, R. M. (1974). *The productive thinking program: A course in learning to think*. Columbus, OH: Merrill.

Crane, J., & Barg, M. (2003). Do early childhood intervention programs really work? Coalition for Evidence-Based Policy. https://pdfs.semanticscholar.org/90f5/fe541b5641037518041e66fe6b5af6d4a51a.pdf

Daniel, V. (2016). Does the intelligence of populations determine the wealth of nations? *Journal of Socio-Economics*, *46*, 27–37.

Daugherty, A. M., Zwilling, C., Paul, E. J., Sherepa, N., Allen, C., Kramer, A. F., et al. (2018). Multi-modal fitness and cognitive training to enhance fluid intelligence. *Intelligence*, *66*, 32–43.

Davies, N. M., Dickson, M., Davey Smith, G., van den Berg, G. J., & Windmeijer, F. (2018). The causal effects of education on health outcomes in the UK Biobank. *Nature Human Behaviour*, *2*, 117–125.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford: Oxford University Press.

Deary, I. J. (2001). *Intelligence: A very short introduction*. Oxford: Oxford University Press.

Deary, I. J. (2012a). Intelligence. *Annual Review of Psychology*, *63*, 453–482.

Deary, I. J. (2012b). 125 years of intelligence in the American Journal of Psychology. *American Journal of Psychology*, *125*, 145–154.

Deary, I. J. (2013). Intelligence. *Current Biology*, *23*, 673–676.

Deary, I. J. (2014). Teaching intelligence. *Intelligence*, *42*, 142–147.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13–21.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Detterman, D. K. (2014a). Introduction to the intelligence special issue on the development of expertise: Is ability necessary? *Intelligence*, *45*, 1–5.

Detterman, D. K. (2014b). You should be teaching intelligence! *Intelligence*, *42*, 148–151.
Detterman, D. K., & Sternberg, R. J. (Eds.) (1982). How and how much can intelligence be increased? Norwood, NJ: Erlbaum.
Dickstein, L. S. (1975). Effects of instructions and premise ordering errors in syllogistic reasoning. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 376–384.
Dillon, J. T. (1988). The remedial status of student questioning. *Journal of Curriculum Studies*, *20*, 197–210.
Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*, 939–944.
Duncan, J. Seitz, R., Kolodny, J., Bor, D., Herzog, H., & Ahmed, A. (2000). A neural basis for general intelligence. *Science*, *289*, 457–460.
Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development*. Philadelphia: Psychology Press.
Ehrenberg, S. D., & Ehrenberg, L. M. (1982). *BASICS: Building and applying strategies for intellectual competencies in students*. Coshocton, OH: Institute for Curriculum and Instruction.
Ennis, R. H. (1986). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. S. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). New York: Freeman.
Ericsson, K. A. (2007). Deliberate practice and the modifiability of body and mind: Toward a science of the structure and acquisition of expert and elite performance. *International Journal of Sport Psychology*, *38*, 4–34.
Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.
Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review*, *85*, 114–121.
Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences*. Hillsdale, NJ: Erlbaum.
Fagan, J. F. (2011). Intelligence in infancy. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence*. New York: Cambridge University Press.
Feuerstein, R., Rand, Y., Hoffman, M., & Miller, R. (1980). *Instrumental enrichment*. Baltimore, MD: University Park Press.
Flavell, J. H. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills*. New York: Academic Press.
Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29–51.
Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*, 171–191.
Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. New York: Cambridge University Press.
Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*, 235–292.
Gage, F. H. (2003). Brain, repair yourself. *Scientific American*, *289*(3), 46–53.
Gagne, R. M. (1967). *Science – A process approach: Purposes, accomplishments, expectations*. Washington: American Association for the Advancement of Science.
Gale, C. R., Batty, G. D., Osborn, D. P., Tynelius, P., Whitley, E., Rasmussen, F. (2012). Association of mental disorders in early adulthood and later psychiatric hospital

admissions and mortality in a cohort study of more than 1 million men. *Archives of General Psychiatry*, *69*, 823–831.

Garber, H. L. (1988). *The Milwaukee Project: Preventing mental retardation in children at risk*. Washington: American Association on Mental Retardation.

Gardner, H. (1995). “Expert performance: Its structure and acquisition”: Comment. *American Psychologist*, *50*, 802–803.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gardner, H., Krechevsky, M., Sternberg, R. J., & Okagaki, L. (1994). Intelligence in context: Enhancing students’ practical intelligence for school. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 105–127). Cambridge, MA: MIT Press.

Geary, D. C. (1996). Biology, culture, and cross-national differences in mathematical ability. In R. J. Sternberg & T. Ben-Zeev (Eds.), *The nature of mathematical thinking* (pp. 145–171). Mahwah, NJ: Erlbaum.

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, *24*, 79–132.

Gray, J. R., & Thompson, P. M. (2004). Neurobiology of intelligence: Science and ethics. *Nature Reviews Neuroscience*, 5, 471–482.

Greenwood, P. M. (2007). Functional plasticity in cognitive aging: Review and hypothesis. *Neuropsychology*, *21*, 657–673.

Greenwood, P. M., & Parasuraman, R. (2015). The mechanisms of far transfer from cognitive training: Review and hypothesis. *Neuropsychology*, *30*(6), 742–755.

Haier, R. J. (2014). The universe, dark matter, and streaming intelligence. *Intelligence*, *42*, 152–155.

Haimovitz, K., Wormington, S. V., & Corpus, J. H. (2011). Dangerous mindsets: How beliefs about intelligence predict motivational change. *Learning and individual differences*, *21*, 747–752.

Hambrick, D. Z., Oswald, F. L., Altman, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34–45.

Hammond, S. I., Müller, U., Carpendale, J. I., Bibok, M. B., & Liebermann-Finestone, D. P. (2012). The effects of parental scaffolding on preschoolers’ executive function. *Developmental Psychology*, *48*, 271–281.

Hansen, K. T., Heckman, J. J., Mullen, K. J. (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics*, *121*, 39–98.

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science*, *24*, 2409–2419.

Harvard University (1983). *Project Intelligence: The development of procedures to enhance thinking skills*. Final report, submitted to the Minister for the Development of Human Intelligence, Republic of Venezuela, October.

Hayes, T. R., Petrov, A., & Sederberg, P. B. (2015). Do we really become smarter when our fluid-intelligence test scores improve? *Intelligence*, *48*, 1–14.

Herrnstein, R. J., Nickerson, R. S., Sanchez, M, & Swets, J. A. (1986). Teaching thinking skills. *American Psychologist*, *41*, 1279–1289.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberg, U. (2009). Enrichment effects on adult cognitive development. *Psychological Science in the Public Interest*, *9*, 1–65.

Heyman, G. D., & Dweck, C. S. (1998). Children's thinking about traits: Implications for judgments of the self and others. *Child Development*, *64*, 391–403.

Hong, Y. Y., Chiu, C., Dweck, C. S., Lin, D., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*, 588–599.

Honzik, M. P., Macfarlane, J. W., & Allen, L. (1948). The stability of mental test performance between two and eighteen years. *Journal of Experimental Education*, *17*, 309–324.

Horacek, H. J., Ramey, C. T., Campbell, F. A., Hoffmann, K., & Fletcher, R. H. (1987). Predicting school failure and assessing early intervention with high-risk children. *American Academy of Child and Adolescent Psychiatry*, *26*(1987), 758–763.

Hultsch, D. F., Hertzog, C., Small, B. J., & Dixon, R. A. (1999). Use it or lose it: Engaged lifestyle as a buffer of cognitive decline in aging? *Psychology of Aging*, *14*, 245–263.

Hunt, E. (2011). *Human intelligence*. Cambridge, UK: Cambridge University Press.

Hunt, E. (2014). Teaching intelligence: Why, why it is hard and perhaps how to do it. *Intelligence*, *42*, 156–165.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, *29*, 340–362.

Jak, A. J. (2011). The impact of physical and mental activity on cognitive aging. In M.-C. Pardon & M. W. Bondi (Eds.), *Behavioral neurobiology of aging* (pp. 273–291). Berlin: Springer.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, *105*, 6829–6833.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training and transfer. *Memory and Cognition*, *42*, 464–480.

Jauk, E., Benedek, M., Dunst, B., & Neubauer, A. C. (2013). The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. *Intelligence*, *41*, 212–221.

Jensen, A. R. (1998). *The g factor*. Westport, CT: Praeger.

Jones, B. F., Palincsar, A. S., Ogle, D. S., & Carr, E. G. (1987). Learning and thinking. In B. F. Jones, A. S. Palincsar, D. S. Ogle, & E. G. Carr (Eds.), *Strategic teaching and learning: Cognitive instruction in the content areas* (pp. 3–32). Alexandria, VA: Association for Supervision and Curriculum Development.

Jones, G., & Schneider, W. (2006). Intelligence, human capital, and economic growth: A Bayesian averaging of classical estimates (BACE) approach. *Journal of Economic Growth*, *1*, 71–93.

Kanyama, I. K. (2014). Quality of institutions: Does intelligence matter? *Intelligence*, *42*, 44–52.

Kaufman, J. C., & Plucker, J. A. (2011). Intelligence and creativity. In R. J. Sternberg, & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 771–783). Cambridge, UK: Cambridge University Press.

Keith, Z. K., & Reynolds, M. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools*, *47*, 635–650.

Kim, K. H. (2005). Can only intelligent people be creative? *Journal of Secondary Gifted Education*, *24*, 57–66.

Klausmeier, H. J. (1980). *Learning and teaching concepts – A strategy for testing applications of theory*. New York: Academic Press.

Klemp, G. O., Jr., & McClelland, D. C. (1986). What characterizes intelligent functioning among senior managers? In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 31–50). Cambridge, UK: Cambridge University Press.

Kosonen, P., & Winne, P. H. (1995). Effects of teaching statistical laws on reasoning about everyday problems. *Journal of Educational Psychology*, *87*, 33–46.

Krueger, J. (2000). Individual differences and Pearson's r: Rationality revealed? *Behavioral and Brain Sciences*, *23*, 684–685.

Lally, J. R., Mangione, P. L., & Honig, A. S. (1998). The Syracuse University Family Development Research Project: Long-range impact of an early intervention with low-income children and their families. In D. R. Powell (Ed.), *Annual advances in applied developmental psychology*, vol. 3, *Parent education as early childhood intervention: Emerging directions in theory, research, and practice*. Norwood, NJ: Ablex.

Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*, 1335–1342.

Lazar, I., & Darlington, R. (1982). Lasting effects of early education: A report from the Consortium for Longitudinal Studies. *Monographs of the Society for Research in Child Development*, 47(2–3), 1–151.

Lee, V. E., & Loeb, S. (1994). *Where do Head Start attendees end up? One reason why preschool effects fade out*. ED368510. Available from the Education Resources Information Center (ERIC).

Lewis, M. (1973). Infant intelligence tests: Their use and misuse. *Human Development*, *16*, 108–118.

Lynn, R. (2012). IQs predict differences in the technological development of nations from 1000 BC through 2000 AD. *Intelligence*, *40*, 439–444.

Lynn, R., & Vanhanen, T. (2012). National IQs: A review of their educational, cognitive, economic, political, demographic, sociological, epidemiological, geographic and climatic correlates. *Intelligence*, *2*, 226–234.

Machado, L. A. (1980). The right to be intelligent. New York: Pergamon Press.

Mackintosh, N. J. (2011). *IQ and human intelligence* (2nd ed.). Oxford: Oxford University Press.

Mackintosh, N. J. (2014). Why teach intelligence? *Intelligence*, *42*, 166–170.

Manuel, H. T. (1962). *Tests of general ability: Inter-American series (Spanish, Level 4, Forms A & B)*. San Antonio, TX: Guidance Testing Associates.

Martin, S. L., Ramey, C. T., & Ramey, S. (1990). The prevention of intellectual impairment in children of impoverished families: Findings of a randomized trial of educational day care. *American Journal of Public Health*, *80*, 844–847.

McCall, R. B., Appelbaum, M. I., & Hogarty, P. S. (1973). Developmental changes in mental performance. *Monographs of the Society for Research in Child Development*, 42(3), 1–84).

McKay, R. H., Condelli, L., Ganson, H., Barnett, B. J., McCouley, C., & Plantz. M. C. (1985). *The impact of Head Start on children, families and communities: Final report of the Head Start Evaluation, Synthesis and Utilization Project*. Washington: US Department of Health and Human Services.

Meeker, M. N. (1969). *The structure of intellect: Its interpretation and uses*. Columbus, OH: Charles E. Merrill.

Meisenberg, G., & Lynn, R. (2011). Intelligence: A measure of human capital in nations. *Journal of Social, Political and Economic Studies*, *36*, 421–454.

Melby-Lervåg̊, M., & Hulme, C.(2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*, 270–291.

Millar, G. (1992). *Developing student questioning skills – A handbook of tips and strategies for teachers*. Bensenville, IL: Scholastic Testing Service.

Mueller, C. W., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*, 33–52.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. www.edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf

Neisser, U. (1997). Rising scores on intelligence tests. *American Scientist*, *85*, 440–447.

Neisser, U. (Ed.) (1998). *The rising curve: Long-term gains in IQ and related measures*. Washington: American Psychological Association.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., & Ceci, S. J. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Nickerson, R. S. (1980). Retrieval efficiency, knowledge assessment and age: Comments on some welcome findings. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging: Proceedings of the George A. Talland Memorial Conference* (pp. 355–366). Hillsdale, NJ: Erlbaum.

Nickerson, R. S. (1987). Project Intelligence: An Account and some reflections. *Special Services in the Schools*, 3(1–2), 83–102.

Nickerson, R. S. (1988/1989). On improving thinking through instruction. In E. Z. Rothkopf (Ed.), *Review of research in education* (vol. 15, pp. 3–58). Washington: American Educational Research Association.

Nickerson, R. S. (1994). The teaching of thinking and problem solving. In R. J. Sternberg (Ed.), *Thinking and problem solving*, vol. 12, *Handbook of perception and cognition* (pp. 409–449). San Diego, CA: Academic Press.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.

Nickerson, R. S. (2004). Teaching reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 410–442). New York: Cambridge University Press.

Nickerson, R. S. (2011). Developing intelligence through instruction. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 107–129). New York: Cambridge University Press.

Nickerson, R. S., Butler, S. F., & Barch, D. H. (in press). Validity and persuasiveness of conditional arguments. *American Journal of Psychology.*

Nickerson, R. S., Perkins, D. N., & Smith, E. E. (1985). *The teaching of thinking*. Hillsdale, NJ: Erlbaum.

Nisbett, R. E. (2009). *Intelligence and how to get it: Why schools and cultures count*. New York: W. W. Norton.

Nottebohm, F. (2002). Why are some neurons replaced in adult brains? *Journal of Neuroscience*, *22*, 624–628.

Nusbaum, E. C., & Silvia, P. (2011). Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence*, *39*, 36–45.

Olds, D., Henderson, C. R., Jr., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., et al. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior: 15-year follow-up of a randomized trial. *Journal of the American Medical Association*, 280(14), 1238–1244.

Otis, A. S., & Lennon, R. T. (1977). *Otis-Lennon School Ability Test (Intermediate Level 1, Form R)*. New York: Harcourt Brace Jovanovich.

Ott, A., van Rossum, C. T., van Harskamp, F., van de Mheen, H., Hofman, A., & Breteler, M. M. (1999). Education and the incidence of dementia in a large population-based study: The Rotterdam study. *Neurology*, *52*, 663–666.

Papageorgiou, E., Christou, C., Spanoudis, G., & Demetriou, A. (2016). Augmenting intelligence: Developmental limits to learning-based cognitive change. *Intelligence*, *56*, 16–27.

Pardon, M.-C., & Bondi, M. W. (Eds.) (2011) *Behavioral neurobiology of aging*. Berlin: Springer.

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, *8*, 293–316.

Park, D. C., & Reuter-Lorenz, P. (2009). The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*, *60*, 173–196.

Payne, J. E., Mercer, C. D., Payne, A., & Davison, R. G. (1973). *Head Start: A tragicomedy with epilogue*. New York: Behavioral Publications.

Perkins, D. N. (1995). *Outsmarting IQ: The emerging science of learnable intelligence*. New York: Free Press.

Piattelli-Palmarini, M. (1994). *Inevitable illusions: How mistakes of reason rule our minds*. New York: Wiley.

Plomin, R., & DeFries, J. C. (1998). Genetics of cognitive abilities and disabilities. *Scientific American*, 275(5), 62–69.

Plomin, R., Pedersen, N. L., Lichtenstein, P., & McClearn, G. E. (1994). Variability and stability in cognitive abilities are largely genetic later in life. *Behavior Genetics*, *24*, 207–215.

Ramey, C. T., Bryant, D. M., & Suarez, T. M. (1985). Preschool compensatory education and the modifiability of intelligence: A critical review. In D. K. Detterman (Ed.), *Current topics in human intelligence*, vol. 1, *Research methodology* (pp. 247–296). Westport, CT: Ablex Publishing.

Ramey, C. T., & Campbell, F. A. (1984). Preventive education for high-risk children: Cognitive consequences of the Carolina abecedarian project. *American Journal of Mental Deficiency*, *88*, 515–523.

Ramey, C. T., & Campbell, F. A. (1994). Poverty, early childhood education, and academic competence: The Abecedarian experiment. In A. C. Huston (Ed.), *Children in poverty: Child development and public policy* (pp. 190–221). New York: Cambridge University Press.

Reynolds, A. J. (1998). The Chicago Child-Parent Center and Expansion Program: A study of extended early childhood intervention. In Jonathan Crane (Ed.), *Social programs that work* (pp. 110–147). New York: Russell Sage Foundation.

Resnick, L. B. (1987). *Education and learning to think*. Washington: National Academy Press.

Rindermann, H., & Becker, D. (2018). Flynn-effect and economic growth: Do national increases in intelligence lead to increases in GDP?. *Intelligence*, *69*, 87–93.

Rips, L. J., & Conrad, F. G. (1983). Individual differences in deduction. *Cognition and Brain Theory*, *6*, 259–285.

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science*, *29*, 1358–1369.

Ross, L., Greene, D., & House, P. (1977). The false consensus phenomenon: An attributional bias in self-perception and social perception processes. *Journal of Experimental Social Psychology*, *13*, 279–301.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118–137.

Rubenstein, M. F. (1975). *Patterns of problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Salthouse, T. A. (2015). Do cognitive interventions alter the rate of age-related cognitive change? *Intelligence*, *53*, 86–91.

Sanders, J. R., & Sonnad, S. R. (1982). *Research on the introduction, use and impact of the* ThinkAbout *instructional television series: Executive summary*. Agency for Instructional Television, January.

Scarmeas, N., Levy, G., Tang, M. X., Manly, J., Stern, Y. (2001). Influence of leisure activity on the incidence of Alzheimer's disease. *Neurology*, *57*, 2236–2242.

Schipolowski, S., Wilhelm, O., & Schroeders, U. (2014). On the nature of crystalized intelligence: The relationship between verbal ability and factual knowledge. *Intelligence*, *46*, 156–168.

Schneider, W., Niklas, F., & Schmiedeler, S. (2014). Intellectual development from early childhood to early adulthood: The impact of early IQ differences on stability and change over time. *Learning and Individual Differences*, *32*, 156–162.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academic Press.

Schwartzman, A. E., Gold, D., Andres, D., Arbuckle, T. Y., & Chaikelson, J. (1987). Stability of intelligence: A 40-year follow-up. *Canadian Journal of Psychology*, 41, 244–256.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool Study through age 40*. Ypsilanti, MI: High/Scope Foundation.

Seeman, T. E., McAvay, G., Merrill, S., Albert, M., & Rodin, J. (1996). Self-efficacy beliefs and changes in cognitive performance: MacArthur studies of successful aging. *Psychology and Aging*, *11*, 538–551.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628–654.

Sigel, I. E. (1973). Where is preschool education going: Or are we en route without a road map? *Proceedings of the 1972 Invitational Conference on Testing Problems: Assessment in a pluralistic society* (pp. 99–116). Princeton: Educational Testing Service.

Squalli, J., & Wilson, K. (2014). Intelligence, creativity, and innovation. *Intelligence*, *46*, 250–257.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.

Stanovich, K. E. (1994). Reconceptualizing intelligence: Dysrationalia as an intuition pump. *Educational Researcher*, *23*, 11–22.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.

Stanovich, K. E., & West, R. (2008). On the failure of cognitive ability to predict myside bias and one-sided thinking biases. *Thinking and Reasoning*, *14*, 129–167.

Sternberg, R. J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills*. San Diego, CA: Harcourt Brace Jovanovich.

Sternberg, R. J. (2012). Intelligence. *Dialogs in Clinical Neuroscience*, *14*, 14–27.

Sternberg, R. J. (2014). Teaching about the nature of intelligence. *Intelligence*, *42*, 176–179.

Sternberg, R. J., & Grigorenko, E. L. (2007). *Teaching for successful intelligence* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Sternberg, R. J., Jarvin, L., & Grigorenko, E. L. (2011). *Explorations of the nature of giftedness*. New York: Cambridge University Press.

Sternberg, R. J., & Wagner, R. K. (Eds.) (1986). *Practical intelligence: Nature and origins of competence in the everyday world*. Cambridge, UK: Cambridge University Press.

Stevenson. H. W., Chen, C., & Lee, S-Y. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science*, *259*, 53–58.

Stevenson, H. W., Lee, S. Y., & Stigler, J. W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science*, 231, 693–699.

Swartz, R. J. (1991). Structured teaching for critical thinking and reasoning in standard subject area instruction. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 415–450). Hillsdale, NJ: Erlbaum.

Swets, J. A., Herrnstein, R. J., Nickerson, R. S., & Getty, D. J. (1988). Design and evaluation issues in an experiment on teaching thinking skills. *American Psychologist*, *43*, 600–602.

Thorell, L. B., Lindqvist, S., Bergman, S., Bholin, G., & Klingberg, T. (2008). Training and transfer effects of executive functions in preschool children. *Developmental Science*, *11*, 969–976.

Tourva, A., Spanoudis, G., & Demetriou, A. (2016). Cognitive correlates of developing intelligence: The contribution of working memory, processing speed and attention. *Intelligence*, *54*, 136–146.

Tsang, S. L. (1988). The mathematics achievement characteristics of Asian-American students. In R. R. Cocking & J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics* (pp. 123–136). Hillsdale, NJ: Erlbaum.

Valenzuela, M. J., Breakspear, M., & Sackdev, P. (2007). Complex mental activity and the aging brain: Molecular, cellular and cortical network mechanisms. *Brain Research Reviews*, *56*, 198–213.

Valenzuela, M. J., Sachdev, P., Wen, W., Chen, X., & Brodaty, H. (2008). Lifespan mental activity predicts diminished rate of hippocampal atrophy. *PLoS One*. https://doi.org/10.1371/journal.pone.0002598

Verghese, J., Wang, C., Katz, M. J., Sanders, A., & Lipton, R. B. (2009). Leisure activities and risk of vascular cognitive impairment in older adults. *Journal of Geriatric Psychiatry and Neurology*, *22*, 110–118.

Verhaeghen, P., Cerella, J., & Basak, C. (2004). A working memory workout: How to expand the focus of serial attention from one to four items in 10 hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1322–1337.

Wai, J. (2013). Investigating America's elite: Cognitive ability, education, and sex differences. *Intelligence*, *41*, 203–211.

Wai, J. (2014). Investigating the world's rich and powerful. *Intelligence*, *46*, 54–72.

Wai, J., & Nisen, M. (2013). The 25 countries with the most brainpower. *Business Insider*. www.businessinsider.com/countries-with-the-most-brainpower-2013-10?r=US&IR=T

Wechsler, D. (1981). *WAIS-R manual: Wechsler Adult Intelligence Scale – Revised*. San Antonio, TX: Psychological Corporation.

Weinert, F. E. (1987). Introduction and overview: Metacognition and motivation as determinants of effective learning and understanding. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 1–16). Hillsdale, NJ: Erlbaum.

Wickelgren, W. A. (1974). *How to solve problems*. San Francisco: W. H. Freeman.

Williams, R. L. (2013). Overview of the Flynn effect. *Intelligence*, *41*, 753–764.

Wrar, C., Deary, I. J., Gale, C. R., & Der, G. (2015). Intelligence in youth and health at 50. *Intelligence*, *53*, 23–32.

Wrulich, M., Brunner, M., Stadler, G., Schalke, D., Keller, U., & Martin, R. (2014). Forty years on: Childhood intelligence predicts health in middle adulthood. *Health Psychology*, *33*, 292–296.

Yang, B., & Lester, D. (2016). Regional differences in intelligence and economic activity: A brief note. *Intelligence*, *54*, 33–36.

Yu, H., McCoach, D. B., Gottfried, A. W., & Gottfried, A. K. (2018). Stability of intelligence from infancy through adolescence: An autoregressive latent variable model. *Intelligence*, *69*, 8–15.

PART III

Intelligence and Group Differences

11 Intellectual Disability

Deborah J. Fidler, Emily Schworer, Molly Swanson, and Susan Hepburn

Intellectual Disability

Intellectual disability (ID) is a concept that is both socially constructed, and also rooted in objective, measurable factors. The socially constructed nature of ID is reflected in the definitional changes that have taken place over the past century, and the evolution of terminology used to refer to the phenomenon of substantially lower than average intellectual capacity. There are also social implications for a diagnosis of ID in community and legal contexts, and an ID diagnosis can have an impact on quality of life by establishing eligibility for support services and governmental benefits (Schalock et al., 2010). And yet, ID can also be rooted in biological etiological factors that lead to decreased intellectual capacity, factors that can often be objectively defined. In this way, the concept of ID involves an important intersection between cultural constructs, biological mechanisms, and the sociohistorical trends that interact with these two dimensions.

In the context of an edited volume on intelligence, ID could simply be construed as globally decreased cognitive capacity. However, the technical definitions of ID that are endorsed by prominent organizations such as the American Association on Intellectual and Developmental Disabilities (AAIDD) have expanded to encompass many other aspects of functioning beyond performance on measures of intelligence. Since the 1960s, the definition of ID has involved a "dual-criterion approach," including both compromised intellectual functioning and a profile of difficulties in other areas of adjustment (Schalock et al., 2010).

According to the AAIDD definition published in 2010, ID is a "disability characterized by significant limitations in both intellectual functioning and in adaptive behavior" that originates before the age of eighteen (Schalock et al., 2010). Within this definition, the AAIDD specifies that "intellectual functioning" or "intelligence" refers to more general mental capacity, and specific areas that include "reasoning, planning, solving problems, thinking abstractly, comprehending complex ideas, learning quickly, and learning from experience" (p. 15). What is notable, however, about this definition (and others) is that diminished capacity in the areas of learning and reasoning are not sufficient to meet criteria for ID according to AAIDD. Instead, the dual-criterion approach specifies that concurrent difficulties with adaptive behavior, including practical skills (activities of daily living), conceptual skills (language and literacy), and social skills (interpersonal skills), are a core part of the definition as well.

Thus, an important distinction should be made between referring narrowly to intellectual impairments and the more formal, comprehensive definition of ID. In this chapter, we will explore the construct of ID and its evolution. We will examine the interaction between biological, social, cultural, and historical factors in the seemingly objective construct of ID. We then discuss issues related to ID comorbidities and current best practices in ID assessment. We conclude with a discussion regarding issues in ID science and practice that are on the horizon, and that will prompt the next set of changes in the evolution of this dynamic construct.

Terminology and Definitions

Terminology

One essential aspect of the socially constructed nature of ID relates to terminology. The term "ID" has had many prior incarnations, most of which are thought to be not only out of vogue, but offensive to current sensibilities. Until the mid-2000s, ID was commonly referred to as "mental retardation," especially in the United States. This term was resoundingly rejected when, in 2007, the American Association on Mental Retardation (AAMR) officially changed its name to the American Association on Intellectual and Developmental Disabilities. This change in terminology was accompanied in the United States by the passing of "Rosa's Law" (PL 111–156) in 2010, legislation that required the use of the term "ID" in legal contexts. However, it is notable that similar terminology changes had taken place repeatedly throughout the past 100 years, for example, when the term "mental deficiency" that was used in the 1950s (American Psychiatric Association [APA] Committee on Nomenclature and Statistics, 1952), was rejected in favor of "mental retardation" in the late 1960s (APA, 1968). These changes were preceded by rejections of what are now considered to be socially unacceptable terms, but at the time were included as technical jargon to connote varying levels of developmental functioning (Biasini et al., 1999). Even today, ID terminology differences continue to exist, as the World Health Organization (WHO) has opted to use the term "Disorders of Intellectual Development" (International Classification of Diseases 11th Revision [ICD-11]; WHO, 2018), and the term "learning disability" is sometimes used in place of ID in some areas within the UK.

Beyond changes in terminology, the criteria for ID have also evolved historically, reflecting social changes within the social sciences and allied health fields. The intersection between social factors and the construct of ID can be observed in the designated IQ boundaries for the definition of ID over the past century. Since the 1920s, the notion of ID has incorporated IQ-based specifications, with an IQ in the area of 70–75 often serving as the cut-off score. In one interesting chapter in ID history during the 1960s, however, the (then called) American Association on Mental Deficiency published a definition update that raised the IQ cut-off point to

85, a level that was also endorsed by the APA at the time. While the intent of this change may have reflected a larger secular move in the 1960s to improve access to services for a wider group of individuals, this change also allowed for approximately 16 percent of the population to have the designation of ID, and included a large proportion of minority groups.

These historical fluctuations reflect several issues. First, they exemplify the arbitrary nature of boundaries and cut-off points in diagnostics, and how bound clinical and technical definitions are with sociohistorical climate. They demonstrate how consequential these arbitrary decisions may be, as marked changes in ID prevalence carry with them enormous social and economic implications. Moreover, they represent the dual-edge nature of a diagnosis of ID, as this designation may bring about benefits such as eligibility for support services and entitlements, but may also bring with it stigma and marginalization.

Current Definitions

The most widely adopted definitions of ID have been produced by three different organizations, including the AAIDD, the WHO, and the APA. At present, each of these definitions includes both intellectual and adaptive functioning elements, with a boundary of about two standard deviations below the mean. However, they vary from one another in subtle ways, which is likely the result of the different professional constituencies that each organization serves.

In the ICD-11, "disorder of intellectual development" is described in a straightforward way, encompassing "significantly below average intellectual functioning" and challenges in adaptive behavior that originate in the developmental period. In contrast, the 2010 AAIDD definition (Schalock et al., 2010) provides a substantial framework for its definition, based on a social-ecological approach that emphasizes the interaction between an individual and the contextual factors in their environment. The definition also emphasizes that the rationale for describing limitations should be for the purpose of providing appropriate supports, and not to marginalize an individual. Several assumptions are established in the opening section of the definition volume, with explicit statements made regarding the importance of understanding intellectual limitations in cultural and developmental context and accounting for linguistic diversity when conducting assessments. The definition, therefore, includes recommendations that the normative cultural environment and linguistic diversity be accounted for when considering a diagnosis of ID. This consideration is important in that concerns have been raised regarding the over-representation of individuals from diverse backgrounds with an ID diagnosis (Fujiura, Yamaki, & Czechowsicz, 1998) and within special education settings (Morgan et al., 2015). Beyond the issue of cultural and linguistic diversity, the AAIDD has several other unique features, including the recognition and valuing of coexisting strengths in the context of ID, which reflects a priority regarding understanding ID as one aspect of an individual's identity, not the entirety of who they are.

In contrast, the APA definition, which is included in the Diagnostic and Statistical Manual, Fifth Edition (DSM-V) section on neurodevelopmental disorders, highlights the developmental nature of ID and places it in a similar category to other disorders that involve atypical early development (APA, 2013). To reflect the developmental emphasis, the APA has taken a more conservative approach to the diagnosis of ID with the addition of a diagnostic category of "global developmental delay." As noted by Brue and Wilmshurst (2016), this category is useful during the very early childhood years, when a child may be evidencing some degree of cognitive delay, but a diagnosis of ID is not yet appropriate.

In comparing the three current major ID definitions, it is clear that they have converged on the idea of the importance of both intellectual and adaptive functioning, and all three definitions discuss an age of onset during the childhood years. There is a notable trend toward focusing on the interaction between an individual and their social, cultural, and interpersonal context. However, each definition is couched within a specific framework, and the variations around these definitions and the amount of guidance provided are a reflection of the authoring organizations and the constituencies that they serve.

Classification of ID

Another intersection between social and biological factors can be observed in the approaches that various professional communities have taken to the classification of ID. The population of individuals with ID is heterogeneous, and a wide range of profiles in social, communication, behavioral, and motoric abilities can be observed in concert with cognitive and adaptive impairments. As a result, one question that has been asked repeatedly in the field of ID has been how to identify subgroups of individuals with ID in order to provide additional information and context regarding an individual's profile beyond the presence of ID. One longstanding approach with roots in the early 1900s involves characterizing the severity of impairment with IQ ranges. For much of the twentieth century, various organizations took an approach wherein an IQ of 55–70 was considered mild ID, 40–55 was designated as moderate ID, 25–40 was designated as severe, and scores below an IQ of 25 were designated as profound ID. This classification system, which came to be seen as too reductionistic, was gradually abandoned by the professional organizations who establish ID definitions – first by the AAIDD definition and then by the APA.

Adaptive behavior continua

In place of classification systems based on IQ, an alternative approach took hold in the 1990s, which involved classifying individuals with ID by the degree of support services they require for participation. This is reflected in more recent definition frameworks from the APA and the AAIDD. In the most recent APA definition of ID (APA, 2013), the classification of mild, moderate, severe, and profound ID is based on adaptive behavior criteria in the areas of conceptual, social, and practical skills. Within this system, someone with mild ID would be expected to "function age-appropriately in personal care" and "need some support with complex daily living tasks in comparison to peers" (p. 34). In contrast, someone with severe ID would require "support for all activities of daily living, including meals, dressing, bathing, and elimination. The individual requires supervision at all times" (p. 36) in the domain of practical skills. This change from classification based on the intellectual capacity dimension to the adaptive behavior dimension marks a recognition that the degree of intellectual impairment may not be an accurate or useful indicator of an individual's overall level of need for support. Using adaptive behavior in the place of intellectual functioning to categorize severity of impairment allows clinicians, the primary users of the DSM-V, to make direct translation to practice in terms of support plans and services needed.

Multidimensional approach

Another novel approach to classification was presented in the 2010 AAIDD definition volume, which employed a "multidimensional classification system" that involved the dimensions of intellectual ability, adaptive behavior, health, participation, and context. Though the intellectual and adaptive behavior components are straightforward, the additional emphasis on health in this classification system accounts for physical and mental health, and ID-etiological factors as well. The dimension of participation relates to the degree of involvement an individual has in their community contexts, including work, education, cultural, avocational, and religious settings. The context element in this classification system allows an account of the environmental and personal influences on an individual with ID. The addition of these elements to an ID classification system is thought to increase the utility of the ID diagnosis and to reduce the stigma associated with ID. Additional specific guidance regarding classification is provided in the 2010 AAIDD definition, which recommends the use of a nominal classification when that information is sufficient, and the use of a severity continuum only when there is a strong rationale and when doing so will affect outcomes. Additional recommendations include avoiding the use of age-referenced scores when possible in order to reduce the stigma of assigning younger developmental age equivalents to individuals with older chronological ages.

Two-group approach

Each of the classification approaches described above has had utility for the constituencies served by the professional organizations that publish definitions and classification systems. From a clinical and treatment standpoint, these approaches have provided guidance regarding supports and services and have reframed the nature of characterizing variability among individuals with ID. At the same time that these systems have become more advanced, yet another classification has been adopted by the scientific community over the past several decades that has been a catalyst for rapid growth in the field of ID research. This approach, which focuses on classifying individuals based on the underlying etiology that causes ID, has origins in the 1800s but grew in its importance in the last few decades of the 1900s. The most influential framework for this classification approach was put forth by Zigler (1967), who theorized that there were two distinct groups associated with ID, an "organic" ID group and a "cultural-familial" ID group. The main distinction between those groups related to the underlying causal factors of the ID diagnosis.

Within this two-group approach, the "cultural-familial" group was understood as representing the lower range of a normal curve distribution of intelligence (Zigler, 1967; Zigler & Hodapp, 1986). Because of this, mild levels of intellectual impairment predominate in this category, with IQ levels below 50 rarely observed. The causal factors contributing to the cultural-familial etiology of ID are generally undetectable (sometimes leading to the label of "idiopathic ID"), and are thought to involve a dynamic interplay between polygenic inheritance and contextual factors (for a summary, see Iarocci & Petrill, 2012). An ID diagnosis in this group tends to be identified in educational, rather than health care settings. Biomedical comorbidities and physical delays are infrequent in this group, and mortality rates are similar to those observed in the general population (Iarocci et al., 2012). Because of the polygenic nature of inheritance of ID, siblings of individuals in the cultural-familial ID group also tend to show lower intellectual functioning as well. An important sociocultural interplay is observed in this group between environmental, economic, and developmental factors. Questions regarding the link between mild ID and socioeconomic status (SES) have been raised for several decades, as research in the 1980s demonstrated a marked increase in the incidence of mild ID in individuals at the lowest SES levels (Broman et al., 1987).

The second group identified in Zigler's two-group approach, which generally comprised individuals with more pronounced intellectual impairments, included individuals with organic or biological causes of ID. A current exhaustive list of these causes can be found in the DSM-V, which identifies prenatal, perinatal, and postnatal etiologies of ID (APA, 2013). The prenatal category includes genetic syndromes, inborn errors of metabolism, brain malformations, and maternal disease, and environmental influences such as exposure to alcohol, drugs, and toxins (APA, 2013). Perinatal factors include disruption during labor and delivery. The postnatal category involves hypoxia-related issues, traumatic brain injury, infections, disorders of myelination, seizure disorders, deprivation, and exposure to toxic substances such as lead or mercury (APA, 2013).

Since Zigler's original theorizing regarding a two-group approach, there has been a marked expansion in developmental research on the organic ID group. In particular, an entirely new set of research questions were posed regarding outcome-specificity in neurogenetic syndromes (Hodapp, 1997), with a particular focus on cognitive and behavioral phenotypes. Along with the rapid increase in the number of genetic causes of ID that have been identified, this work has led to an expansion of the field of ID research, with federal funding initiatives targeting specific neurogenetic syndrome groups (e.g., the National Institutes of Health's Down syndrome INCLUDE project and Centers for Collaborative Research in Fragile X Syndrome) and notable rise in the number of publications that focus on syndrome-specific outcomes (Hodapp & Dykens, 2009). With respect to intellectual functioning, work in this area has demonstrated the complexity of cognitive functions, with unique profiles of attention, memory, executive function, and other essential aspects of cognition that present uniquely in different disorders.

An etiology-based approach to ID was initially embraced by the research community, while greater hesitancy came from the clinical and educational communities. For researchers, studying varying outcomes associated with different neurogenetic syndromes made it possible to ask questions regarding the connections between genetic factors and behavioral outcomes in ways that were impossible using an IQ- or adaptive behavior–continuum approach to classification. Examining etiology-related differences has also allowed researchers to recognize that ID occurs in the context of nuanced profiles of relative strengths and challenges in a variety of developmental domains, including language and communication, social relatedness, emotion and behavior regulation, and motor development, and that these profiles are more likely to occur in individuals with a particular neurogenetic condition. However, the utility of an etiology-based classification system was unclear to some in practice communities, with initial concerns raised regarding the potential "balkanization" of special education (Forness & Kavale, 1994), and questions posed regarding how to translate this information into interventions that would serve lower incidence groups. Over time, an appreciation of the possibilities afforded by an etiology-based approach has begun to take hold, with numerous syndrome-specific educational interventions developed that target phenotypic profiles in Down syndrome (Bennett, Holmes, & Buckley, 2013; Lemons et al., 2015; Pulina et al., 2015), fragile X syndrome (McDuffie et al., 2016), and other neurogenetic disorders (Singh et al., 2017). There is also promise that in characterizing phenotypic profiles carefully in adolescence and adulthood, it may be possible to investigate then whether early manifestations of those outcomes are identifiable earlier in childhood. Such an approach may make it possible to target syndrome-specific profiles early in development, and offer opportunities to strengthen cognitive and other developmental foundations with cascading effects downstream.

Prevalence

Just as the construct of ID is yoked to sociohistorical factors in terms of definition and terminology, social-contextual factors also play a role in estimating the prevalence of ID. According to one recent study, approximately 1.2 percent of the US population has a diagnosis of ID (Maenner et al., 2016). This estimate is based on caregiver responses to questions on the National Survey of Children's Health and the National Health Information Survey (NHIS) between 2011 and 2013. The 1.2 percent prevalence number is higher than previous estimates that were generated from the NHIS between 2006 and 2008, which placed the prevalence of ID at approximately 0.7 percent (Boyle et al., 2011). However, Maenner et al., note that these two timeframes of data collection overlap with the terminology change that took place in the mid-2000s. Caregivers in Boyle and colleagues' (2011) study were asked to respond to survey questions that contained the term "mental retardation," but the caregivers in Maenner and colleagues' (2016) study responded to questions that included the ID terminology. Maenner and colleagues (2016) suggest that the apparent increase in prevalence may have more to do with the reduced stigma associated with ID terminology than an actual increase in the number of individuals with an ID diagnosis.

While cultural factors may impact our ability to estimate ID prevalence accurately, there is also evidence for true prevalence differences cross-nationally that are linked to social factors, such as variability in wealth and access to resources. A meta-analysis of data collected across fifty-two studies between 1980 and 2009 estimated an ID prevalence of 1.04 percent (Maulik et al., 2011). Using the World Bank's estimates of income per capita for 2010, cross-national differences observed were related to degree of wealth and poverty. A higher ID prevalence was observed in countries with lower gross national income per capita (approximately 1.64%), and the lowest prevalence was observed in countries with the highest gross national income per capita (approximately 0.92%). Maulik and colleagues (2011) hypothesized that this may be the result of cross-national differences in access to maternal and child health care, and greater risk for biomedical factors such as iodine deficiency and birth-related infections in more impoverished geographic locations. Thus, even an issue as straightforward as estimating the basic prevalence of ID is confounded by social issues related to stigma in reporting of ID, and greater vulnerability to ID due to a lack of access to resources.

Comorbidity of ID with Other Conditions

Though a diagnosis of ID has important social implications, an additional contextual factor that intersects with ID is an increased vulnerability to comorbid diagnoses. Prevalence estimates of psychiatric disorders in people with ID range from 16 to 55 percent and vary across studies, depending on the diagnostic system applied, the methods used to assess symptoms, and the characteristics of the study sample (Fletcher, Barnhill, & Cooper, 2016). Diagnostic systems that incorporate clinical judgment and specify adaptations for evaluating a person with ID tend to report higher prevalence estimates. For example, in a large, multifaceted epidemiological study of adults with ID in the UK ($N = 1{,}023$), Cooper and colleagues (2007) reported that 40.9 percent of adults with ID met criteria for a psychiatric disorder based on clinical diagnosis. Conversely, if the diagnostic criteria from ICD-10 or DSM-IV was applied without considering the developmental functioning of the individual, the prevalence estimate fell to 15–16 percent (Cooper et al., 2007). Even these more conservative estimates indicate approximately a fourfold increase in risk for mental health problems, as compared to individuals without ID (Einfeld, Ellis, & Emerson, 2011).

Unfortunately, timely identification of co-occurring psychiatric conditions in persons with ID is a significant challenge in most communities. Several factors contribute to a delay in diagnosis, including clinician characteristics (e.g., lack of training in adapted assessment, little clinical experience with ID), methods (e.g., overreliance on self-report, limited access to range of informants and/or history, lack of measures validated for use with people with ID), and misconceptions held by clinicians and caregivers (Buckles, 2016). Clinical misconceptions include "diagnostic overshadowing," or the tendency for experienced clinicians to attribute psychiatric symptoms to ID itself and thus miss the opportunity to identify and intervene with evidence-based care (Reiss, Levitan, & Szyszko, 1982). Caregivers may focus on the challenging behaviors demonstrated by the person with ID and pursue behavioral and environmental interventions, not recognizing the potential for an underlying psychiatric/medical condition (Carr & Owen-DeSchryver, 2007).

Cognitive limitations associated with ID also pose obstacles in the assessment process. Difficulties with insight, talking about feelings, and realistically appraising one's skills and challenges are all thought to impede clinical interviews with adults with ID in the mild/moderate ranges of severity (Sovner, 1986). Individuals who are minimally verbal are least likely to be evaluated for co-occurring psychiatric disorders (Tager-Flusberg & Kasari, 2013).

Clinical practice guidelines for assessing mental health in persons with ID of different levels of severity are now provided through the DM-ID (Diagnostic Manual – Intellectual Disability; Fletcher et al., 2016). Modifications in assessment practices include simplifying language used in clinical interviews, incorporating visual supports, collecting information from multiple informants, and considering adapted criteria that reflect the developmental functioning of the individual being assessed (Fletcher et al., 2016). In addition to assessing sleep, appetite, and energy level, practitioners are guided to consider the etiology of ID in a particular individual, as neurogenetic conditions associated with ID may also be associated with increased risk for a particular psychiatric condition (Fletcher et al., 2016). For example, individuals with fragile X syndrome are at increased risk for social anxiety (Smith et al., 2012), and people with Prader-Willi syndrome are at particular risk for ritualistic behaviors (Dykens & Roof, 2008). For individuals with idiopathic ID, there could be unknown/unrecognized etiological factors that increase vulnerability to social-emotional functioning in addition to intellectual functioning. For example, alcohol exposure in utero is associated with both intellectual impairment and increased risk for attention deficit hyperactivity disorder (Olson, Morse, & Huffine, 1998). Infants who experience toxic stress may show slower growth intellectually as well as in the development of emotional regulation skills, which are thought to be a foundational set of capacities for emotional well-being (Oh et al., 2018). These factors should be integrated into the larger picture of treatment for individuals with ID who present with additional behavioral and adjustment issues.

Assessment and ID

Intelligence testing in the context of ID can be used for a variety of purposes, including diagnosis, classification, and educational planning (Brue & Wilmshurst, 2016; Schalock et al., 2010). However, while IQ assessment remains fundamental to both ID science and practice, there are numerous issues raised in the use of commercially available measures. These issues relate to both the psychometric properties of standardized assessments for use in ID and their decontextualized nature.

Validity of IQ Measures in Various ID Subgroups

IQ assessments are, for the most part, designed and normed based on the skill levels of the general population of children. Because of this, additional care must be taken when using these standardized measures with individuals with ID, as additional factors beyond intelligence may impact performance, and therefore, compromise psychometric integrity (Coolican, Bryson, & Zwaigenbaum, 2008; Tzuriel, 1992). This issue is becoming particularly salient in the area of cognitive assessment in neurogenetic syndromes. As there is new momentum in the direction of biomedical and pharmacological treatment studies to improve overall aspects of cognition in diagnostic groups such as Down syndrome (Gardiner, 2015) and fragile X syndrome (Wang et al., 2015), there is a growing recognition that current available gold standard assessments may not be valid for use in treatment studies that include these groups. For example, cognitive assessments that rely on verbal or motoric response with individuals with ID may produce findings that are confounded by expressive language and motor delays observed in some individuals in these populations. In addition to the phenotype-related issues in ID assessment, validity questions have been raised regarding the role of motivation during assessment in ID and more broadly (Duckworth et al., 2011).

Test-retest reliability

One potential threat to the utility of standardized measures for individuals with ID relates to reliability, or the degree to which measures will produce the same results across multiple administrations. Test-retest reliability is determined by estimating the degree to which differences in an individual's scores reflect differences in true scores relative to test procedures or characteristics (measurement error; Coaley, 2014). A critical factor that impacts reliability involves a measure's procedures. While it is impossible to account for all sources of measurement error, there are important steps that should be taken in ID assessment to minimize predictable sources of measurement error that result from developmental and behavioral issues that can serve as confounds. In particular, assessments that inadequately account for self-regulation characteristics often associated with ID may generate a substantial degree of measurement error with test procedures that assume competence in this area. With the increased likelihood of compromise in numerous domains of development, assessment in ID must account for the presence of much greater risk for measurement error than expected in the population of typically developing children.

Construct validity

In addition to capturing the degree of measurement error in the selected assessments (via test-retest reliability), it is critical to consider a measure's construct validity, or the degree to which an assessment captures the stated constructs of interest (Coaley, 2014). Evaluating construct validity involves identifying whether and how much a specific assessment elicits responses that minimize confounds from irrelevant domains and capture the putative construct in question.

For assessment in ID, the issue of construct validity is highly relevant, as co-occurring developmental presentations may compromise the validity of a measure. For example, a core feature of the behavioral phenotype associated with Down syndrome involves challenges in expressive language, and a general profile of relative competence in receptive language compared to expressive language (Abbeduto, Warren, & Conners, 2017; Malak et al., 2015). As such, cognitive assessments that require verbal responses will likely produce invalid data, as any underlying cognitive construct being evaluated will be confounded by the expressive modality required to generate an answer. Similarly, early development in Down syndrome is associated with delays in achieving gross and fine motor milestones (Fidler et al., 2005; Palisano et al., 2001), and developmental assessments that require a high degree of motoric proficiency or motor planning will suffer from similar compromises to construct validity. Overall, then, there are important potential threats to the psychometric properties of assessments when used with individuals with ID. In response, there has been a recent movement to examine these properties in specific diagnostic groups, in order to provide more accurate assessment and to capture treatment effects in upcoming studies.

Alternative Approaches to Assessment

Another concern regarding assessment in ID relates to the assumptions underlying the tests themselves. Traditional IQ assessments are designed with the presumption of stable and fixed levels of intelligence that can be measured via standardized approaches. Yet, an important premise of intervention and treatment work in ID is that developmental and behavioral functioning, including intelligence, is modifiable with effective instructional techniques. An important alternative approach has gained popularity that has roots in the work of Lev Vygotsky in the 1930s. This approach focuses on learning potential, rather than what an individual can produce at a single time point and under artificial social circumstances. These approaches fall under the broader category of learning potential assessments, and they began to take hold in the 1960s (Kozulin, 2005). One major development in this alternative to standardized testing is dynamic assessment, which is based on a pretest, intervention, posttest format that closely analyzes the individual's learning process and potential (Lidz & Peña, 1996). Dynamic assessment takes into account educational planning and views the child's strengths and challenges from a learning perspective (Lidz, 1987; Missiuna & Samuels, 1989). This approach involves an initial assessment, followed by the establishment of learning objectives, and then there is an examination of response to instruction (Grigorenko & Sternberg, 1998; Lidz, 1987; Missiuna & Samuels, 1989).

Another more general component of best practice in assessment of ID involves placing assessments in the context of a broader body of information. Beyond IQ tests, additional information may be gathered via parent or teacher questionnaires, observations of the child, developmental history, and family history (Shulman et al., 2011). This type of approach reduces the risk of misclassification based on arbitrary cutoff points used in standardized test scores (Sternberg & Grigorenko, 2000). Providing continued assessments across time also ensures that assessment is not solely used to diagnose and classify, but that the child's educational planning and environment are also benefitting from the use of assessment, which ultimately contributes to a supportive educational environment for individuals with ID (Shulman et al., 2011).

Conclusions and Future Directions

When considering where the field of ID science and practice has been, and directions forward, it is clear that the construct has been influenced by changes prompted by both advances in biomedical and other natural sciences, as well as sociohistorical trends that have evolved over the past century. Changes in our understanding of the causes of ID are closely yoked to rapid innovation in the field of genetics, advanced diagnostic procedures, and increasing awareness among medical professionals regarding the range of conditions associated with ID. Changes in approaches to ID classification systems have involved a response to societal trends that emphasize inclusive practice rather than marginalization of individuals with ID. Other changes in classification have been the result of more rigorous work relating genotype to phenotype. The assessment of ID has transformed in ways that address the needs of the educational community, which places an emphasis on learning potential. A call for even more change to assessment in ID is now coming from the scientific community, which is in need of psychometrically valid assessments for treatment work. In these ways, ID remains a construct where many trends intersect. In looking to the future, these junctures will likely influence the field of ID in several critical areas in upcoming years.

Treatment

As a result, in part, of the expansion of research on the link between neurogenetic syndromes and cognitive phenotypic outcomes, there has been a growing interest in the development of targeted treatments that can address the needs of individuals with specific ID diagnoses. These treatments are both behavioral and pharmacological, and they represent a new frontier in our understanding of the stability and malleability of intelligence and related aspects of neuropsychological functioning. For many of these current studies, IQ or aspects of cognition such as attention and memory have been identified as targeted constructs. The pharmacological work in this area has been informed by advances in mouse modeling work that shows preliminary evidence of improvement in, and even the potential "rescue" of, aspects of cognition and information processing. To date, some clinical trial work that includes individuals with Down syndrome or fragile X syndrome has led to modest effects (Hart et al., 2017). However, this work is still very much in its infancy, and null effects may be attributable, in part, to psychometrically invalid outcome measures in the target areas of interest. If these treatment approaches are shown to be effective, and, of course, safe, the discussion regarding the nature of ID in upcoming handbooks will likely look radically different than it does today.

New Frontiers for Inclusion

A second area that will likely undergo important changes over the next several decades relates to inclusive practices in educational and community settings. While great progress has been made with respect to disability rights in the United States and other countries, there are still many ways in which individuals with ID remain at the margins. One new, but rapidly growing area of interest relates to inclusive higher education (Butler et al., 2016; Plotner & Marshall, 2015). Once thought to be out of reach for individuals with ID, there is a growing number of higher education institutions that have developed programs that facilitate college and university attendance for individuals with ID after the completion of secondary education. Intriguing conversations in the literature have been raised regarding how best to support these experiences (Cook, Klein, & Chen, 2015; Griffin et al., 2016), and a variety of models are already in use. Some institutions have taken a "mixed/hybrid" approach, where students with ID are involved in campus-related activities and take some coursework along with the broader community of students, but they also are involved in life skill classes (Cook et al., 2015). Other models involve more focused supports for an individual or a small group of individuals with ID that coach them toward success on their educational pathway (Cook et al., 2015). In the context of these growing programs, the inclusion of individuals with ID has challenged perceptions of what the boundaries are for individuals who show challenges related to intellectual functioning and adaptation.

Summary

In looking to the future, the foundations have been laid for important progress in treatment and societal inclusion for individuals with ID. It is difficult to predict exactly what form those innovations will take, but if the past is predictive of the future, we can be certain that many aspects of the construct of ID and the field of ID research will look very different in twenty years than they do today.

References

Abbeduto, L., Warren, S. F., & Conners, F. A. (2007). Language development in Down syndrome: From the prelinguistic period to the acquisition of literacy. *Mental retardation and developmental disabilities research reviews*, *13*(3), 247–261.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Washington: American Psychiatric Association.

American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington: American Psychiatric Association.

American Psychiatric Association Committee on Nomenclature and Statistics. (1952). *Diagnostic and statistical manual of mental disorders*. Washington: American Psychiatric Association.

Bennett, S. J., Holmes, J., & Buckley, S. (2013). Computerized memory training leads to sustained improvement in visuospatial short-term memory skills in children with Down syndrome. *American Journal on Intellectual and Developmental Disabilities*, *118*(3), 179–192.

Biasini, F. J., Grupe, L., Huffman, L., & Bray, N. W. (1999). Mental retardation: A symptom and a syndrome. In S. D. Netherton, D. Holmes, & C. E. Walker (Eds.), *Child and adolescent psychological disorders: A comprehensive textbook* (pp. 6–23). New York: Oxford University Press.

Boyle, C. A., Boulet, S., Schieve, L. A., Cohen, R. A., Blumberg, S. J., Yeargin-Allsopp, M., et al. (2011). Trends in the prevalence of developmental disabilities in US children, 1997–2008. *Pediatrics*, *127*(6), 1034–1042.

Broman, S. H., Nichols, P. L., Shaughnessy, P., & Kennedy, W. (1987). *Retardation in young children: A developmental study of cognitive deficit*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Brue, A. W., & Wilmshurst, L. (2016). *Essentials of intellectual disability assessment and identification*. Hoboken, NJ: John Wiley & Sons.

Buckles, J. (2016). The epidemiology of psychiatric disorders in adults with intellectual disabilities. In C. Hemmings & N. Bouras (Eds.), *Psychiatric and behavioral disorders in intellectual and developmental disabilites* (pp. 34–44). Cambridge, UK: Cambridge University Press.

Butler, L. N., Sheppard-Jones, K., Whaley, B., Harrison, B., & Osness, M. (2016). Does participation in higher education make a difference in life outcomes for students with intellectual disability? *Journal of Vocational Rehabilitation*, *44*(3), 295–298.

Carr, E. G., & Owen-DeSchryver, J. S. (2007). Physical illness, pain, and problem behavior in minimally verbal people with developmental disabilities. *Journal of Autism and Developmental Disorders*, *37*(3), 413–424.

Coaley, K. (2014). *An introduction to psychological assessment and psychometrics*. Los Angeles: SAGE.

Cook, R. E., Klein, M. D., & Chen, D. (2015). *Adapting early childhood curricula for children with special needs*. Edinburgh: Pearson.

Coolican, J., Bryson, S. E., & Zwaigenbaum, L. (2008). Brief report: Data on the Stanford–Binet intelligence scales in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *38*(1), 190–197.

Cooper, S. A., Smiley, E., Morrison, J., Williamson, A., & Allan, L. (2007). Mental ill-health in adults with intellectual disabilities: Prevalence and associated factors. *British Journal of Psychiatry*, *190*(1), 27–35.

Duckworth, A. L., Quinn, P. D., Lynam, D. R., Loeber, R., & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences*, *108*, 7716–7720.

Dykens, E. M., & Roof, E. (2008). Behavior in Prader-Willi syndrome: Relationship to genetic subtypes and age. *Journal of Child Psychology and Psychiatry*, *49*(9), 1001–1008.

Einfeld, S. L., Ellis, L. A., & Emerson, E. (2011). Comorbidity of intellectual disability and mental disorder in children and adolescents: A systematic review. *Journal of Intellectual and Developmental Disability*, *36*(2), 137–143.

Fidler, D. J., Hepburn, S. L., Mankin, G., & Rogers, S. J. (2005). Praxis skills in young children with Down syndrome, other developmental disabilities, and typically developing children. *American Journal of Occupational Therapy*, *59*(2), 129–138.

Fletcher, R. J., Barnhill, J., & Cooper, S. A. (2016). An introduction to the DM-ID–2. In *Diagnostic manual – intellectual disability 2: A textbook of diagnosis of mental disorders in persons with intellectual disability* (pp. 1–12). New York: NADD Press.

Forness, S. R., & Kavale, K. A. (1994). The Balkanization of special education: Proliferation of categories for" new" behavioral disorders. *Education and Treatment of Children*, *17*(3), 215–227.

Fujiura, G. T., Yamaki, K., & Czechowicz, S. (1998). Disability among ethnic and racial minorities in the United States: A summary of economic status and family structure. *Journal of Disability Policy Studies*, *9*(2), 111–130.

Gardiner, K. J. (2015). Pharmacological approaches to improving cognitive function in Down syndrome: Current status and considerations. *Drug Design, Development and Therapy*, *9*, 103–125.

Griffin, M. M., Wendel, K. F., Day, T. L., & McMillan, E. D. (2016). Developing peer supports for college students with intellectual and developmental disabilities. *Journal of Postsecondary Education and Disability*, *29*(3), 263–269.

Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, *124*, 75–111.

Hart, S. J., Visootsak, J., Tamburri, P., Phuong, P., Baumer, N., Hernandez, M. C., et al. (2017). Pharmacological interventions to improve cognition and adaptive functioning in Down syndrome: Strides to date. *American Journal of Medical Genetics Part A*, *173*(11), 3029–3041.

Hodapp, R. M. (1997). Direct and indirect behavioral effects of different genetic disorders of mental retardation. *American Journal on Mental Retardation*, *102*(1), 67–79.

Hodapp, R. M., & Dykens, E. M. (2009). Intellectual disabilities and child psychiatry: Looking to the future. *Journal of Child Psychology and Psychiatry*, *50*(1–2), 99–107.

Iarocci, G., & Petrill, S. A. (2012). Behavioral genetics, genomics, intelligence, and mental retardation. In J. A. Burack, R. M. Hodapp, G. Iarocci, & E. Zigler (Eds.), *The Oxford handbook of intellectual disability and development* (pp. 13–29). New York: Oxford University Press.

Kozulin, A. (2005). Learning potential assessment: Where is the paradigm shift? In D. B. Pillemer & S. H. White (Eds.), *Developmental psychology and social change* (pp. 352–367). New York: Cambridge University Press.

Lemons, C. J., King, S. A., Davidson, K. A., Puranik, C. S., Fulmer, D., Mrachko, A. A., et al. (2015). Adapting phonological awareness interventions for children with Down syndrome based on the behavioral phenotype: A promising approach? *Intellectual and Developmental Disabilities*, *53*(4), 271–288.

Lidz, C. S. (1987). *Dynamic assessment: An interactional approach to evaluating learning potential*. New York: Guilford Press.

Lidz, C. S., & Peña, E. D. (1996). Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to Latino American preschool children. *Language, Speech, and Hearing Services in Schools*, *27*(4), 367–372.

Maenner, M. J., Blumberg, S. J., Kogan, M. D., Christensen, D., Yeargin-Allsopp, M., & Schieve, L. A. (2016). Prevalence of cerebral palsy and intellectual disability among children identified in two US national surveys, 2011–2013. *Annals of Epidemiology*, *26*(3), 222–226.

Malak, R., Kostiukow, A., Krawczyk-Wasielewska, A., Mojs, E., & Samborski, W. (2015). Delays in motor development in children with Down syndrome. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, *21*, 1904–1910.

Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., & Saxena, S. (2011). Prevalence of intellectual disability: A meta-analysis of population-based studies. *Research in Developmental Disabilities*, *32*(2), 419–436.

McDuffie, A., Machalicek, W., Bullard, L., Nelson, S., Mello, M., Tempero-Feigles, R., et al. (2016). A spoken-language intervention for school-aged boys with fragile X syndrome. *American Journal on Intellectual and Developmental Disabilities*, *121*(3), 236–265.

Missiuna, C., & Samuels, M. (1989). Dynamic assessment of preschool children with special needs: Comparison of mediation and instruction. *Remedial and Special Education*, *10*, 53–62.

Morgan, P. L., Farkas, G., Hillemeier, M. M., Mattison, R., Maczuga, S., Li, H., & Cook, M. (2015). Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. *Educational Researcher*, *44*(5), 278–292.

Oh, D. L., Jerman, P., Silvério Marques, S., Koita, K., Purewal Boparai, S. K., Burke Harris, N., et al. (2018). Systematic review of pediatric health outcomes associated with childhood adversity. *BMC Pediatrics*, *18*(1), 83. https://doi.org/10.1186/s12887-018-1037-7

Olson, H. C., Morse, B. A., & Huffine, C. (1998). Development and psychopathology: Fetal alcohol syndrome and related conditions. *Seminars in Clinical Neuropsychiatry*, *3*(4), 262–284.

Palisano, R. J., Walter, S. D., Russell, D. J., Rosenbaum, P. L., Gémus, M., Galuppi, B. E., et al. (2001). Gross motor function of children with Down syndrome: Creation of motor growth curves. *Archives of Physical Medicine and Rehabilitation*, *82*(4), 494–500.

Plotner, A. J., & Marshall, K. J. (2015). Postsecondary education programs for students with an intellectual disability: Facilitators and barriers to implementation. *Intellectual and Developmental Disabilities*, *53*(1), 58–69.

Pulina, F., Carretti, B., Lanfranchi, S., & Mammarella, I. C. (2015). Improving spatial-simultaneous working memory in Down syndrome: Effect of a training program led by parents instead of an expert. *Frontiers in Psychology*, *6*, 1265.

Reiss, S., Levitan, G. W., & Szyszko, J. (1982). Emotional disturbance and mental retardation: diagnostic overshadowing. *American Journal of Mental Deficiency*, *86*(6), 567–574.

Schalock, R. L., Borthwick-Duffy, S. A., Bradley, V. J., Buntinx, W. H., Coulter, D. L., Craig, E. M., et al. (2010). *Intellectual disability: Definition, classification, and systems of supports* (11th ed.). Washington: American Association on Intellectual and Developmental Disabilities.

Shulman, C., Flores, H., Iarocci, G., & Burack, J. A. (2011). Intellectual disability: Concepts, definitions, and assessment. In P. A. Howlin, T. Charman, & M. Ghaziuddin (Eds.), *Handbook of developmental disorders* (pp. 365–402). Thousand Oaks, CA: SAGE Publications Inc.

Singh, N. N., Lancioni, G. E., Myers, R. E., Karazsia, B. T., Courtney, T. M., & Nugent, K. (2017). A mindfulness-based intervention for self-management of verbal and physical aggression by adolescents with Prader–Willi syndrome. *Developmental Neurorehabilitation*, *20*(5), 253–260.

Smith, L. E., Barker, E. T., Seltzer, M. M., Abbeduto, L., & Greenberg, J. S. (2012). Behavioral phenotype of fragile X syndrome in adolescence and adulthood. *American Journal Intellectual Development Disability*, *117*, 1–17. https://doi.org/10.1352/1944-7558-117.1.1COMP: Link

Sovner, R. (1986). Limiting factors in the use of DSM-III criteria with mentally ill/mentally retarded persons. *Psychopharmacology Bulletin*, *22*, 1055–1059.

Sternberg, R. J., & Grigorenko, E. L. (2000). Theme-park psychology: A case study regarding human intelligence and its implications for education. *Educational Psychology Review*, *12*(2), 247–268. https://doi.org/10.1023/A:1009079500495

Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. *Autism Research*, *6*(6), 468–478.

Tzuriel, D. (1992). The dynamic assessment approach: A reply to Frisby and Braden. *Journal of Special Education*, *26*, 302–324.

Wang, H., Pati, S., Pozzo-Miller, L., & Doering, L. C. (2015). Targeted pharmacological treatment of autism spectrum disorders: Fragile X and Rett syndromes. *Frontiers in Cellular Neuroscience*, *9*, 55.

World Health Organization (2018). *International Classification of Diseases 11th Revision (ICD-11)*. Geneva: World Health Organization.

Zigler, E. (1967). Familial mental retardation: A continuing dilemma. *Science*, *155*(3760), 292–298.

Zigler, E., & Hodapp, R. M. (1986). *Understanding mental retardation*. Cambridge, UK: Cambridge University Press.

12 Prodigies and Savants

David Henry Feldman and Martha J. Morelock

In this chapter we consider the relationship between general intelligence and prodigies and savants. At first glance the relationship would seem to be straightforward: Prodigies have high general intelligence, savants have low general intelligence. While this relationship generally holds, there is more to the story. As we will see, there is a more nuanced set of relationships between general and specific intelligences in prodigies and savants that not only helps us understand these unusual forms of intellectual functioning, but also helps resolve long-standing debates in the field between general versus specific intelligence.

Because prodigies and savants have rarely been studied together, we will review each literature separately, attempting to provide a current summary of what is known and understood about each. For example, prodigies appear in a wider array of fields than savants, and there are some areas where the two do not overlap; for example, there are no calendar prodigies and there are no savants in chess. We will also attempt to provide a framework for joint study of the two phenomena that may shed light on each as well as on their possible relationships to each other.

We should note that the two subfields of research that deal with savants and prodigies are different in several ways, and that these differences influence how much is known and how confident we can be in research findings to date. For savants, there is a substantial research tradition that goes back more than a century and is part of the medical field (Treffert, 1989, 2006, 2008, 2009, 2013, 2014). The techniques for doing research tend to reflect the deficit/remediation orientation of a medical approach. Over the years there has been a sustained interest in and commitment to research that may provide intervention to or relief for some of the burdens that most savants carry. For prodigies, research stretches back almost as long but has been sporadic and relatively uncommon. Although there were a small number of studies in the early decades of the previous century (e.g., Baumgarten, 1930; Revesz, 1916/1970), the empirical base of knowledge about prodigies is not large, and almost all of it is based on case studies by psychologists. Prodigies are generally assumed to be blessed with greater gifts than most and are typically not seen as requiring resources to ameliorate their "condition."

Defining Prodigies and Savants

There is relative consensus on how to define a savant but less agreement on the definition of a child prodigy. A savant (formerly referred to as an "idiot savant") is a person (not necessarily a child) who displays an island of exceptional mental performance in a sea of disability (L. K. Miller, 1989, 1999; Treffert, 1989, 2006, 2008, 2009). The syndrome can be either congenital, or acquired by a normal person after injury or disease to the central nervous system. The skills can appear – and disappear – suddenly and inexplicably. The area of exceptionality for savants can be remarkable in contrast to their generally low level of functioning in other areas (i.e., "talented savant"), or it can be so extreme as to be spectacular even if it had been viewed in a normal person (i.e., "prodigious savant"; Treffert, 1989, 2006). For example, a calculating savant may be able to multiply numbers of many digits by other numbers of many digits in their head as quickly as a computer. Or a calendar savant may be able to produce the day or the week for any day in the past or the future with only a few seconds' delay, with uncanny (if not perfect) accuracy. There have been artistic savants whose works are considered to be of professional quality. In spite of such exceptionalities, most savants are unable to live independently and require major support from family and/or society to survive.

Unlike research into the savant, prodigy research has generated a fair amount of disagreement over definitional issues. Until late in the last century, there was no scientific or technical definition of the child prodigy. Dictionary definitions referred to the origin of the word "prodigy" as an omen or portent, an event out of the usual course of nature.[1] The earliest definitions of prodigies were not limited to children but rather referred to an event that was cause for wonder and/or for impending changes that were not necessarily welcome. During the decades when psychometric definitions of intelligence were dominant, prodigies were defined as exceptionally high–IQ children (cf. Hollingworth, 1942; Tannenbaum, 1993). For Hollingworth, an IQ exceeding 180 put the child in the range of what would be required to be considered a prodigy.

In recent decades, an effort to provide a more technical definition of the child prodigy for purposes of research has stimulated both the desired research and some disagreement over just what constitutes a prodigy (Edmunds & Noel, 2003; Hulbert, 2005; Morelock & Feldman, 1993, 2003; Ruthsatz & Detterman, 2003; Shavinina, 1999). The definition proposed in Feldman with Goldsmith (1986) defined a prodigy as a child younger than ten years of age who performs at an adult professional level in a highly demanding field. This definition was intended to be explicit and precise enough to be tested empirically. For example, if further research revealed that children, although performing extraordinarily well for children, still did not reach adult professional levels of performance until well after ten years of age, that finding would tend to weaken the part of the definition that is age specific. For the most part, research on child prodigies has used the 1986 definition either as a guide or as a foil (e.g., Kenneson, 1998; McPherson, 2006, 2007, 2016; Radford, 1990; Shavinina, 1999).[2] For the purposes of this chapter, we will use a variation of the definition

1 *Webster's Third New International Dictionary*, 1961, s.v. "prodigy."

2 There have also been several books written by journalists, critics, and historians, or the individuals themselves, about child prodigy lives. These works have added valuable information about specific cases but are not social science research as such. Examples of works in this tradition are Clynes (2015); Conway and Siegelman (2005); Kanigel (1991); Rolfe (1978); Ruthsatz and Stephens (2016); Solomon (2012); Wallace (1986); Weiner (1953); and the many books about Mozart (e.g., Hildesheimer, 1982/1977).

proposed in 1986, recognizing that there is some disagreement as to its adequacy. A prodigy is defined as a child who, at a very young age (typically younger than ten years old), performs at an adult professional level in a highly demanding, culturally recognized field of endeavor. A prodigy's performance is assessed as being at a professional level based on the standards of their field as well as the reaction of the general audience, reflected, for example, in sales of paintings and positive reviews of performances.

Although both prodigies and savants are very rare, there are no solid estimates of the frequencies of their occurrence in the general population. Most identified savants are males, although there have certainly been exceptions (e.g., Selfe, 1977). It has been estimated that savant syndrome occurs six times as often in males as in females (Hill, 1977). Traditionally, most prodigies have been males as well, although that has changed dramatically in the past thirty years (Feldman with Goldsmith, 1986; Goldsmith, 1987).

Prodigies

Research on Child Prodigies

The contemporary field of research with child prodigies began with the publication of a study of six boys under the age of ten in the fields of music, chess, and writing (and a child, labeled an "omnibus prodigy," who had not yet settled into a specific area) (Feldman with Goldsmith, 1986). The boys were between three and eight years of age when first studied, and were followed for as many as ten years. The study focused on each child's specific and general abilities, experiences with their teachers and their families, and development in their specific field in the context of their more general development. This is the study that proposed the working definition described in the section Defining Prodigies and Savants. The findings most frequently cited from this research are that a child prodigy has a mix of child- and adult-like qualities; that prodigies require the sustained efforts of at least one parent, teachers, and others to support the development of their talent; that the process requires several years even in the most extreme cases; that the talents of prodigies are at least partly natural and inborn (the more extreme the case, the more nearly completely inborn the talents are likely to be); and that prodigies' talents tend to be domain specific and require above average but not necessarily extreme intelligence.

One study of eight prodigies (as defined above) in chess explored the extent to which proficiency at the level of a professional tournament player as a child predicted how well these chess players performed as young adults (Howard, 2008). The research was intended to shed light on the issue of natural talent as well as the role of practice in achieving world-class levels of performance. The study also dealt with an issue that often is cited as a reason to be skeptical of the prodigy phenomenon: the fact that relatively few child prodigies become successful adult performers in their original field of endeavor. In chess, at least, the child performers were highly likely to become successful adult performers in the same domain.

The results of this study support the importance of natural talent in the field of chess as a critical ingredient in success and that a prodigy is difficult to explain without recourse to a substantial natural talent base from which to work (Feldman, 1995, 2008; Morelock & Feldman, 1993, 1999; Winner, 1996). Most of the children have achieved a high level of international success in spite of the fact that they are not likely to have practiced as long as many players who have performed less well. On a number of measures, the child prodigy chess players exceeded in skill other high-level players in chess. For example, they needed fewer games to reach master levels, required fewer years to achieve grand master status, and were younger when they received grand master ratings. One of the eight became a world champion, although other known world champions were not necessarily identified as child prodigies under the present definition.

Another chess study, this one of the Polgar sisters as compared with another chess prodigy and eight grand masters (Howard, 2008), showed that hours of practice alone does not predict chess performance. The three Polgars had similar numbers of hours of practice, for example, but varied substantially in the level of chess expertise they achieved. And the fourth prodigy, with fewer hours than the Polgars and a later start, exceeded all three sisters in chess rating. While not direct evidence for the importance of natural talent, the study modulates the common claim (e.g., Gladwell, 2008) that practice is the key to expertise.

A case study (Ruthsatz & Detterman, 2003) explored the importance of general intellectual ability (IQ) in the performance of a piano prodigy, arguing that IQ contributes significantly to the six-year-old's ability to perform at a high, professional concert level in his chosen domain. Along with "domain-specific skills," a well above average IQ (an attained score in what would typically be considered the gifted range) was found to contribute to the child's overall performance. Most striking was the child's general and specific musical memory capabilities. The study also tended to discount the most common alternative explanation for the child's exceptional level of performance, namely, practice (Ericsson, Krampe, & Tesch-Romer, 1993), inasmuch as the child had not yet received formal training in music. Overall, this study points to a combination of elevated IQ, domain-specific natural abilities, and practice as implicated in high-level performance within the field of music, a conclusion that we will affirm at the end of this chapter when we summarize the state of current knowledge and theorizing about prodigies and savants.

A more recent study of eight prodigies across several fields (art, math, music, gastronomy) provides a more complex picture of the prodigy profile (Ruthsatz & Urbach, 2012). For the eight participants, IQ ranged from above average to highly superior, but working memory was consistently extremely high, as was attention to detail. In addition, autistic behavior (based on the AQ test for autism) was found in four of the eight families (three in prodigies, five in first-degree or second-degree relatives). The relative frequency of autism raised the possibility of a common quality shared with savants, who also show a much higher than average incidence of autistic tendencies.

A case study in another domain (writing) was carried out by Edmunds and Noel (2003). The study focused on the writing that their subject produced during a period of about twelve months, from about age five in 1999 to about age six. This child (Geoffrey) was interested in math and science and much of his writing reflected these interests, although his first thirty-page work was based on the then popular Pokémon cartoon books and was written for Geoffrey's younger brother. The authors report that this work was done very quickly and in a "rush of creative energy" (Edmunds & Noel, 2003, p. 188), which was to become Geoffrey's way of writing.

All told, Geoffrey wrote 129 works during this brief period, totaling more than 1,500 handwritten pages. Reproduced here is part of the final work, a letter to one of his mentors, which communicates his astonishing levels of understanding of math and science concepts and a remarkable ability to communicate them in writing, as well as some childish playfulness:

Dear Jim,
I am into math but also science. Here's the math part. I know addition, addition with tens and ones, multiplication, multiplication with tens and ones, division, and division by zero!! Here's how that works. 5 [divided by] 0 = undefined, or, the answer is undefined. I can do algebra, addition with tens, ones, hundreds, thousands, and millions up to infinity. . . . I also have a bunch of questions. What is calculus? . . . How do you get –0 if it exists?

Now, some science. I do theoretical physics just like you. I am working on a unified theory. Are you? And if you're not, what's the theory you're working on anyways? . . . My unified theory is broken up into many parts, each part the size of special relativity . . . E = sp, meaning energy = speed of light pulses. It is the theoretical answer to why Pikachuic electricity is so fast. . . . I really know my geometry, even though I'm in grade 1! I know that a rhombicosidodecahedron has 240 forces. A rhombicosidodecahedron is the largest known polyhedron. It is huge!

XOX
Geoffrey

Edmunds and Noel (2003) analyzed examples of Geoffrey's writing over and noted areas of major change in style and sophistication. Using standard measures of language, Geoffrey's level exceeded high school students' norms, and showed tendencies toward transformation and innovation in language that are unusual at any age.

As to the question of intelligence in the traditional psychometric sense, Geoffrey had been given a Wechsler Intelligence Scale for Children – Third Edition (WISC-III) test and scored "moderate-to-high," with an IQ of 128. On the Raven's, he scored higher, above the 99th percentile for age thirteen (Edmunds & Noel, 2003, p. 192). Informally, the authors noted an unusual memory ability that allowed Geoffrey to recall, in detail, work that he had done several months prior to the interviews. Overall, the authors found that the most striking quality that Geoffrey displayed was a "dogged persistence" to learn. This persistence is what Kevin Kearney, father of Michael, who graduated from college at age ten, called a "rage to learn" (Kearney & Kearney, 1998; Morelock, 1995). It appears in the most extreme cases of prodigious achievement.

Edmunds and Noel (2003) preferred the term "precocity" to prodigy, emphasizing rapid early mastery of knowledge and focusing less on the mysterious and elusive qualities of the prodigy and the difficulties in defining a prodigy precisely. Terminology and emphasis notwithstanding, their case study adds significantly to the existing literature on prodigies. Writing prodigies are rare even among the range of prodigies, and the approach that Edmunds and Noel have taken to understanding Geoffrey's abilities in the context of his domain of expertise and his development adds richness and detail to the small body of knowledge in the scholarly literature.

Two more recent studies of musical prodigies (by far the most frequent type of prodigy) have provided important data on the specific abilities that contribute to their extreme levels of performance (Comeau et al., 2017a; Comeau et al., 2017b). Using both general psychometric as well as music-specific tasks, Comeau and colleagues (2017a) compare the performance of an eleven-year-old piano prodigy (LN) with a number of other musicians, including other music prodigies. They also do a retrospective comparison with one of the earliest cases reported in scientific literature (Revesz, 1916/1970), that of Ervin Nyireghazi, also a piano prodigy. The results tend to confirm the basic claims that a prodigy performs at an adult professional level, has a high but not necessarily exceptionally high IQ, has exceptionally strong working memory skills, and benefits greatly from devoted parents and teachers. The specific music tasks (some of which duplicate the tasks from more than a century earlier) show a distinct pattern for LN, with much higher performance on some (e.g., pitch accuracy), less than average performance on others (e.g., improvisation). A striking finding was that on many tasks LN and Nyireghazi scored similarly, in spite of being tested more than a hundred years apart.

The second study (Comeau et al., 2017b) tried to gather empirical data to address the claim that music prodigies perform at the level of a professional musician. Using audio clips of prodigy and professional musicians playing the same pieces, musicians and nonmusicians judged which were which. For the most part prodigies and professional musicians were judged similarly, although trained musicians were better at distinguishing between them than nonmusicians. The study also found that "older" prodigies (11–14) were harder to distinguish from professional musicians than "younger" prodigies (under 10).

Theoretical Interpretations

There have been a small number of more interpretive or theoretical efforts to try to comprehend and make sense of the prodigy phenomenon. Prodigies have fascinated and inspired awe and wonder for millennia, but there has been little advance in explanation beyond divine inspiration, reincarnation, or magical incantation. Some of the more conceptual/theoretical work has centered on definitional issues, such as in the Edmunds and Noel (2003) study described above. The term "prodigy" continues to carry powerful associations stemming from its ancient meaning as something "out of the usual course of nature." Consequently, there has often been considerable aversion to the term both within and outside the scholarly community (Radford, 1990).

One early response to the definitional issue was simply to place the prodigy within the range of IQs from lowest to highest, with the child prodigy at the highest extreme of the distribution (i.e., above 180 IQ), as Leta Hollingworth (1942) did in her classic work on extremely high IQ. By placing the prodigy under the umbrella of IQ, its many complexities and associations with nonscientific traditions could be wiped away. It also put prodigies squarely into the psychometric IQ tradition. Unfortunately, the prodigy did not fit well under this definition; an IQ of 180 (or even several standard deviations lower) was rarely required for a child to become a prodigy, and the astonishing performance of children in specific domains could not be explained by high general intelligence alone.

Feldman proposed a revised definition of the prodigy, placing the phenomenon within an evolutionary and cultural historical framework (Feldman with Goldsmith, 1986), which he termed "co-incidence." The construct of co-incidence was intended to acknowledge the mysterious nature of the prodigy phenomenon and to recognize that interpretations that seem irrational and unscientific, such as reincarnation and astrology, are understandable in the face of the baffling reality that the prodigy represents. Reducing the prodigy to extreme high IQ, Feldman argued, diminishes its complexity, ignores the fact that prodigies occur only in a small number of domains, and tends to discourage further research. It also was unsupported by empirical data: Only one of the six cases in the study would have qualified using Hollingworth's definition (above 180 IQ).

It is assumed in the co-incidence framework that child prodigies are naturally endowed with extraordinary talent. Even the most extreme talent, however, cannot fully account for the prodigy. The child's family is essential (particularly a parent who is totally devoted to the development of the child's talent), as are their teachers (who must balance the astonishing capability of the child with the need to guide and direct the child's mastery of critical skills and knowledge, in proper sequence); the current state of the child's chosen domain must match the child's talents (as it is claimed that domains, as well as children, undergo developmental transitions and transformations); the broader social/cultural context in which a field channels resources, sets standards, responds to pressures from inside and outside, and confers status that can increase or decrease the likelihood that a prodigy's talent will be recognized and celebrated must be congenial, as well as the period of history in which all of the other forces interact (a war, pestilence, or a great economic boom can have profound influences on opportunities or the lack of them; Simonton, 1994).

A number of scholars have criticized the co-incidence framework, and in doing so, have added some important additional conceptual distinctions and possible areas of further research (Edmunds & Noel, 2003; Ruthsatz & Detterman, 2003; Shavinina, 1999). Edmunds and Noel, for example, believe that precocity is a better designation than child prodigy to avoid the issues that tend to come along with the term. The advantage of the focus on precocity is that it invites close attention to the specific behavior of the child in relation to what is normative for the domain, for age peers, or in relation to more advanced students of the domain. Psychologist and educator Julian Stanley also promoted the term "precocity" in advocating accelerated education for intellectually precocious youth, including youths who could reason exceptionally well mathematically or verbally, and those showing exceptional spatial and mechanical talent (Brody & Stanley, 2005; Lubinski & Benbow, 2006; Lubinski, Benbow, & Morelock, 2000; Lubinski et al., 2001; Stanley, 1996, 2000).

Ruthsatz and Detterman (2003) found that co-incidence tends to diminish the importance of psychometric intelligence in accounting for the prodigy's achievements; in their case study of a six-year-old musical prodigy, they found that the child scored an IQ of 132 on the 1985 version of the Stanford-Binet intelligence test, although his pattern of scores was idiosyncratic, with a range from 114 (abstract reasoning) to 158 (short-term memory). The argument that general intelligence as traditionally assessed – that is, through an IQ test – is implicated in this child's superior performance in music is consistent with data from other studies (e.g., Feldman, with Goldsmith, 1986; Simonton, 1999). For a child prodigy (as contrasted with a calculating savant, for example), an IQ in the above-normal range seems to be necessary.

Shavinina (1999) comes at co-incidence from a different angle, finding it inadequate in its ability to explain the actual mental and emotional processes of development and experience that are distinctive to the gifted and to the prodigy. Shavinina's proposed addition to the set of considerations when trying to comprehend the reality of the prodigy is a function of a phenomenon called "age sensitivity," which in turn is involved with "sensitive periods" in the child's development. These notions are adapted from research and theory done by Leites (1960, 1996), with use of terms somewhat different from Western scholarly research. "Sensitive periods" (Bornstein & Krasnegor, 1989; Thompson & Nelson, 2001), for example, refer to universal processes that help explain why children during a period of years are particularly receptive to and particularly adept at learning languages, much less so thereafter. Sensitive periods as used in Western psychological studies do not refer to individual differences between and among children, but this is how Shavinina (1999) uses the term.

Terminology aside, Shavinina's emphasis on the distinctive cognitive and emotional qualities and experiences that may be involved in producing a prodigy is a welcome one. In Shavinina's terminology, for the prodigy, a "sensitive period" of intense involvement with a domain changes from a more typical "developmental" sensitive period to an "individual" one. In other words, for the prodigy the often intense but fleeting passions of growing children may transform into a lifelong career, as in the case of a child who was fascinated by birds and became a highly renowned ornithologist as an adult (Shavinina, 1999).

Brain Imaging Research on Prodigies

Although it would seem like an obvious choice for research, there have been few studies of brain function and/or brain development in prodigies. With the availability of powerful imaging techniques like functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and others, prodigy cases may be able to shed light on some of the most enduring issues in the study of intelligence (Morelock, 2013). Questions of both anatomical and functional differences between prodigy brains and more typical brains appear to be compelling areas of research. Since its beginning more than a century ago, the question of one versus more than one form of intelligence has remained controversial. Given that the prodigy tends to be a child with extreme ability in a single field, knowing what brain areas tend to be implicated compared with those of brains in less gifted children might help address the domain-general versus domain-specific question. Are prodigies' brains anatomically distinct in detectable ways? Are there distinctive brain areas responsible for different prodigy fields – for example, for music, for chess, for visual art?

As compelling as these questions may be, we know of no research directly addressing them. There are, however, some studies on related topics that may be relevant to prodigies. A number of studies examined mathematically gifted students as compared with less gifted ones (e.g., O'Boyle, 2008a, 2008b; Singh & O'Boyle, 2004). In these studies, the brains of mathematically precocious children and adolescents were studied morphologically, developmentally, and functionally. Distinctive processes and patterns of activation were found for the mathematically talented children, as well as evidence of enhanced development of the right cerebral hemisphere and possible enhanced connectivity and integrative exchange between right and left hemispheres (Singh & O'Boyle, 2004). It is reasonable to expect that similar, and perhaps more pronounced, differences between mathematical prodigies and others would be likely to occur.

A related area of research has been carried out with "calculating prodigies," one of the traditional areas in which astonishing performance has been observed going back several centuries (Smith, 1983). That these calculating savants were called prodigies has led to some confusion about the phenomenon. For most of the history of Western mathematics, arithmetic was a major activity. In more recent centuries, complex mathematical reasoning has become increasingly more central to the field. Thus, centuries ago a calculating savant (who could, for example, divide or multiply large sums rapidly) was called a "mathematics prodigy," where today such a child or adult would be labeled a "calculating savant."

An article reviewing research on Rudiger Gamm, in which he is called a "calculating prodigy," illustrates the problem. The title of the article (Butterworth, 2001) is "What makes a prodigy?" when it perhaps should have been "What makes a savant?" As the article says, "Gamm is remarkable in that he is able (for example) to calculate 9th powers and 5th roots with great accuracy, and he can find the quotient of 2 primes to 60 decimal places" (Butterworth, 2001, p. 11). The analysis of Gamm's brain activation as compared with six nonexpert calculators revealed (using PET scan procedures) distinctly different patterns. The problem is that by contemporary standards, Gamm is a calculating savant, not a child prodigy, particularly because he did not begin his calculating efforts until he was twenty.

There have also been brain imaging studies of trained musicians versus less trained or untrained individuals, revealing reliable differences between and among the various levels of training and experience (e.g., Schlaug et al., 1995a, 1995b), showing that trained musicians have a larger than average corpus callosum (as was true of the mathematically precocious children) as well as other differences in brain morphology and activation. Studies of the effects of musical training on cortical development also have shown that training affects organization and reorganization of brain circuitry without resolving the question of plasticity and/or inborn susceptibility to training effects as the main source of the change (Baeck, 2002).

General and Specific Abilities in Prodigies

A small number of studies of child prodigies in the fields of art and music have been carried out by scholars with a background in the specific field rather than in social science research. One such study (Kenneson, 1998) of musical prodigies was done by Claude Kenneson, a professor of music at the University of Alberta in Canada. Kenneson did not consider his subjects' academic intelligence as a separate topic, but it can be indirectly accessed from his account of their experiences. For example, Canadian cellist Shauna Rolston received bachelor's and master's degrees in music history and music performance with distinction from Yale University and later became a professor of cello at the University of Toronto. Academic achievements of this sort are unlikely without substantial academic ability, and we can assume with confidence that Shauna Rolston possessed such abilities. Similarly, cellist Yo Yo Ma studied at Columbia and Harvard. As Kenneson writes, "It was at Harvard, where he [Ma] distinguished himself studying the humanities, that he realized that music has as much to do with philosophy, history, psychology, and anthropology as it has to do with playing an instrument well" (Kenneson, 1998, p. 330).

The advantages are significant when a study is carried out by someone who is deeply involved and highly accomplished in a field where prodigies are found. One of the very few additional examples in the literature of a study by a scholar with training and experience in both the domain of interest and in social science research is that of Milbrath (1998), who studied visual art.

Milbrath's study bears directly on issues of intelligence and talent, although not in the traditional psychometric sense. Milbrath studied several highly talented drawing prodigies over several years, giving her the opportunity to analyze change over time and the contributions of various aspects of intellectual functioning to the drawings that children produced. Examples of drawings by one of Milbrath's subjects are shown in Figures 12.1–12.4.

A question that interested Milbrath was the role that natural talent plays in the development of exceptionally talented visual artists. Taking Piaget's notions of figurative and operative knowledge as a starting point, Milbrath asked if these processes might help explain how her very young subjects could possibly have produced drawings as sophisticated as they did.

In Piaget's theory of intelligence, figurative and operative knowing are reciprocal processes that, together, provide the basis for construction of knowledge (Feldman, 2000), functioning similarly in all people. As an artist, Milbrath wondered if figurative and operative knowing might vary from person to person, with future artists tending to have more acute figurative processes (sharper perceptions, a more acute sense of color, etc.) while at the same time being less controlled than others by operative processes of ordering, categorizing, and discerning logical relationships. The other way in which Milbrath thought artistic prodigies might differ from others

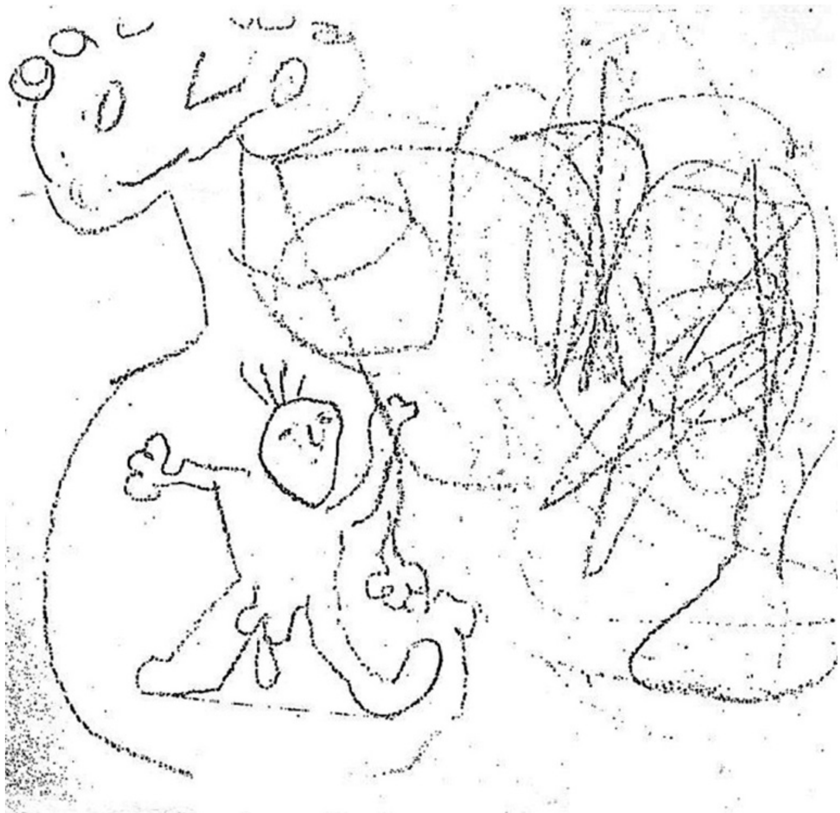

Figure 12.1 *Drawing by two-year-old Peregrine (from Milbrath, 1998, fig. 3.7b).*

less talented is in their continued emphasis on sensorimotor intelligence even as other children move toward more advanced (in the Piagetian sense) cognitive developmental processes.

Milbrath found support for her hypotheses and shed light on one of the current controversies in the field. A number of scholars who have studied high-level performance in several fields (sports, music, visual arts, chess, and others) claim that "deliberate practice" is the best explanation for differences in levels of expertise (Ericsson, 1996; Howe, Davidson, & Sloboda, 1998). These scholars argue that about 10,000 hours of well-planned and guided practice is the variable that separates exceptional from less exceptional performers. For Milbrath, the age and quality of her subjects' work would make deliberate practice an unlikely source of explanation for their work (although, to be sure, her subjects spent a great deal of time practicing their craft).

Milbrath found that the developmental course of talented children's drawing is qualitatively distinct from that of less talented children, with the difference primarily in attentiveness, awareness, and preoccupation of the talented children to the figural qualities of objects. Talented children are also less controlled by the conceptual structures that constrain less talented children, leading the less talented children to emphasize what they "know" more than what they "see."

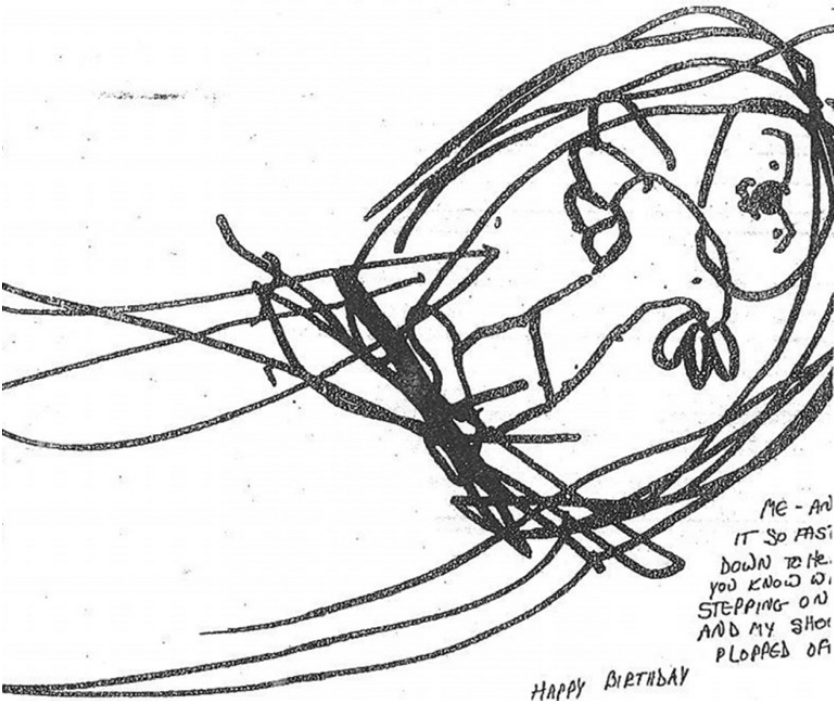

Figure 12.2 *Drawing by two-year-old Peregrine (from Milbrath, 1998, fig. 3.7a).*

Figure 12.3 *Drawing by eight-year-old Peregrine (from Milbrath, 1998, fig. 6.25b).*

Figure 12.4 *Drawing by eleven-year-old Peregrine (from Milbrath, 1998, fig. 4.10b).*

Savants and Intelligence

According to Darold Treffert (2008, 2014), a physician and one of the leading scholars of savant syndrome, the first recorded case of savant syndrome was reported in the scientific literature almost 160 years ago, although it was about 120 years ago that Dr. J. Langdon Down described savant syndrome as a distinct condition. As compared with research on child prodigies, there has been a great deal more work done over more than a century of activity. The vast majority of savant studies have come from the medical research community, although a significant number of studies have also been reported by psychologists. More recently, brain studies have begun to appear in the scientific literature.

There is a sufficiently large base of research on savant syndrome, as it tends to be labeled, since Treffert's 1989 book (it had been originally labeled "idiot savant"), to consider savants in specific domains: calendar calculation, music, mathematics, art (primarily drawing), and memory. There are also occasional cases in other areas, such as sensory sensitivity, mechanical aptitude, and language (L. K. Miller, 1999). There has been a good deal of interest in savant cases as they relate to both general psychometric intelligence and more specific cognitive processes. There are also several films that have portrayed the savant, from the 1988 commercial film *Rain Man*, starring Dustin Hoffman, to a documentary called *A Real Rainman*, based on the late Richard Wawro, an autistic savant who was a remarkable visual artist (Zimmerman, 1989). The life of Kim Peek, the savant who was actually a real-life inspiration for the character Dustin Hoffman played in *Rain Man*, has also been documented in two fascinating accounts by his father, Fran Peek (1997; Peek with Hanson, 2007).

General and Specific Abilities in Savants

From the earliest studies, savants have been described as severely lacking in general intellectual abilities, with an area of superior ability that stands out relative to their overall low functioning, or more rarely, stands out relative to the broad population. It is the latter kind of case that has drawn the most attention from the research community (and, not surprisingly, from the media). In recent decades, the degree of severity of the overall intellectual deficit appears often to be less than was originally believed (in IQ terms, savant cases were originally thought to have IQs around 20–40, but several studies have shown savants with IQs near or even above normal; Treffert, 2009); the appearance of Daniel Tammet (2006, 2009) in the literature has further supported the possibility of both high IQ and extreme savant skills appearing in the same person.

Savant research has also shed light on the question of the viability of theories of multiple intelligences (e.g., Gardner, 1983; Sternberg, 1985). Treffert (2009), for example, believes there is evidence among some savants that supports the existence of several intelligences in the areas where savants appear: music, mathematics, visual art, mnemonics, and perhaps others. Although Treffert acknowledges that most savants are known to have low IQ scores, he finds that fact to be of limited value in explaining the remarkable ways that "intelligence" sometimes manifests itself in savants. For example, Treffert describes a concert by Leslie Lemke, a blind, autistic musical savant whose IQ measures in the 35–55 range:

> At this particular concert Leslie was asked to play a piece he had never heard before with the other pianist, rather than waiting for the piece to conclude and then play it back as he usually does. The other pianist began playing. Leslie waited about three seconds and then did indeed play the piece with the other pianist, separated only by those three seconds. . . . Leslie was parallel processing, just as some very intelligent, but rare, interpreters are able to translate what a speaker is saying into another language simultaneously. . . . That would not be possible if the level of IQ of 35–55 was an accurate barometer of his over-all intelligence. He exceeds that level by far . . . which signals that more than a single "intelligence" was at work during that complex performance. (Treffert, 2008, pp. 2–3)

Brain researcher Allan Snyder (2009) proposes that all individuals have savant skills, but most of us have inhibited these skills through adoption of and preference for the reasoning and abstract thinking that is adaptive in our highly technological and rationalized environments. Thus, we normally respond to our experience not in terms of the stream of information and sensory details bombarding us but, rather, in terms of conceptual mindsets. Using magnetic techniques to "turn off" higher mental processes of the brain, he and his coworkers have demonstrated that savant-like abilities are sometimes latent in normal subjects.

Robyn Young (1995) investigated the talents and family backgrounds of fifty-one savants recruited throughout Australia and the United States. The selection of savants included prodigious and talented savants as well as those with "splinter skills" – levels of interest and competence only marginally above the level of general functioning. Young found the parents and siblings of the savant participants to be exceptionally able, with above-average IQ and frequency of high-level skills, though not necessarily the same skills as those displayed by the savants. In addition, there was a family predisposition toward high achievement, possibly genetically predisposed and/or part of a tradition, which provided encouragement and reinforcement for savant skills. The researcher concluded that savants have an underlying biological predisposition toward high general ability that is tempered by neurological impairment. The resultant savant skills are encouraged through familial support.

Research on Savants' Intelligence and Related Topics

Young, incorporating psychometric measures into the study, found peaks and valleys in the Wechsler Adult Intelligence scale (WAIS) profiles of the savant sample. The researcher consequently took exception to the widely held notion that savants manifest islands of extreme capability showcased against a backdrop of overall severely deficient intellect. Among the fifty-one savants, sixteen had a subtest score at least one standard deviation above the population mean, and 60 percent had at least one subtest one standard deviation above the Full-Scale score. Highest scores were revealed in Block Design, Object Assembly, and Digit Span; lowest scores were found on Comprehension, Coding, and Vocabulary. These patterns are compatible with strengths and weaknesses of savant functioning documented in the literature (i.e., verbal/conceptual weaknesses and perceptual strengths). In addition, the level of precocity exhibited by the savants (i.e., prodigious or talented) was found to be positively correlated with the level of general cognitive ability, as indexed by IQ.

The idea that savant cognition is best described as islands of extreme capability showcased against a backdrop of overall severely deficient intellect emerged from the earliest writings on savants. A case study by Scheerer, Rothmann, and Goldstein (1945) was the first to document features of savant functioning that thereafter were repeatedly observed. These include (1) minimal abstract reasoning ability and almost exclusive reliance on concrete and literal patterns of expression and thought, (2) lack of metacognition, (3) extraordinary memory, (4) flattened affect, and (5) limited creativity. Elaboration and examples of each of these follow.

Scheerer and colleagues (1945) wrote of one savant who memorized and sang operas in several languages yet had no comprehension of the conceptual and symbolic meaning of the words. Still, the question of abstract reasoning in savants is a complex one. Studies show that savants have an immediate, seemingly intuitive access to the underlying structural rules and regularities of their domain, whether it be music (L. K. Miller, 1999; Treffert, 1989), mathematical calculation (Hermelin & O'Connor, 1986; O'Connor & Hermelin, 1984), or art (O'Connor & Hermelin, 1987). Furthermore, these are the same rules and regularities as those applied by practitioners of normal or high reasoning ability who are skilled in the same area.

It appears, therefore, that even though most savants can't reason conceptually, they can abstract to a degree – at least in circumscribed and domain-specific areas (O'Connor, 1989; L. K. Miller, 1999). L. K. Miller (1999) suggests that what is missing in savants is a conceptual system that can reconstruct domain-specific knowledge, transferring it into a more generalized framework, affording a decontextualized representation containing less perceptual detail but better adapted to varied application (see Karmiloff-Smith, 1992).

Savants appear to be incapable of metacognition. They cannot reflect on their internal thinking processes or explain how they arrived at correct responses to posed questions (Scheerer et al., 1945). When asked to account for how they can do whatever it is that they do, they frequently respond with something irrelevant. O'Connor (1989) reports that one calendar calculator who was able to render remarkably fast responses to date questions was, nevertheless, usually unable to add or subtract without pencil and paper. Yet, when asked how he managed his calendar feats (e.g., giving the correct answer to a question such as "On what day of the week did September 1744 fall?"), he responded simply, "I make all sorts of mathematical calculations, don't I?" Some savants are able to articulate rule-based strategies. Those who do so tend to have higher IQs than do their counterparts (Hermelin & O'Connor, 1986). Savant Daniel Tammet, who reports having a measured IQ of 150 on the WAIS (top 1% of the population on that measure), has an exceptional ability to describe what he sees in his head and to reflect on his cognitive processes (Tammet, 2009). This has prompted Allan Snyder's comment that Tammet "could be the Rosetta Stone" in terms of what we can learn from him about savant cognition (Johnson, 2005).

All savants have extraordinary memories. Savant mnemonists are notable solely for their impressive memory for miscellaneous or mundane happenings (e.g., some savants have been known to remember weather conditions for each day of most of their lives). In other savants, it is the norm for their incredibly powerful memories to be limited to their domains of achievement.

Savants exhibit a restricted range of emotion, precluding the experience of heightened passion, excitement, or sentiment (Treffert, 1989, 2006). In the case of musical savants, for example, this usually comes across in performance as shallow imitative expressiveness lacking subtlety or innuendo. However, there have been some cases of musical savants demonstrating emotional connection with the music they were performing (L. K. Miller, 1989, 1999; Viscott, 1970). In one such case (Viscott, 1970), the savant exhibited more expanded verbal abilities than is commonly the case with savants and this ability may have allowed for an interpretive response to the music. As another possible explanation, emotional response to music can be, to some extent, the direct result of the physiological changes it evokes (Winner, 1982). Music has been found to affect pulse, respiration, blood pressure, and the electrical resistance of the skin, while also delaying the onset of fatigue (Mursell, 1937). These types of changes also occur during emotional experience. The question is whether the emotional response seen in musical savants is more a straightforward reflection of specific physiological effect than is the case with musicians more conceptually and interpretively involved in the performance of their music.

Earlier research findings suggested that savants are incapable of being creative in the sense of producing original work. Treffert (1989) concluded that while musical savants might imitate, improvise, or embellish based on preestablished constraining musical rules, they are generally incapable of composing. Sacks (1995) later distinguished between two different kinds of creativity, acknowledging as creative the individuality of savant ability based on perceptual talent while recognizing that even the prodigious savant does not achieve a higher order of creativity involving the invention of new ideas and new ways of seeing things. Daniel Tammet appears to be an exception once again. In his recent (2009) book, he brings together research on the brain and neuroscience, concluding with a theory of "hyperconnectivity" to account for autistic functioning as well as creativity. In addition, he describes an original language which he has been creating since childhood called "Mänti," based on the lexical and grammatical structures of Baltic and Scandinavian languages.

Supporting Sacks' observation is evidence that musical savants with more highly developed language capacities are more likely to compose music. One musical savant, "L. L.," studied by L. K. Miller (1999), developed more complex language over a period of months, with capacities evolving from simple monosyllabic or echolalic responses to conversational generation of requests, comments, and more sophisticated responses to questions. At the beginning of this period, L. L. remained musically confined to renditions of songs and melodies written by others, with little inclination to improvise or compose. At the end of the study, however,

L. L. announced and played an original composition. This concordance of the development of expanded language skills with the onset of musical creativity led Miller to speculate that music and language are not independent (see also Patel, 2008).

More Recent Research and Interpretation of the Savant Phenomenon

Research has intensified and increased greatly during recent years, with some important new findings and interpretations of savant skills and how they develop. There have been advances in two areas that bear directly on savants and intelligence. One of these is of general interest and deals with all savants; this work tends to show that previously assumed constraints on IQ and other capabilities do not always hold for savants – that there is more diversity and greater plasticity in savant development than was previously believed (L. K. Miller, 1999; Treffert, 1989, 2006, 2008, 2009). The other advance is specific to calendar savants; there are now plausible explanations for how calendar savants are able to achieve their remarkable results (Thioux et al., 2006) as well as some research on the ways that general intellectual level may interact with savant capabilities over the course of development (Cowan et al., 2004). We will review these recent areas of research for what they may tell us about savants and intelligence.

Plasticity and Diversity in Savants

While, in general, it remains true that savants tend to be impaired in most areas other than their special skill, it is less true than was believed until quite recently. In a review of research, L. K. Miller (1999, 2005) found considerable variation among savant cases within a skill area as well as variation from specialty to specialty. Treffert (2006, 2008, 2009) reported similar findings. Nonetheless, there do seem to be certain abilities that are implicated in each specific savant domain. These tend to be present in all cases, whether of the more profound sort, with performance comparable to that of a person not afflicted with disabilities, or more “splinter” skills that are exceptional in relation to the other areas of functioning of the savant but not necessarily exceptional when compared with the best performers in that field.

L. K. Miller (2005) reports that among musical savants, preestablished component skills of absolute pitch, aural melody retention, aptitude for harmonic analysis, and ability to reproduce what is heard tend to be present. For drawing savants, visual memory for detail, awareness of perspective, and an ability to depict what is seen are the common skills. Among calendar savants, event memory and attribution of personal meaning to date and numerical information are typically found.

Along with the typical strengths, there are typical weaknesses: Recognition of previously seen figure drawings was no better among drawing savants than for other mentally impaired individuals (O’Connor & Hermelin, 1987). Musical savants have difficulty with same versus different judgments, even with musical notes that they can identify perfectly. And savants rarely have general intellectual abilities above normal. For calendar calculation in particular, there appears to be a relationship between the development of calendar calculation knowledge and IQ, with higher IQ associated with more extensive and more accurate calendar calculating skills (O'Connor, Cowan, & Samella, 2000, cited in L. K. Miller, 2005).

In a study of two young calendar savants aged five and six years, Cowan and colleagues (2004) explored the relationship of general intellectual ability (IQ) to calendar calculation development. As children, the two boys were remarkable in their skills, but not as adept or as accurate as most adult calendar savants. When retested two years later, neither boy had improved in calendar calculation, and the hypothesized reason for their lack of improvement (indeed, their diminished interest in calendar calculation) was attributed to their normal and exceptional IQs (scored on the Wechsler III – UK Edition); one child had a Full Scale IQ of 105, the other, 141. These robust scores on a standard IQ appeared to give the boys options to pursue other interests typically not available to a savant. The early stimulus for calendar activity was probably a physical limitation that isolated the boys (one had a hearing problem, the other a visual one). Both boys had become more social and were pursuing activities more typical of boys their age. Although these results are from only a single study of two boys, they suggest that lower IQ or general intellectual ability of the sort assessed on an IQ test may constrain development in other areas.

L. K. Miller (1999), summarizing studies of calendar savants by Hermelin and O'Connor (1986; O'Connor & Hermelin, 1987) and others, reports some evidence for IQ-related differences (range 50–114), with higher IQ associated with better performance: a wider range of calendar knowledge and better application of rules in other tasks. The finding was particularly robust when based on the Performance subscale of the WAIS.

In a study of one of the most impressive young calendar calculating savants, Thioux and colleagues (2006) tried to account for the child's performance with a series of studies that led to an explanatory model for his behavior. The model includes three components: memory of fourteen calendars stored in the form of fourteen verbal associative networks; processes that access these fourteen calendars through "anchoring years" close to the present; finally, simple arithmetic operations based on calendar rules to match past and future with a year already associated with a calendar. Here is how Thioux and colleagues describe their findings:

> Our working hypothesis is that the appearance of savant skills is determined not only by the presence of circumscribed interests but also by a specific profile of neuropsychological abilities including, in the case of calendar skills, strong rote memory and good elementary calculation ability. . . . The model presented here suggests that calendar skills may rely mostly on parietal areas of the brain because this region is important both for simple calculation (addition and subtraction) and for rote verbal memorization of multiplication facts, which we believe is a process quite similar to memorizing date-weekday association. . . . In summary, we propose that two conditions are necessary and probably sufficient for the development of savant skills: (a) the presence of circumscribed foci of interests with a predilection for repeating behaviors and (b) the relative preservation of parietal lobe learning abilities. (pp. 1167–1168)

Two other areas where savant syndrome research has influenced the field of intelligence are the venerable issues of one versus several intelligences, typically described as “g” versus “s” theories of intelligence; and the related question of the existence of distinct “modules” that are innately available and that are designed to respond to and process specific kinds of information (e.g., musical, linguistic, spatial, social, etc.). Within the savant syndrome research community, there has been growing consensus that an adequate theory of intelligence needs to be able to account for the reality of savant behavior, and this consensus leads to a tendency to embrace one or another form of “multiple intelligence” theory (Gardner, 1983; L. K. Miller, 1999, 2005; Treffert, 1989, 2006, 2008, 2009).

L. K. Miller (1999) concludes an extensive review of the savant research literature with the argument that the existence of savants supports multiple-intelligence frameworks:

> The traditional notion that savants represent exceptionality in the context of general mental retardation has been modified in recent definitions. The consistent finding of at least some intact component skills in savants stands in contrast to the inconsistent evidence for special motivational conditions or tutoring. This suggests that modular explanations of savant behavior are likely to fare better than those stressing more generic factors in skill acquisition. . . . [T]he types of skills found in savants . . . are at best loosely congruent with current modular models (e.g., Gardner, 1983). (p. 36)

Taking this conclusion more cautiously, Treffert, whose career has been spent studying and working with savants, sees the general versus specific theories of intelligence issue as far from resolved: Arguing for comparative studies involving prodigies, genius, and savants, Treffert (2009) calls for such research since:

> the interface between genius, prodigies and savants is an important, and in some ways a very narrow one, those persons should be included also in the multidisciplinary, multimodality compare and contrast studies. Such studies can shed light on the debate regarding general intelligence versus separate intelligences. (p. 1355)

On the other hand, in describing the more extreme “prodigious” cases of savant abilities, Treffert (1989) leaves little doubt that a theory that includes separate intelligences as well as general intelligence is necessary: “In the prodigious savant . . . the skills are so spectacular, and the inherent access to the rules and ‘language’ behind those skills so extensive, that there must be, at least as part of the reason, a genetic endowment that somehow is preserved apart from, and that exists separately from, overall intelligence” (p. 222). These recent efforts calling for a theory that transcends the either/or debate over one versus more than one intelligence appear to be moving toward a more nuanced view (see Chapter 22, Intelligence and the Cognitive Unconscious, this volume). Based on both prodigy and savant research, the existence of relatively isolated, relatively specific natural abilities seems likely to be confirmed. The existence of at least some domain-general abilities is also likely to be affirmed. The questions become more about how the specific and the more general abilities interact, influence each other, and explain the range and diversity of intellectual profiles found in our species.

The related topic of modules (Fodor, 1983) and/or modularization (Karmiloff-Smith, 1992) of functions has tended to play itself out largely around the topic of language development, an area of deficit in virtually all savant cases. For this reason, much of the work on modules is only indirectly relevant to savants. There have been only a few language savants, however, and these have been controversial and closely studied because of their potential direct relevance to the modularity issue.

The case of "Christopher" has been at the center of the discussion in recent years. Christopher is a remarkable language savant who can read, write, and translate between and among more than a dozen languages. Smith and Tsimpli (1995) wrote a book about Christopher, in which they claim that his abilities provide compelling evidence for a "language module" that functions quite independently from general intelligence. Follow-up work (Tsimpli & Smith, 1999) responds to criticism of their claims that Christopher proves by his amazing abilities the existence of such a language module. The disputed evidence turns on whether Christopher is sufficiently impaired in general intelligence to support the claim that his language abilities (which are indeed protean) function independently of "cognitive prerequisites" associated with a mental age of about five years.

When Christopher's intelligence was tested, his performance IQ was consistently lower than his verbal intelligence (Bates, 1997), with scores on nonverbal tests ranging from 42 to 76 and verbal scores all above average. The question is what specifically are the prerequisites of cognitive development that may underlie first-language acquisition, and there is no clear consensus on this question. If Smith and Tsimpli (1995) are right, Christopher functions in language areas substantially independent of general cognitive development, thus supporting the modularity claim. If not, then his first-language acquisition was enabled normally, that is, bootstrapped off general cognitive functions available between three and five years of age in normally developing children.

A key issue is that Christopher's abilities in his first language (English) are unremarkable; what is remarkable is his ability to learn second languages. It may be that the same abilities are involved with both processes or that there are differences between them. At the least, learning a first language is (logically) prerequisite to learning the second, and so on. The arguments are complex and technical, but the conclusions reached at this point seem tentative. There is evidence that some functions of language are independent of more general cognitive development and general intelligence, and there is some evidence that learning one's first language depends at least in part on at least some of the functions attributed to general cognitive development. Tsimpli and Smith (1999) offered a reasonable summary of the current situation: "Language is only partially modular. It also belongs in the central system. This is not just vague anarchic agnosticism; we have made explicit suggestions about which parts of language belong in which domain" (p. 213). Although the questions of specific versus general functions, and modules versus general intelligence, are not fully resolved, research with savants has helped to sharpen the issues and provide important data that bear directly on the issues.

Brain Studies of Savants

Because savants are often in institutional care, they are frequently the responsibility of the medical community. The desire to learn about the source of the savant's abilities and disabilities has led to studies of brain function, morphology, and development. Although not many studies exist, there is a sufficient number to offer some provisional interpretations of brain and central nervous system involvement in savants.

Current imaging technologies provide clear views of savant brain architecture, allowing comparisons to be made with normal brains. Brain function, however, has been more difficult to access because most technologies require that subjects remain immobile during the procedure (e.g., computed tomography [CT], magnetic resonance imaging [MRI]). Some newer techniques (e.g., positron emission tomography [PET], functional magnetic resonance imaging [fMRI], and single photon emission computed tomography/computed tomography [SPECT-CT]) allow activity (e.g., drawing) during the imaging procedure. The newest ones (e.g., diffusion tensor imaging, diffusion tensor tracking) provide information about brain connectivity between hemispheres and other parts of the brain, as well as images of brain fibers, that is, the "wiring" of the brain. Near-infrared spectroscopy allows the subject to perform music or paint while wearing an infrared cap (Treffert, 2009).

Young's (1995) previously referenced work was the largest study of savants to date and included fifty-one cases (12 "prodigious," 20 "talented," and 19 with "splinter" skills). All had neurological impairments but preserved neurological capacity for information processing in their specific area of skill. A process of atypical brain development may account for some savants, that is, left-brain dysfunction (language, abstract reasoning, reflection) with right-brain compensation. This applies to both congenital and acquired savant skills. Comparable compensatory brain functioning has been found in other populations, as well. B. L. Miller and colleagues (B. L. Miller et al., 1998; B. L. Miller et al., 2000) and Hou and colleagues (2000), studying frontotemporal dementia patients, found that this condition generally involved loss of function in the left temporal lobe with enhanced functioning of the posterior neocortex (Treffert, 2009).

There is also growing acknowledgment of greater than previously believed plasticity in brain development and function. As has been found in studies of brain development in normal subjects (cf. Thompson & Nelson, 2001), savants appear to recruit and reassign brain materials for the specialized purposes of their skill (Treffert, 2009). The ability of the brain to recruit resources from areas that are not usually devoted to the functions that savants develop appears in both congenital and acquired cases. These findings, should they be confirmed by future studies, have implications for our understanding of intelligence and how its more general and more specific forms are developed.

General Conclusions

The past few decades have seen significant progress in research with prodigies and savants. The field of prodigy studies has been revived and, although not large, has produced a steady flow of research and some important new findings and interpretations. The area of savant studies has seen a marked increase in activity, stimulated in part by the availability of new technologies for brain imaging that include the possibility of studying savants while they are actively engaged with their skill area. In this concluding section, we summarize some of the noteworthy advances in each area of study and put forward some provisional generalizations about the ways in which more general and more specific kinds of intelligence interact, placing what appear to be opposite extremes within a single interpretive framework.

Progress in Prodigies Research

For prodigies, there is considerable evidence that extremely high IQ is not a prerequisite for prodigious achievement. The more likely relationship between IQ and child prodigies is that IQ in the average range sets the lower boundary between prodigy and savant. For some domains (e.g., mathematics, physics), an IQ much higher than average is probably a necessary prerequisite for prodigious achievement (cf. Simonton, 1999), while for visual art an extremely high IQ may be an impediment to the emphasis on the figurative aspects of knowledge essential for that kind of endeavor (cf. Milbrath, 1998).[3]

Recent research tends to affirm that child prodigies can be found among girls, in some fields more frequently than boys. There were few girls found in research studies before the 1980s, although there have been some famous girl prodigies in the public eye for centuries (cf. Goldsmith, 1987, 2000). In the visual arts, though no cases had been documented in scientific case studies before 1980 (there were autistic girl artists like Nadia; see Selfe, 1977), artists like Wang Yani (Ho, 1989) and the cases in Milbrath (1998) are mostly girls.

There has been progress in distinguishing between mathematical prodigies and calculating savants (sometimes called calculating prodigies). Historically (cf. Smith, 1983), calendar calculators and arithmetic calculators were called prodigies. Since diagnostic procedures were not available to determine how many such cases were also autistic, mentally impaired, or both, there is no way to be sure, but recent child prodigy studies have found no cases of individuals younger than ten years old that would meet the definition of adult professional performance in the domain of mathematics as it is now practiced. It appears likely that the widely held belief that there have been mathematics prodigies is inaccurate, and that the cases so labeled were actually calculating savants of various IQ levels or even high-IQ individuals with apparent savant-like skills.

This labeling dilemma is worth pondering in more depth. As a case for definitional discussion, consider George Parker Bidder (1806–1878), one of the most brilliant nineteenth-century English civil engineers. Bidder is recorded as having been able, by the age of ten, to solve calculations such as dividing 468,592,413,563 by 9,076 (Campbell, 2005). The question arises: Was Bidder a savant, a high-IQ savant (autistic or non-autistic), a prodigy, or a high-IQ individual with savant-like skills?

It is clear from the level of his adult achievement that Bidder possessed sufficient general cognitive ability to be considered a "prodigy" or even a high-IQ savant rather

3 Although Milbrath's interpretation of the interplay between figurative and operative processes is plausible, a case like Leonardo da Vinci seems to contradict it. A man of immense intelligence as well as an artist of great stature, Leonardo may be an exception that proves the rule.

than either a talented or prodigious savant, as classically defined. Bidder's later achievements in engineering, debate, and politics, with all that implies in the sense of complex professional and social demands (Clark & Linfoot, 1983), rules out the classical savant possessing extraordinary skills standing in stark contrast to overall handicap, or even the notion of his being a high-IQ autistic savant like Daniel Tammet, since such would imply considerable social deficits.

Was he a non-autistic savant? In 1856, Bidder (1856) made a presentation to the Institution of Civil Engineers, carefully laying out the principal operations and algorithms involved in his mental computation. As a very simple example, he reported that to multiply two three-digit numbers, he started from the left, multiplying first the hundreds together, and adding each successive product to the total so as to hold as few intermediate sums in his head during the calculation as possible (Clark & Linfoot, 1983). He carried in his head key results from earlier calculations, learned to use successive approximations, and deduced new rules as he went along. Unlike Tammet and other savants, whose numerical abilities are largely intuitive and unconscious, Bidder's calculations were conscious and explicitly logical. He was capable of analyzing them and explicating them, and even believed that his methods could be taught to children to improve their mental arithmetic. Bidder also reported that he visualized numbers as shapes in his mind, a predilection that he attributed to the fact that he began to calculate before he learned to write (Clark & Linfoot, 1983). Daniel Tammet also reports that numbers appear in three-dimensional shapes in his mind. Unlike Bidder, however, Tammet reports that these shapes spontaneously chunk together to generate a mathematical solution. He then reads off the "numerical landscape," a process typical of savant skills (Snyder, 2009).

Was Bidder, then, a prodigy? The deciding rule of thumb would be whether at that time, arithmetic calculation was considered a culturally recognized domain of achievement ripe for prodigies, with associated standards for professional-level performance. While Bidder, as a child, developed a national reputation as a "calculating boy" who performed at local fairs and even, at one point, for the queen, calculating alone failed to parlay itself into a professional path. Bidder required a viable profession, such as engineering, for him to use his calculating skills productively and to contribute to society.

Ultimately, what we can conclude is that Bidder was a high-IQ individual with savant-like domain-specific skills. His introspective reports and later professional achievement leave no doubt that his skills reflected robust executive functioning and extraordinary conscious analytical and logical skills harnessed in the process of calculation. Nevertheless, his childhood domain of achievement did not allow for the emergence of a prodigy whose level of performance could be assessed as equal to that of an adult professional, since standards for "adult professionals" did not exist – nor did adult professionals exist in the field of mathematical calculation at that point in history.

Availability of appropriate resources, technologies, instruction, and opportunities for recognition enable or constrain the expression of prodigy possibilities, as do broader cultural and historical contexts that may impact opportunities and possibilities. In the extreme, a war on home soil is certain to constrain organized development and recognition of exceptional performance in all prodigy fields. On the other hand, the same conditions may make the appearance of prodigious achievement more likely in other domains; Joan of Arc may be an example from history of a prodigy in military leadership (Feldman with Goldsmith, 1986).

Research on prodigies bears on the general versus specific intelligence issue, although it does not support an either/or resolution. The prodigy reveals a complex relationship between more general and more specific aspects of intelligence (as does the savant, as we discuss in the section The Interplay of General and Specific Intellectual Abilities). For the prodigy, an IQ in the average range (minimally about 90–110) seems necessary as a contributor to the amazing performance that is the hallmark of the child prodigy. The general intelligence aspect of prodigy performance seems to give the child access to the social, cultural, and specific traditions of the domain, to allow for generalization and reflection, as well as give the child access to the social, emotional, and pedagogical dimensions of the field. These broader aspects of the knowledge domain and its context provide access to and a basis for the child's progress in reaching the higher levels of their domain.

The more specific aspects of intelligence help determine which domain the child will engage, and which specific areas the child will pursue (e.g., in music, instrument choice, musical genre, pedagogical tradition, performance venues, and the like). Specific talents for particular kinds of activities (e.g., chess versus visual art) are related to but not determined by general intellectual abilities. It is in the interplay between more general abilities and more specific talents that the child prodigy's area of achievement will crystallize. Both general as well as specific aspects of intelligence are involved in the choice of domain, the kind of activity within that domain, and the level of achievement ultimately reached through their sustained interplay.

Progress in Research with Savants

Savants are now seen more as a source of knowledge about brain and cognitive functions and less as anomalies (Treffert, 1989, 2006). Whereas most research on child prodigies remains based on single or small case studies, savant research now includes larger samples, some experimental studies, and several sustained research centers with systematic programs of research. What has emerged from this heightened activity is a better understanding of savant syndrome, recognition that the constraints on savant performance are not as severe as once believed, and an understanding that general intelligence is likely to be a moderating variable that helps determine how and why a savant does what he (or occasionally she) does.

Perhaps the greatest advances in understanding of the savant mind have been with calendar savants (and calendar "prodigies" and "calculators," who tend to have higher IQs). It now seems likely that the severity of the disabilities that accompany the specific talents of the savant, as well as the degree of general intellectual impairment, largely determine the initial involvement in calendar activity, the degree of skill, and the range of the savant's capabilities, as well as the likelihood that a savant will continue their preoccupation with the activity into adulthood (cf. Cowan et al., 2004).

The main reasons for continuing to pursue savant-like activities are that they provide a sense of competence and that they are recognized and admired in the (typically) institutional context (L. K. Miller, 1999; Treffert, 2006, 2014). If a savant at some point is able to function in the wider community, the likelihood of sustaining and enhancing the specific savant skills diminishes (Cowan et al., 2004). The greater the constraints from other limitations and/or impairments, the greater the likelihood that the savant will sustain and continue to pursue greater achievements in the circumscribed domain in which they can succeed.

A second advance, also with calendar savants, is in research that has led to a plausible framework to account for their amazing abilities. In a series of elegant studies, Thioux and colleagues (2006) were able to construct a relatively straightforward cognitive model to explain how "Donny" (one of the fastest and most accurate calendar savants on record) was able to perform his feats. For Donny, fourteen calendar types were stored in long-term memory; these types were accessed through a set of anchoring years close to the present, and a few simple arithmetic calculations link the fourteen models with any past or future year. An overall IQ that is not severely retarded, and at least nominal access to the knowledge domain, complete the picture. The model does not demean or lessen the remarkable achievement of the savant, but it does go a long way toward demystifying how and why that achievement occurs.

A third advance, following anecdotal reports from Treffert's (2014) decades long observations of individual savants, is evidence of modest creative abilities in some savants and some indications of modest development in their abilities over time. A study by Pring and colleagues (2012), for example, showed that a group of nine savant artists performed better on a standard, if limited, test of creativity than a group of nine individuals with mild learning disabilities, as well as a group of nonartists. Only a group of nine talented art students performed better. And Treffert (2013, 2014) has observed improvements in the abilities of musical savants to improvise over time.

Finally, brain imaging studies have provided important information on the likely source of savant abilities. Specific areas of the brain that have known functions and that are influenced by various anatomical and/or developmental variations have been found. The picture that is emerging is one that provides a plausible set of possible brain compensation and regeneration processes for savant syndrome and some of its more specific manifestations. In a recent review of brain imaging and related research with savant syndrome and autism spectrum disorders, Bokkon et al., (2013) proposed a right hemisphere visual processing hypothesis to help explain common patterns of activity found among a variety of participants in several studies. Quoting Kunda and Goel (2011), Bokkon and colleagues conclude that "certain individuals with autism may 'think visually' [and] should be taken seriously as a cognitive model and receive more focused and sustained attention in behavioral and neurobiological experiments" (p. 75). This conclusion may apply to savant syndrome as well.

Savant syndrome is often associated with left-brain dysfunction (specifically left anterior temporal lobe or LATL), which leads to right brain compensation. The conditions can appear very early, even prenatally, or they can appear later as in cases of frontotemporal dementia (FTD) when the functions of a normal brain deteriorate as part of the aging process. In most right-handed individuals, this part of the brain is responsible for language and semantic processing, symbolic representation, and reflection. For the savant, the absence, diminishing, or deterioration of these functions is associated with the kinds of activity characteristic of the savant, particularly the autistic savant.

One way to test whether this interpretation of brain functions (LATL) involved in savant syndrome may be correct is to artificially suppress normal brain functioning through repetitive transcranial magnetic stimulation (rTMS; Snyder et al., 2003) of the suspected areas. Results from such studies have shown that savant skill–related capabilities often increased under these conditions (Bokkon et al., 2013; Snyder, 2009).

Although the number of studies of brain functioning and brain-related events in savant behavior is still small compared with research into other aspects of intelligence, the techniques and technologies are promising and advancing rapidly, making it likely that more results will be forthcoming. We should know a great deal more about the brains of savants and others with savant-like skills in the not too distant future (Treffert, 2009, 2014).

The Interplay of General and Specific Intellectual Abilities: Transcending the General Versus Specific Intelligence Issue

Given these findings, it appears that a picture of the way in which various degrees and varieties of intelligence interact to produce both prodigies and savants is emerging. In this respect, research with extreme cases has shed light on the long-standing debate between advocates of a more general interpretation of intelligence (typically IQ) and those who favor a more multiple intelligence–oriented view (e.g., Gardner's [1983] multiple intelligences (see also Gardner, Kornhaber, & Wake [1996], Sternberg's [1985] triarchic theory). In this final section, we summarize how more general and more specific forms of intelligence jointly contribute to the appearance of the kinds of individuals we have called prodigies and savants.

If we assume that human evolution of intellectual abilities has had variations and redundancies built into the system over time, as is true of other species, it seems likely that our brains include more than one way to respond to the challenges of our environments (Snyder, 2009). Most of our primate ancestors were specialized to habitat (although importantly not all; cf. Bruner, 1971). For humans, however, a distinctive feature of our evolution has been that it has equipped us to adapt to and thrive in highly varied environments. What we call general intelligence seems to be one of the main sources of this distinctly human capability (Feldman, 2003).

The tendency of evolution to "hedge its bets" with many variations and combinations of general and specific abilities helps explain humanity's selective advantage over its competitors for resources (Feldman with Goldsmith, 1986). The extreme examples of specific ability without general support from IQ (most savants) is an example of "niche" evolution that produced people capable of keeping track of the calendar, of telling the time, of remembering names and locations, of calculating sums in important transactions, of carrying and sharing cultural traditions such as stories, songs, and poems, and no doubt many other narrowly circumscribed and specific abilities. A savant may be anachronistic given modern technologies for doing the things that they were uniquely able to do historically, but they point to a natural source of specialized talent.

A picture is emerging of intelligences as varying along a continuum of general to specific, with numerous possibilities for combinations that reveal how these combinations may have evolved and how they have been utilized through history. Physical evolution appears to have produced both general (IQ-like) and highly specific (savant-like) abilities; in some individuals a given individual may possess one or the other kind of intelligence and others may be blessed with substantial doses of both. Perhaps an extremely high-IQ individual with no specific talents might tend to function primarily using general, abstract, logical reasoning, while the most constrained savants (e.g., those who can say the day of the week of any date on the calendar) reflect a tendency to evolve highly specific cognitive skills. Depending on their strength, the degree of general versus specific abilities, and their interaction, a prediction can be made about the possible outcome for a given person, especially at the extremes (Feldman, 1999, 2003, 2016; Morelock, 2013).

For individuals who have low (30–50 IQ or so) general ability, but who have a powerful specific ability in a particular area (e.g., music), the probability of a musical savant is likely (given availability of appropriate technology and exposure), but more creative musical ability may prove difficult if not impossible. For individuals with moderate impairment of general ability (50–80 IQ or so), a musical savant, with appropriate encouragement and support (Treffert, 2009, 2013, 2014), may be capable of improvisation and creative expression comparable to that of a professional musician. For individuals whose general abilities are in the average range (80–110 IQ or so), the kinds of achievements that are associated with prodigies may be possible in some fields (like music and visual art). For individuals whose general abilities (IQ 120–150) are exceptional, along with strong interests and abilities in certain specific areas (e.g., physics, mathematics), the probabilities of becoming notable achievers in those fields are substantial (Simonton, 1999).

Inspired by the study of prodigies and savants and the ways in which general and specific intelligences are involved in their amazing accomplishments, a coherent interpretation of human abilities has begun to emerge. The issue of general versus specific ability can now be transcended and replaced by an integrated, dynamic view that turns on the interplay among general and specific intelligences as they express themselves in social, cultural, historical, and evolutionary contexts (Feldman, 2016).

References

Baeck, E. (2002). The neural networks of music. *European Journal of Neurology*, 9(5), 449–460.

Bates, E. (1997). On language savants and the structure of the mind. *International Journal of Bilingualism*, *1*(2), 163–186.

Baumgarten, F. (1930). *Wunderkinder psychologische untersuchungen*. Leipzig: Johann Ambrosious Barth. (Untranslated)

Bidder, G. P. (1856). On mental calculation. *Minutes of Proceedings of the Institution of Civil Engineers*, Vol. 15, session 1855–56, February 19 and 26.

Bokkon, I., Salari, V., Scholkmann, F., Dai, J., & Grass, F. (2013). Interdisciplinary implications on autism, savantism, Asperger syndrome and the biophysical picture representation: Thinking in pictures. *Cognitive Systems Research*, *22–23*, 67–77.

Bornstein, M. H., & Krasnegor, N. A. (Eds.) (1989). *Stability and continuity in mental development: Behavioral and biological perspectives*. Hillsdale, NJ: Erlbaum.

Brody, L. E., & Stanley, J. C. (2005). Youths who reason exceptionally well mathematically and/or verbally: Using the MVCT: 4 Model to develop their talents. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 20–37). New York: Cambridge University Press.

Bruner, J. (1971). The nature and uses of immaturity. *American Psychologist*, *27*, 1–22.

Butterworth, B. (2001). What makes a prodigy? *Nature Neuroscience*, *4*(1), 11–12.

Campbell, J. I. D. (2005). *Handbook of mathematical cognition*. New York: Psychology Press.

Clark, E. F., & Linfoot, J. J. (1983). George Parker Bidder: The calculating prodigy. *Institute of Mathematics and Its Applications*, *23*, 68–71.

Clynes, T. (2015). *The boy who played with fusion*. New York: Houghton Mifflin Harcourt.

Comeau, G., Lu, Y., Swirp, M., & Mielke, S. (2017a). Measuring the musical skills of a prodgy: A case study. *Intelligence*, *66*, 84–97.

Comeau, G., Vuvan, D. T., Picard-Deland, C., & Peretz, I. (2017b). Can you tell a prodigy from a professional musician? *Music Perception*, *35*, 199–209.

Conway, F., & Siegelman, J. (2005). *Dark hero of the information age: In search of Norbert Wiener, the father of cybernetics*. New York: Basic Books.

Cowan, R., Stainthorp, R., Kapnogianni, S., & Anastasiou, M. (2004). The development of calendrical skills. *Cognitive Development*, *19*(2), 169–178.

Edmunds, A. L., & Noel, K. A. (2003). Literary precocity: An exceptional case among exceptional cases. *Roeper Review*, *25*(4), 185–194.

Ericsson, K. A. (Ed.) (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games*. Mahwah, NJ: Erlbaum.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Feldman, D. H. (1995). Intelligence in prodigies. In R. Sternberg (Ed.), *Encyclopedia of intelligence* (pp. 845–850). New York: Macmillan.

Feldman, D. H. (1999). A developmental, evolutionary perspective on gifts and talents. *Journal for the Education of the Gifted*, *22*(2), 159–167.

Feldman, D. H. (2000). Figurative and operative processes in the development of artistic talent. *Human Development*, *43*, 60–64.

Feldman, D. H. (2003). A developmental, evolutionary perspective on gifts and talents. In J. Borland (Ed.), *Rethinking gifted education* (pp. 9–33). New York: Teachers College Press.

Feldman, D. H. (2008). Prodigies. In J. Plucker & C. Callahan (Eds.), *Critical issues and practices in gifted education* (pp. 501–512). Waco, TX: Prufrock Press.

Feldman, D. H. (2016). Two roads diverged in the musical wood: A coincidence approach to the lives and careers of Nyiregyhazi and Menuhin. In G. McPherson (Ed.), *Musical prodigies: Interpretations from psychology, education, musicology, and ethnomusicology* (pp. 115–133). Oxford: Oxford University Press.

Feldman, D. H., with Goldsmith, L. T. (1986). *Nature's gambit: Child prodigies and the development of human potential*. New York: Basic Books.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Gardner, H. (1983). *Frames of mind*. New York: Basic Books.

Gardner, H., Kornhaber, M., & Wake, W. (1996). *Intelligence: Multiple perspectives*. Fort Worth, TX: Holt, Rinehart and Winston.

Gladwell, M. (2008). *Outliers: The story of success*. New York: Little Brown.

Goldsmith, L. T. (1987). Girl prodigies: Some evidence and some speculations. *Roeper Review*, *10*(2), 74–82.

Goldsmith, L. T. (2000). Tracking trajectories of talent: Child prodigies growing up. In R. C. Friedman & B. M. Shore (Eds.), *Talents unfolding: Cognition and development* (pp. 89–118). Washington: American Psychological Association.

Hermelin, B., & O'Connor, N. (1986). Idiot savant calendrical calculators: Rules and regularities. *Psychological Medicine*, *16*, 1–9.

Hildesheimer, W. (1982/1977). *Mozart*. New York: Vintage Books.

Hill, A. L. (1977). Idiots-savants: Rate of incidence. *Perceptual and Motor Skills*, *44*, 161–162.

Ho, W. (Ed.) (1989). *Yani: The brush of innocence*. New York: Hudson Hills Press.

Hollingworth, L. (1942). *Children above 180 IQ*. Yonkers-on-Hudson, NY: World Book. (Reprinted by Arno Press, 1975.)

Hou, C., Miller, B., Cummings, J., Goldberg, M., Mychack, P., Bottino, B., & Benson, F. (2000). Artistic savants. *Neuropsychiatry*, *13*, 29–38.

Howard, R. W. (2008). Linking extreme precocity and adult eminence: A study of eight prodigies at international chess. *High Ability Studies*, *19*(2), 117–130.

Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences*, *21*, 399–406.

Hulbert, A. (2005). The prodigy puzzle. *New York Times Magazine*, November 30, 64–71.

Johnson, R. (2005). A genius explains. *The Guardian*, February 12. www.theguardian.com/theguardian/2005/feb/12/weekend7.weekend2

Kanigel, R. (1991). *The man who knew infinity: A life of the genius Ramanujan*. New York: Washington Square Press.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Kearney, K., & Kearney, C. (1998). *Accidental genius*. Juneau, AK: Woodshed Press.

Kenneson, C. (1998). *Musical prodigies: Perilous journeys, remarkable lives*. Portland, OR: Amadeus Press.

Kunda, M., & Goel, A. (2011). Thinking in pictures as a cognitive account of autism. *Journal of Autism and Developmental Disorders*, *41*, 1157–1177.

Leites, N. S. (1960). *Intellectual giftedness*. Moscow: APN Press.

Leites, N. S. (Ed.) (1996). *Psychology of giftedness of children and adolescents*. Moscow: Academia.

Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science*, *1*(4), 316–345.

Lubinski, D., Benbow, C. P., & Morelock, M. J. (2000). Gender differences in engineering and physical sciences among the gifted: An inorganic-organic distinction. In K. Heller, F. Mönks, R. Sternberg, & R. Subotnik (Eds.), *International handbook of giftedness and talent* (2nd ed., pp. 633–648). New York: Pergamon Press.

Lubinski, D., Webb, R. M., Morelock, M. J., & Benbow, C. P. (2001). 1 in 10,000: A longitudinal study of the profoundly gifted. *Journal of Applied Psychology*, *86*, 718–729.

McPherson, G. E. (Ed.) (2006). *The child as musician: A handbook of musical development*. Oxford: Oxford University Press.

McPherson, G. E. (2007). Diary of a child prodigy musician. In A. Williamson & D. Coimbra (Eds.), *Proceedings of the International Symposium on Performance Science 2007* (pp. 213–218). Porto, Portugal: Association of European Conservatories.

Milbrath, C. (1998). *Patterns of artistic development in children: Comparative studies of talent*. New York: Cambridge University Press.

Miller, B. L., Boone, K., Cummings, L. R., & Mishkin, F. (2000). Functional correlates of musical and visual ability in frontotemporal dementia. *British Journal of Psychiatry*, *176*, 458–463.

Miller, B. L., Cummings, J., Mishkin, F., Boone, K., Prince, F., Ponton, M., et al. (1998). Emergence of artistic talent in fronto-temporal dementia. *Neurology*, *51*, 978–982.

Miller, L. K. (1989). *Musical savants: Exceptional skill in the mentally retarded*. Hillsdale, NJ: Erlbaum.

Miller, L. K. (1999). The savant syndrome: Intellectual impairment and exceptional skill. *Psychological Bulletin*, *125*(1), 31–46.

Miller, L. K. (2005). What the savant syndrome can tell us about the nature and nurture of talent. *Journal for the Education of the Gifted*, *28*(3–4), 361–374.

Morelock, M. J. (1995). The profoundly gifted child in family context. Unpublished doctoral dissertation, Tufts University, Medford, MA.

Morelock, M. J. (2013). Prodigies, passion, persistence, and pretunement: Musings on the biological bases of talent. In S. B. Kaufman (Ed.), *The complexity of greatness: Beyond talent or practice* (pp. 83–101). New York: Oxford University Press.

Morelock, M. J., & Feldman, D. H. (1993). Prodigies and savants: What they have to tell us about giftedness and human cognition. In K. A. Heller, F. J. Monks, & A. H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp. 161–181). Oxford: Pergamon Press.

Morelock, M. J., & Feldman, D. H. (1999). Prodigies. In M. Runco & S. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 1303–1320). San Diego, CA: Academic Press.

Morelock, M. J., & Feldman, D. H. (2003). Prodigies, savants and Williams syndrome: Windows into talent and cognition. In F. J. Monks, K. A. Heller, R. J. Sternberg, & R. Subotnik (Eds.), *International handbook for research on giftedness and talent* (2nd ed., pp. 455–469). Oxford: Pergamon Press.

Mursell, J. (1937). *The psychology of music*. New York: W. W. Norton.

O'Boyle, M. (2008a). Adolescent psychopathology and the developing brain. *Journal of Youth and Adolescence*, *37*, 481–483.

O'Boyle, M. (2008b). Mathematically gifted children: Developmental brain characteristics and their prognosis for well-being. *Roeper Review*, *30*, 181–186.

O'Connor, N. (1989). The performance of the "idiot savant": Implicit and explicit. *British Journal of Disorders of Communication*, *24*, 1–20.

O'Connor, N., & Hermelin, B. (1984). Idiot savant calendrical calculators: Math or memory? *Psychological Medicine*, *14*, 801–806.

O'Connor, N., & Hermelin, B. (1987). Visual and graphic abilities of the idiot savant artist. *Psychological Medicine*, *17*, 79–80.

Patel, A. D. (2008). *Music, language and the brain*. New York: Oxford University Press.

Peek, F. (1997). *The real Rain Man*. Salt Lake City, UT: Harkness.
Peek, F., with Hanson, L. L. (2007). *The life and message of the real Rain Main: The journey of a mega-savant*. Port Chester, NY: Dude Publishing/National Professional Resources.
Pring, L., Ryder, N., Crane, L., & Hermelin, B. (2012). Creativity in savant artists with autism. *Autism*, *16*, 45–57.
Radford, J. (1990). *Child prodigies and exceptional early achievers*. New York: Free Press.
Revesz, G. (1916/1970). *The psychology of a musical prodigy*. Freeport, NY: Books for Libraries Press.
Rolfe, L. (1978). *The Menuhins: A family odyssey*. San Francisco: Panjandrum Books.
Ruthsatz, J., & Detterman, D. K. (2003). An extraordinary memory: The case study of a musical prodigy. *Intelligence*, *31*, 509–518.
Ruthsatz, J., & Stephens, K. (2016). *The prodigy's cousin*. New York: Penguin Random House.
Ruthsatz, J., & Urbach, J. B. (2012). Child prodigy: A novel cognitive profile places elevated general intelligence, exceptional working memory and attention to detail at the root of prodigiousness. *Intelligence*, *40*, 419–426.
Sacks, O. (1995). *An anthropologist on Mars*. New York: Alfred A. Knopf.
Scheerer, M., Rothmann, E., & Goldstein, K. (1945). A case of "idiot savant": An experimental study of personality organization. *Psychology Monograph*, *58*, 1–63.
Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995a). In vivo evidence of structural brain asymmetry in musicians. *Science*, *267*, 699–701.
Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995b). Increased corpus callosum size in musicians. *Neuropsychologica*, *33*, 1047–1055.
Selfe, L. (1977). *Nadia: A case of extraordinary drawing ability in an autistic child*. New York: Academic Press.
Shavinina, L. (1999). The psychological essence of the child prodigy phenomenon: Sensitive periods and cognitive experience. *Gifted Child Quarterly*, *43*(1), 25–38.
Simonton, D. K. (1994). *Greatness: Who makes history and why*. New York: Guilford Press.
Simonton, D. K. (1999). Talent and its development: An emergenic and epigenetic model. *Psychological Review*, *106*(3), 435–457.
Singh, H., & O'Boyle, M. W. (2004). Inter-hemispheric interaction during global-local processing in mathematically gifted adolescents, average-ability youth, and college students. *Neuropsychology*, *18*(2), 371–377.
Smith, N. V., & Tsimpli, I. (1995). *The mind of a savant: Language learning and modularity*. Cambridge, MA: Blackwell.
Smith, S. B. (1983). *The great mental calculators: The psychology, methods, and lives of calculating prodigies past and present*. New York: Columbia University Press.
Snyder, A. (2009). Explaining and inducing savant skills: Privileged access to lower level, less processed information. *Philosophical Transactions of the Royal Society*, *364*, 1399–1405.
Snyder, A., Mulcahy, E., Taylor, J., Mitchell, D., Sachdev, P., & Gandevia, S. (2003). Savant-like skills exposed in normal people by suppressing the left fronto-temporal lobe. *Journal of Integrative Neuroscience*, *2*, 149–158.
Solomon, A. (2012). *Far from the tree: Parents, children, and the search for identity*. New York: Scribner's.
Stanley, J. C. (1996). SMPY in the beginning. In C. P. Benbow & D. Lubinski (Eds.), *Intellectual talent: Psychometric and social issues* (pp. 225–235). Baltimore, MD: Johns Hopkins University Press.

Stanley, J. C. (2000). Helping students learn only what they don't already know. *Psychology, Public Policy, and Law*, *6*, 216–222.

Sternberg, R. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Tammet, D. (2006). *Born on a blue day: Inside the extraordinary mind of an autistic savant*. New York: Free Press.

Tammet, D. (2009). *Embracing the wide sky: A tour across the horizons of the mind*. New York: Free Press.

Tannenbaum, A. (1993). History of giftedness and "gifted education" in world perspective. In K. A. Heller, F. J. Monks, & A. H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp. 3–27). Oxford: Pergamon Press.

Thioux, M., Stark, D. E., Klaiman, C., & Schultz, R. T. (2006). The day of the week when you were born in 700 ms: Calendar computation in an autistic savant. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1155–1168.

Thompson, R., & Nelson, C. (2001). Developmental science and the media. *American Psychologist*, *56*(1), 5–15.

Treffert, D. (1989). *Extraordinary people: Understanding "idiot savants*." New York: Harper & Row.

Treffert, D. (2006). *Extraordinary people: Understanding savant syndrome*. Omaha, NE: iUniverse.

Treffert, D. (2008). Myths that persist: Savant syndrome 2008. www.wisconsinmedicalsociety.org/professional/savant-syndrome/resources/articles/myths-that-persist-savant-syndrome-2008

Treffert, D. (2009). The savant syndrome: An extraordinary condition. A synopsis: Past, present, future. *Philosophical Transactions of the Royal Society*, *364*, 1351–1357.

Treffert, D. (2013). Savant syndrome: A compelling case for innate talent. In S. B. Kaufman (Ed.), *The complexity of greatness: Beyond talent or practice* (pp. 103–118). New York: Oxford University Press.

Treffert, D. (2014). Savant syndrome: Realities, myths and misconceptions. *Journal of Autism and Developmental Disorders*, *44*, 564–571.

Tsimpli, I., & Smith, N. (1999). Modules and quasi-modules: Language and theory of mind in a polyglot savant. *Learning and Individual Differences*, *10*(3), 193–215.

Viscott, D. S. (1970). A musical idiot savant. *Psychiatry*, *33*, 494–515.

Wallace, A. (1986). *The prodigy: A biography of William James Sidis, America's greatest child prodigy*. New York: Dutton.

Wiener, N. (1953). *Ex-prodigy: My childhood and youth*. Cambridge, MA: MIT Press.

Winner, E. (1982). *Invented worlds*. Cambridge, MA: Harvard University Press.

Winner, E. (1996). The rage to master: The decisive role of talent in the visual arts. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games* (pp. 271–301). Mahwah, NJ: Erlbaum.

Young, R. (1995). Savant syndrome: Processes underlying extraordinary abilities. Unpublished doctoral dissertation, University of Adelaide, South Australia.

Zimmerman, R. (Writer) (1989). *A Real Rainman* [VHS Film]. Maple Plain, MI: Simitar Entertainment.

13 Intellectual Giftedness

Sally M. Reis and Joseph S. Renzulli

The study of gifts and talents and how various innate abilities interact with one's environment, personality, educational opportunities, family support, and life experiences has fascinated psychologists, educators, and parents for decades. Many are intrigued with the reasons that one child with remarkably high potential born into a particular family in a particular environment grows up to become a neurosurgeon while a child of similar intellectual potential who lives in the same community and attends the same schools decides to drop out of high school. What have researchers and scholars learned in the past few decades about the nature of talent development and intellectual giftedness? What general concepts are widely accepted about intellectual giftedness? How is it defined and how can it be developed? What combinations of genetic abilities and talents interact with one's personality and environment to produce intellectual giftedness?

In this chapter, these questions, none of which can be answered simply, are discussed and current research about intellectual giftedness is summarized. As the research reviewed in this chapter points out, core concepts about giftedness relate to the diversity of expression of gifts and talents, as there are no more varied groups of people than those labeled intellectually gifted (Pfeiffer, Shaunessy-Dedrick & Foley-Nicpon, 2018). Those labeled gifted as children and/or adults are found in every ethnic and socioeconomic group and culture. They exhibit an unlimited range of personal and learning characteristics and differ in effort, temperament, educational and vocational attainment, productivity, creativity, risk taking, introversion, and extraversion (Renzulli & Park, 2002; Renzulli & Reis, 2014). They have variable abilities to self-regulate and sustain the effort needed to achieve personally, academically, and in their careers (Housand & Reis, 2009). And despite the label that this diverse population has been given, within the population some do and some do not demonstrate high levels of accomplishment in their education or their chosen professions and work (Reis & McCoach, 2000; Renzulli & Park, 2002). And some simply are not willing to devote the time and focus necessary to develop and realize their talents, for it is clear to those who study giftedness that developing extraordinary levels of talent requires extraordinary levels of effort and focus, beyond that which is expended by individuals' intellectual and educational peers.

Despite the broad diversity within the population, however, several common themes emerge about intellectual giftedness and the academic and work conditions necessary for its development. We begin our review of research related to intellectual giftedness with a discussion of these themes, summarizing highlights about research on intellectual giftedness in the United States, including some work that is considered seminal, and presenting an overview of what we believe to be some interesting and potentially important theories to date. We conclude the chapter with some interesting new research-based trends related to emerging ideas in defining and developing academic gifts and talents. It is important to understand, however, that scholars and researchers continue to debate definitions of giftedness, how intellectual giftedness develops, and the characteristics of diverse groups of high-potential youth that will help educators and psychologists to identify and nurture intellectual gifts and talents. To illustrate the challenges associated with both defining and identifying giftedness in students, four brief case studies are introduced.

Four Case Studies

Andrew was identified as a gifted student in first grade. Highly verbal and the son of two university professors, he read at age four, was exceptionally analytical, and excelled in nursery school and first grade, particularly in his verbal skills. His energy and enthusiasm for learning were identified by all of his teachers and both his kindergarten and first grade teacher referred him for the gifted program in his school despite the fact that formal identification for most students did not usually occur until fourth grade. Andrew excelled in the primary grades, but with each year that passed, he struggled more with schoolwork that depended on his ability to write. In fourth grade, despite very high abilities, he had begun to express his difficulties in writing. At this point, his classroom teacher suspected that Andrew might have a learning disability and discussed dysgraphia with his parents for the first time. Dysgraphia, a learning disability connected to the graphomotor aspect of writing, is often identified by examining and evaluating writing samples for word and letter spacing (e.g., how and if the letters fit on the line and the quality of what is written). Students with dysgraphia often struggle with holding pencils and writing for long periods of time. Andrew's teachers described behaviors such as shaking his hands and constantly stretching and rubbing his hands, wrists, or fingers while writing. Andrew began to use overly simplistic language and very short sentences in his minimal writing. When questioned orally, he responded with fluency and insight, but when he had to write in class, his work resulted in short, stilted responses with limited description. As Andrew matured, his lack of attention in class and academic struggles intensified, despite his scores at the 99th percentile in IQ on both verbal and figural areas. His fourth-grade teacher and the special education teacher suggested a series of academic recommendations in both special and gifted education as part of an individual education plan for Andrew.

Sofía was in second grade when her teachers recommended her for participation in the gifted program. She was highly verbal and was reading independently at the seventh-grade level, excelling in every aspect of her academic work. Gifted program participation in her school was not dependent on IQ test scores, and Sofía was identified based on her achievement tests (99th percentile in all academic areas), teacher nominations, leadership and creativity, and classroom work. Sofía was a high-achieving student throughout elementary and secondary school and graduated in the top three of her class, earning entrance to an Ivy League university. However, prior to her freshman year in high school, her parents moved and she transferred to a new school district that required an IQ test for formal identification as gifted. Her score was 119, well below the cutoff for gifted program entrance in the new school district. Despite being a star in the gifted program in her former district, she was denied entrance to the program in her new district. Sofía, however, excelled in all of her Advance Placement (AP) and honors courses, graduated in the top five of her class, scored over 1500 on the SATs, completed a complex and highly evaluated senior year project, and ultimately entered and graduated as a dean's list student at a highly prestigious university.

Kendra was a shy, quiet fifth-grader who had been identified as gifted in second grade in a school in which a 130 cutoff score on an individually administered IQ test was used to determine which students would be identified for and participate in the gifted program. An avid reader and introvert, she displayed few characteristics related to most traditional notions of giftedness. Although she loved to read, she did not initially appear to display verbal precocity. Her teachers did not report observing any characteristics of intellectual advancement, nor of problem-solving, reasoning, insight, or other commonly acknowledged characteristics of academic giftedness. Kendra was primarily known for being quiet, kind, and an advanced reader who did not like to discuss or share what she was reading, perhaps due to her shyness. As she grew up, she remained a quiet and passive learner who, despite her intelligence, rarely spoke in class and achieved well but was not outstanding in any one particular area.

Jelani was identified as gifted in third grade; however, his schoolwork and grades frustrated both his parents and teachers for years following his identification and placement in a gifted program. Always a child of very high potential, Jelani's grades fluctuated in elementary, middle, and senior high school. To qualify as gifted in his district, Jelani had to achieve an IQ score above 130 on an aptitude assessment in addition to demonstrating high achievement in the classroom. He enjoyed discussing his ideas with others and was highly verbal, but he had poor work habits in required subjects. As the years progressed, Jelani's work became less and less impressive, and his teachers questioned his identification as gifted. His writing was considered below average and the only class in which he consistently excelled was math. Jelani disliked reading anything that was unrelated to his interests. His grades varied, from top marks in math and technology to failing grades in subjects that did not interest him. Although he took advanced math classes in middle and high school and achieved a near-perfect score on the math section of the SAT, during his junior year of high school, Jelani's teachers and parents labeled him an "underachiever" because of his fluctuating performance in, and attitudes about, school. He rarely displayed characteristics of a gifted student in classes in which he did not have an interest. His technology and math teachers realized his potential and saw his talents in problem-solving, persistence, and creativity. Few other teachers noted any positive characteristics and he continued to underachieve in school, attaining below-average grades.

Common Themes Related to Intellectual Giftedness

As these brief case studies illustrate, giftedness and high potential are manifested differently in various young people and it is challenging to determine who should and should not be identified as intellectually gifted, especially in childhood. No standard pattern of intellectual giftedness and talent exists among high-potential children. No specific formula identifies the "right" combination of aptitude, and home and school environmental prompts and supports, needed to enhance or produce intellectual giftedness. We know some, but not all, of the most likely combinations of genetic traits and environment interactions that produce a desired outcome, such as the development of specific talents or gifts. Access to opportunities also matters, as it is widely suggested that increasing opportunities will increase achievements (Ceci & Papierno, 2005; Gagne, 2000; Renzulli & Reis, 2014). A child with high scientific aptitude, who likes science, and whose parents are scientists, will have more opportunities, resources, and encouragement in science than a child with the same cognitive aptitude who does not like science and whose parents do not have similar patterns of education and interest in this area. The child with interest in and parental support for science is, of course, more likely to seek a college degree and perhaps a career in this area but may be just as likely to pursue other interests. However, the nuances related to the development of intellectual giftedness are many and varied, and the child with high aptitude, interest, and parental support may subsequently encounter negative school experiences in science, deflating her interests and derailing her from the science pipeline. If positive elementary and secondary school experiences enhance scientific interests, negative college experiences (e.g., a first low grade in organic chemistry or an understanding of the struggles associated with earning a PhD and finding work in research in this field) may also lower or change aspirations and careers choices. Gifts and talents emerge in conjunction with a series of environmental events and personality variables, and, of course, chance factors (Tannenbaum, 1991) and the desire and work ethic to develop one's potential, despite obstacles and barriers (Reis, 2005).

Any discussion of intellectual giftedness must acknowledge the importance of factors such as environment, opportunities, and chance in the development of this construct. This is even true in persons of the highest levels of cognitive ability, as suggested by Lubinski, Benbow, and Kell (2014) and Lubinski and colleagues (2001), who found variability in the accomplishments of this group. Lubinski and colleagues (2001) investigated the patterns of those in the top 1 percent or higher of cognitive abilities and identified variation in both development trajectories and important life accomplishments. They found that the likelihood of earning a doctorate, earning exceptional compensation, publishing novels, securing patents, and earning tenure at a top university varied as a function of the individual differences in childhood cognitive abilities assessed decades earlier, suggesting the need to study the importance of both genetic and environmental origins of exceptional abilities, a finding also discussed by Terman decades earlier (1925–1959).

In this current review of research on intellectual giftedness, several important themes emerge. The first is that giftedness is an open, dynamic, intentional system that is capable of building increasingly complex behaviors through self-organization and self-direction (Dai, 2010; Dai & Renzulli, 2008; Renzulli, 2005, 2012). Themes that guide this chapter include the many ways in which intellectual giftedness develops; the multipotentiality of giftedness; the ways that culture defines and influences giftedness; and the presence and importance of nonintellectual components of intellectual giftedness in determining performance. Other themes relate to the assessment of intellectual giftedness which, according to Sternberg (2004), is too often validated almost exclusively against the societally approved criteria resulting in appearance of validity that may not exist within a specific sociocultural group, and the importance of understanding that there is no right or wrong way to define intellectual giftedness. Some theorists believe that we can identify gifted individuals across domains, even in children at a young age, as if there is a golden chromosome that enables one to be identified with the right assessment tools. Others believe that giftedness occurs within a domain, such as those who are scientifically or mathematically gifted. Different conceptions of giftedness across cultures (Phillipson & McCann, 2007) suggest emerging research and understandings of the ways in which languages and cultures influence and contribute to giftedness.

Recently, in a thoughtful and scholarly examination, Dai (2010) systematically summarized definitions of giftedness and proposed a new framework for the field of gifted education by identifying essential tensions that revolve around three core questions. These questions relate to what we know about the respective roles of natural ability, environment and experiences, and personal effort in talent development, how we identify gifted and talented students and how we assess and research the process of gifted and talent development. Dai was especially interested in how the goals of gifted education are defined and implemented to promote excellence. Dai suggests a contextual, developmental approach as a more viable alternative to the traditional psychometric one often used in schools and challenges researchers and practitioners in the field to move beyond tensions between defining and identifying gifted children and toward the current focus on talent development. The themes highlighted below emerge across many contemporary conceptions of giftedness, highlighting the need for a contextual developmental discussion, and illustrating the challenges associated with both defining giftedness and identifying intellectually gifted individuals.

Intellectual Giftedness Is Developmental

Over four decades ago, Renzulli summarized research suggesting that giftedness exists in certain people, at certain times, and under certain circumstances (Renzulli, 1978, 2005). This notion of giftedness argues against considering giftedness as a trait such as eye color or something that a child has or does not possess. Currently, increasing numbers of other researchers also support developmental constructs of giftedness. For example, Gagne's (2000) Differentiated Model of Giftedness and Talent (DMGT) is another developmental theory that distinguishes giftedness from talent and discusses how outstanding natural abilities (gifts) can develop into specific expert skills (talents). Gagne believes that those labeled as gifted have the potential for extraordinary work and that those who are subsequently identified as talented develop their inherent potential for contributions. He identifies six components that interact in multiple ways to foster the transition of moving from having natural abilities (giftedness) to systematically developed skills. These components include the gift itself, chance, environmental catalysts, intrapersonal catalysts, learning/practice, and the outcome of talent (Gagne, 2000). Many of the chapter authors in two seminal books on conceptions of giftedness edited by Sternberg and Davidson (1986, 2005) also identify similar themes related to the developmental nature of intellectual giftedness. Csikszentmihalyi, Rathunde, and Whalen (1993) concur on the developmental nature of talents, as discussed later in the section Interesting Directions in Intellectual Giftedness and Talent Development, as do Dai (2010) and Subotnik, Olszewski-Kubilius, and Worrell (2011, 2012). One of the leading scholars in the area of intelligence, Simonton (2005, 2008), for example, proposed a model of giftedness in which talents result from the coming together of genetic components that develop on individual domain-specific trajectories. These genetic components include any and all characteristics needed to develop a particular gift, such as superior visual-spatial skills or a high degree of mathematical creativity in gifted mathematicians. Simonton suggested further that the absence or late development of a key trait would prevent or delay the development of a given talent. This model provides an explanation for why individuals begin to demonstrate talents at different times, and why certain types of talents emerge earlier while others emerge later in life.

Intellectual Giftedness Is Multidimensional

Few, if any, researchers or theorists who have studied intelligence or intellectual giftedness continue to believe that giftedness is unidimensional rather than multidimensional. Similar to psychologists who believe in the multidimensional aspects of intelligence (Carroll, 1993; Gustafsson & Undheim, 1996), theorists who study intellectual giftedness (Gagne, 2000; Gardner, 1983, 2016a; Renzulli, 2005; Sternberg, 1985, 2005) agree that we must look beyond the traditional early notions stating that intellectual giftedness can be equated with a high score on one assessment such as an IQ test. In fact, recent research on assessment has found that large, significant discrepancies among verbal, figural, and quantitative reasoning abilities as measured by standardized IQ tests are more common among high- and low-ability students than among average-ability students (Lohman, Gambrell, & Lakin, 2008; Shavinia, 2001; Sternberg, 2005). Lohman and colleagues (2008), for example, examined the score profiles of students obtaining stanine scores of 9 on at least two batteries of a standardized achievement test. They found that the percentage of these highly able students demonstrating an “extreme” or significant weakness in at least one of the three tested areas – verbal, spatial, or quantitative reasoning – was equal to the percentage of students with more even profiles. They noted that this finding suggests that gifted programs using a single composite IQ score for identification may miss many highly able students whose scores are brought down by a single area of relative weakness. In summary, the most influential theorists, Gardner, Sternberg, and Renzulli, conceptualize intelligence as a dynamic system that is a product of abilities, contextual influences, and complex mental processes that interact to produce giftedness, meaning that identification should be based on a broader spectrum of measurements, such as tests, teacher and parent nominations, and student work portfolios.

Several multiple conceptions of intellectual giftedness have been suggested by many researchers; these range from general, broad, and overarching characterizations to more specific definitions of giftedness identified by *specific* actions, products, or abilities within certain domains (Sternberg & Davidson, 2005). This research, generally conducted during the past few decades, supports a more broad-based conception of giftedness as a combination of multiple qualities, in addition to intellectual potential, which includes nonintellectual traits such as motivation and creativity (Renzulli, 1978, 2005; Sternberg & Lubart, 1995) and positive beliefs in self (Reis, 2005).

Intellectual Giftedness Is Influenced by Culture, Gender, and Environment

As illustrated by the case studies and earlier discussion in this chapter, those labeled intellectually gifted are a varied group with differing cognitive profiles, learning disabilities, attention deficits, varied learning styles, issues related to procrastination and perfectionism, and faster or slower processing speeds. This diverse group is influenced by many conditions and may demonstrate asynchronous (uneven) development, cognitive and/or academic relative strengths and weaknesses, or learning disabilities (Reis, Neu, & McGuire, 1997). Sternberg's (2004) work suggests that different patterns of giftedness exist and change over time and his research on the influence of culture and environment is present across his latest publications, as discussed in the section about his work, entitled Triarchic Theory, Balance Theory of Wisdom, and Successful Intelligence Applied to Cognitive Giftedness.

The notion of intellectual giftedness itself has different meanings for different people, and discussions about these meanings are influenced by the culture, environment, and context in which the gifts emerge as well as the values associated with each (Simonton, 1998). Not surprisingly, within different cultures, contexts, and environments, the outcomes of intellectual giftedness vary. Cultural influences can negatively or positively affect the choices and products that emanate from one's gifts, and the ability to select, shape, and/or adapt one's environment (Sternberg, 1996, 2004).

Gender also has an impact on giftedness, as little doubt exists that gifted males in many cultures far surpass gifted women in accomplishment and professional attainments (Reis, 1998, 2005). Reis explored the paths leading to female talent realization in women in a study of twenty-two American women who gained eminence in diverse fields over a decade (Reis, 1998). Each eminent woman was recognized as a major contributor in her field, and several achieved the distinction of being the first or one of the first women to achieve eminence in her domain, such as theater, politics, academe, literature and poetry, science, musical composition, government, business, environmental sciences, art, education, and other fields. Reis proposed a theory of talent development in women (Reis, 2002, 2005) that includes abilities (intelligence and special talents), personality traits, environmental factors, and personal perceptions, such as the social importance of the use of one's talents to make a positive difference in the world. Underlying this theory is the belief that talent is developed in women of high potential through systematic work, active choices, and individual, sustained effort (Renzulli, 1978, 1986, 2012). Most of these women made difficult choices about their personal lives in order for their creative productivity to emerge, including whether to divorce or refrain from marrying, to forgo having children or to have fewer children than they might otherwise have had, to live alone, or any combination of these characteristics (Reis, 1998). These decisions were usually consciously made to support a lifestyle conducive to the production of highly challenging work. Within multicultural societies, it is usually the views held by the dominant culture about gender that guide the ways that giftedness is defined and measured, and research summarized in this chapter shows the links among culture, environment, and gender and the development of intellectual giftedness.

Noncognitive Aspects of Intellectual Giftedness

In addition to cognitive contributors to the development of high performance, a number of other factors referred to by Renzulli (2005) as "intelligences outside the normal curve" have also been found to play a role in the accomplishments of intellectually gifted young people and adults. Factors such as creativity, motivation, courage, optimism, sense of power to change things, empathy, and physical and mental energy are aspects of the gifts that we respect in the work of people such as Rachel Carson, Nelson Mandela, Mother Teresa, and Martin Luther King, Jr. (Renzulli, 2005). Combined with other noncognitive skills such as collaboration, leadership, organization, planning, and self-efficacy, what emerges is a picture of giftedness that extends far beyond the "golden chromosome" theory that would lead us to believe that some people are preordained to be gifted (Renzulli, 2005).

And what about creativity and the role that it plays in developing giftedness? Renzulli (2005) included creativity as one of the three clusters in his widely recognized "three-ring conception" of giftedness, calling it co-cognitive, as opposed to a noncognitive cluster. Creativity is one of the clusters of abilities that is necessary for creative productive giftedness to develop in Renzulli's definition. This type of giftedness involves the development of innovative ideas or products that will be valued by their targeted audiences. Creative-productive giftedness also requires above-average intelligence, sensible risk taking, a sense of purpose, and the motivation to work hard for a prolonged period of time. Creativity bridges the gap between childhood and adult giftedness and it involves the same capacity across one's life span, even though one's creative goals and accomplishments change with age. According to current research reviewed in this chapter, gifts and talents emerge in conjunction with a series of personality variables, and other nonintellectual factors including motivation, task commitment, focus, and desire and work ethic to develop one's potential, despite obstacles and barriers that are encountered.

Important American Contributions to Research on Intellectual Giftedness

Four seminal theoretical contributions related to research on intellectual giftedness are summarized in this section, including the historically important work of Lewis M. Terman, and the recent works of Joseph Renzulli, Howard Gardner, and Robert J. Sternberg.

Genetic Studies of Genius: Lewis M. Terman's Contributions

Lewis M. Terman edited a five-volume series entitled *Genetic Studies of Genius* between 1925 and 1959, resulting in a body of work that is widely acknowledged to be a seminal contribution to the field of intellectual giftedness. The background of the use of the word "genius" in the title stems from his publication in 1916 of the Stanford-Binet Intelligence Scale, based on the work of Alfred Binet, who had devised a scale commissioned by the French government to identify children who needed help in school. Terman conducted longitudinal research on a sample of over 1,500 boys and girls, with very few minorities, who were nominated by their teachers as highly intelligent and who subsequently scored over 140 on the Stanford-Binet Intelligence Scale. Terman and his colleagues tested these students, who may have been those who demonstrated what Renzulli has called "schoolhouse giftedness" as they performed well academically in the classroom. This procedure for selection illustrates a continuing debate related to the study of intellectual giftedness, which is how intellectual giftedness is defined and measured by various scales and tests and who is nominated by their teachers for participation in the gifted program.

Terman's research resulted in several important findings. The high-IQ children he studied longitudinally were physically and emotionally healthy, and most did well in school and college and had successful professional careers. But as Renzulli (1978) pointed out over forty years ago, the longitudinal findings of Terman's work also produced some interesting results that raise questions about how potential translates into actualized giftedness. During the period in which Terman's research was conducted, most women became homemakers rather than pursuing full-time careers and achieving college degrees, resulting in different career profiles from those of the men in his study. Also, almost one-third of the men in the sample did not realize their expected potential and might even have been labeled underachievers, as they did not complete the level of education or attain the career goals that might have been expected in their professional lives. Few in the sample would later be labeled geniuses but many did achieve eminence across various fields and domains. The problematic legacy of Terman's work is his preexisting notion that an IQ score is the same as giftedness and that the highest IQ is equal to genius. This legacy has not been helpful to the field of gifted education and those researchers, psychologists, and educators who believe that giftedness is developmental and includes noncognitive attributes such as creativity and motivation and focus.

Three-Ring Conception of Giftedness: Joseph Renzulli

For many years following the publications of Terman's work, psychologists and educators continued to equate intellectual giftedness with high scores on an intelligence or IQ test. It is important to remember that pioneers in intelligence assessment such as Binet believed that both genetic and environmental factors contributed to intellectual ability and would not have supported the subsequent practice of Terman, who equated intelligence with a number achieved on one intellectual assessment. Intelligence and measurement theory were developed simultaneously and often conflated, meaning that scores on standardized measurements of intellectual ability were widely interpreted as also measuring intelligence in the decades following Terman's work.

Renzulli's (1978) definition helped to move the focus of previous discussions from an examination of gifted individuals to an examination of gifted behaviors and suggested the inclusion of nonintellectual components in giftedness. He defined giftedness as reflecting an interaction among three basic clusters (popularly known as the three-ring conception of giftedness) of human traits – above-average ability, high levels of task commitment, and high levels of creativity – stating that individuals capable of developing gifted behavior are those possessing or capable of developing this composite set of traits and applying them to any potentially valuable area of human performance. He also distinguished between schoolhouse or high academic giftedness and creative-productive giftedness, arguing that many individuals who excel in school and are labeled gifted do not make creative contributions as adults because they lack both creativity and task commitment for creative-productive giftedness (Renzulli, 1986). His definition became widely used and adapted by some states and school districts across the country.

Renzulli (2002) continued the work on his three-ring conception by examining personality and environmental factors that contribute to socially constructive behaviors reflected in the works of people who have made contributions to the greater good in all walks of life. These interactive factors are depicted by the houndstooth background of his three-ring conception (see Figure 13.1). Renzulli identified six variables contributing to giftedness that will form the basis for his newest research on how these specific traits are manifested, the extent to which they exist, and the ways

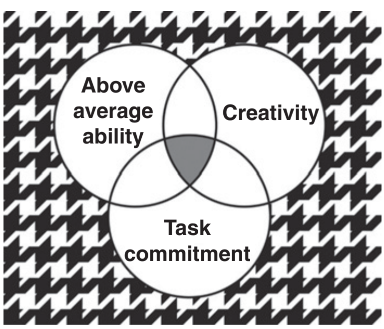

Figure 13.1 *Three-ring conception of giftedness with houndstooth background.*

they interact with one another. He believes that these variables, coupled with abilities, creativity, and task commitment, are the key to both explaining and nurturing the kind of genius that has been used for the betterment of mankind.

The first of the six variables is optimism, defined as the belief that the future holds good outcomes. Optimism can be considered an attitude associated with expectations of a future that is socially desirable, to the individual's advantage, or to the advantage of others. It is characterized by a sense of hope and a willingness to work long hours for a cause. The second variable is courage, the ability to face difficulty or danger while overcoming physical, psychological, or moral fears. Courage is characterized by integrity and strength of character, the most salient marks of those creative people who actually increase social capital. The third is romance with a topic or discipline that occurs when an individual is passionate about a topic or discipline. The passion of this romance often becomes an image of the future in young people and provides the motivation for a long-term commitment to a course of action. The fourth is sensitivity to human concerns, a trait that encompasses one's abilities to comprehend another's world and to accurately and sensitively communicate such understanding through action. Altruism and empathy also characterize this trait. The fifth is physical/mental energy, or the amount of energy an individual is willing and able to invest in the achievement of a goal, a crucial issue in high levels of accomplishment. In the case of eminent individuals, this energy investment is a major contributor to task commitment. Charisma and curiosity are frequent correlates of high physical and mental energy. The last trait Renzulli identified in his more recent work is vision/sense of destiny, which although complex and difficult to define, may best be described by a variety of interrelated concepts, such as internal locus of control, motivation, volition, and self-efficacy. When an individual has a vision or sense of destiny about future activities, events, and involvements, this vision serves to stimulate planning and becomes an incentive for present behavior.

Application of Multiple Intelligence to Gifted Contributors and GoodWork: Howard Gardner

Gardner's (1983) theory of multiple intelligences (MI) proposes seven relatively autonomous but interactive intelligences, based on his work with individuals exhibiting extreme cognitive abilities (or deficits) in particular areas, such as music or math, but not general cognitive superiority. The seven intelligences initially proposed by Gardner were linguistic, logical-mathematical, musical, spatial, bodily-kinesthetic, interpersonal, and intrapersonal. Linguistic intelligence relates to a person's ability to read, write, and speak, and along with logical-mathematical intelligence composes the traditional conception of intelligence. Musical intelligence is related to one's ability to create, communicate, and understand sound, whereas spatial intelligence is revealed through perceiving, manipulating, and recreating visual and spatial objects. Gardner's idea of bodily kinesthetic intelligence refers to the use of the body's strength, agility, balance, grace, and control of movements as demonstrated in persons such as Jackie Joyner Kersey, a well-known Olympic athlete. Interpersonal and intrapersonal intelligence both involve social skills relating to understanding emotions regarding others and the self, respectively. Naturalist intelligence, or the ability to care for and nurture living things in nature, has since been added to Gardner's theory, but is less likely to be widely accepted as the original components of MI theory (Gardner, 1995a, 2006, 2016a).

How does Gardner define intellectual giftedness? Gardner (1995b applied his MI theory to an analysis of the intelligences of creative leaders of the twentieth century, explaining that outstanding performance emanated from a particular intelligence. Gardner believed, for example, that Mahatma Gandhi excelled in intrapersonal intelligence and Einstein in logical-mathematical intelligence. Although these individuals excelled in one particular intelligence, Gardner theorized that most individuals exhibit some balance across levels of the various intelligences (Gardner, 1995b, 2006. In 1996, psychologists Mihaly Csikszentmihalyi, William Damon, and Howard Gardner began focusing on qualitative research and developing practical materials focused on the meaning of good work, effective collaboration, digital citizenship, and civic participation. The GoodWork Project (Gardner, 2016b) continues Gardner's long-standing efforts to probe the nature of good work across different professional domains and promote good work, described as high quality, socially responsible, and meaningful. Gardner and his collaborators have focused on determining how best to increase the incidence of good work by introducing a GoodWork Toolkit, an instrument that consists of actual ethical dilemmas faced by professionals that can be used in schools as well as other settings.

Triarchic Theory, Balance Theory of Wisdom, and Successful Intelligence Applied to Cognitive Giftedness: Robert J. Sternberg

Robert J. Sternberg developed his own multidimensional conception of intelligence, the triarchic theory of intelligence (Sternberg, 1985). According to this theory, intelligence is the interplay between analytical, creative, and practical abilities in a given sociocultural environment. Analytical abilities are those most traditionally associated with intelligence and involve evaluating and analyzing information. Creative and practical abilities differ from traditional conceptions of intelligence as they are more associated with generating new ideas and applying knowledge in a given context. More recently, Sternberg's (2015) theory of successful intelligence explained how individuals can optimize their different strengths while compensating for their relative weaknesses. Successful intelligence shifts away from ability or aptitude measurement and relies on individualized assessments of achievement. According to this theory of successful intelligence, intelligence can be transformed into the development of expert performance in a given field and is measured by how a person develops their abilities by adapting, shaping, and selecting different environments.

Sternberg is one of the few cognitive psychologists who has conducted research on the ways his theories of intelligence apply to giftedness (Sternberg, 2005). Gifted individuals, according to Sternberg, demonstrate three common attributes that comprise his definition of intelligence (Sternberg, 1985, 1996), including analytical giftedness, demonstrated by an ability to analyze and evaluate one's own ideas and those of others; creative giftedness, an ability to generate one or more major ideas that are novel and of high quality; and practical giftedness, an ability to convince people of the value and practicality of ideas. According to Sternberg (1990), individuals possess patterns of strengths and weaknesses, although their patterns may change over time. Many tasks require all three kinds of thinking but that does not mean that people in general, or gifted people in particular, are equally adept at all three types. Rather, gifted individuals capitalize on their strengths and compensate for or correct their weaknesses. (Sternberg, 1996). People may show different patterns of skills in general, and of giftedness, in particular periods over the course of their lives. Sternberg, after three decades of work, still had unanswered questions about the nature of superior intelligence and why some gifted individuals had such a positive impact on the world, while others did not. Sternberg studied individuals such as Gandhi, Martin Luther King, Jr., and Mother Teresa and compared them to Stalin and Hitler, finding that they didn't differ much in intelligence but rather in wisdom.

Sternberg (1998) also proposed a balance theory of wisdom, defining wisdom as the use of one's intelligence, creativity, common sense, and knowledge, mediated by positive ethical values toward the achievement of the common good. His balance theory of wisdom is achieved through the recognition and development of intrapersonal, interpersonal, and extrapersonal interests over time to achieve a balance among one's adaptation to existing environments, shaping of existing environments, and selection of new environments. Sternberg believes that wise decisions require more than intelligence and explicit knowledge, including tacit, or implicit, knowledge gained through experiences. Wisdom, according to Sternberg's theory (Sternberg, 1990), draws on a personal understanding of balance, the balance among multiple interests, immediate and lasting consequences, and environmental responses. Balance needs to exist for intrapersonal interests, extrapersonal interests, and environmental responses and these interests are not weighed equally. Relative weights exist that are determined by the extent to which a particular alternative contributes to the achievement of a common goal. Wisdom involved identifying the common goal and persuading others of its value.

Sternberg's theory of successful intelligence (1996) may be the most interesting and exciting extension of his voluminous work, as he links intelligence to intellectual giftedness, defining successful intelligence as one's ability to set and accomplish personally meaningful goals in one's life, given one's cultural context. A successfully intelligent person accomplishes goals by identifying their strengths and weaknesses, and then by capitalizing on the strengths and correcting or compensating for the weaknesses. According to Sternberg (1996), strengths and weaknesses emerge across various creative, analytical, practical, and wisdom-based abilities. In particular, individuals need to be creative in order to generate novel and useful ideas; analytical in order to ascertain that the ideas they have (and that others have) are good ones; practical in order to apply those ideas and convince others of their value; and use wisdom in order to ensure that implementation of the ideas will help ensure a common good through the mediation of positive ethical principles. And most important to educators, Sternberg believes that wisdom can and should be taught (2001).

Interesting Directions in Intellectual Giftedness and Talent Development

Talent Development in Some or All Young People

Research on the development of intellectual giftedness has demonstrated that talents develop across multiple domains and over time, with the right combination of innate talent, parental support, expert teaching, and the desire of the individual to apply the effort necessary to develop innate talents (Bloom, 1985; Csikszentmihalyi et al., 1993; Renzulli, 1978, 2012). Some research, such as Bloom's classic work, examined the childhoods and backgrounds of highly accomplished individuals across domains to identify common features contributing to their talent development. High levels of talent development require constant attention, nurturing, and focused effort and task commitment. Whether or not a talent ultimately develops seems to depend on many factors, including abilities, creativity, effort, motivation to achieve, societal support and appreciation of the talent area, environmental support and opportunities, and chance or luck (Bloom, 1985; Csikszentmihalyi et al., 1993). Research also suggests that supportive experiences at school, in the community, and at home are critical forces in transforming potential into fully developed talents (Bloom, 1985; Csikszentmihalyi et al., 1993). For example, Csikszentmihalyi and his colleagues (1993) studied intellectually talented teens, identifying a variety of factors that contribute to the development of their talents, including enjoyment of classes and activities, having adults help them establish both short- and long-term goals, and encouraging student engagement and commitment to their talent areas during critical periods of development, such as adolescence. Talent development research conducted by Bloom (1985) and Csikszentmihalyi and colleagues (1993) demonstrates that outstanding talent is developed by individuals over long periods of time and is influenced by a variety of factors, such as the personal characteristics of the talented person and an individual's support systems.

Bloom and colleagues (1985) studied musicians, athletes, and scholars who achieved high-level public recognition, focusing on the significant factors in the development of talent and the contributions of home and school. A positive family environment as well as support and encouragement from parents or family members with a personal interest in the talent field were found to be essential in the development of exceptional accomplishment in a talent area. Bloom also found that talented individuals across domains demonstrate certain qualities such as a strong interest and emotional commitment to a particular talent field, a desire to reach a high level of attainment in the talent field, and a willingness to put in the great amounts of time and the effort needed to reach very high levels of achievement in the talent field. Bloom found that the psychological factors involved in the development of outstanding talent often occur over a long period of time and are influenced by a variety of individuals and factors, including the personal characteristics of the talented person and a strong support system. Parents instill the value of working hard during the early years. In the second phase (the precision phase), a master coach or teacher helps the talented individual to master the long-term systematic skills necessary to hone the talent, focusing on technical mastery, technique, and excellence in skill development. Finally, in the third phase (the elite years), the individual continues to work with a master teacher and practice many hours each day to turn training and technical skills into personalized performance excellence. During the final phase, a realization occurs of how significant the activity has become in one's life.

Csikszentmihalyi and colleagues (1993), in a classic five-year longitudinal study, examined the experiences of 200 talented teenagers in athletics, art, music, and science to identify similarities and differences between teens who developed and used their talents in adulthood, as opposed to those who drifted away from their talents to pursue work that required only average skills. The researchers described the need for talented teenagers to acquire a set of "metaskills" that allowed them to work with intense concentration and curiosity in order to develop their talents. Talent, these researchers learned, was developmental and affected by contextual factors in the environment. Talent was nurtured by the acquisition of knowledge of the domain, motivation provided by the family and persons in the specialized field of talent, and discipline created by a set of habits resulting in long-term concentrated study and superior performance. The talented teenagers studied shared certain personal characteristics, including the ability to concentrate, which led to both achievement and endurance, and an awareness of experience that enhanced their understandings. They experienced flow, a "state in which people are so involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at great cost, for the sheer sake of doing it" (p. 4). When immersed in pleasurable work, these teenagers pursued work as a reward in itself. Csikszentmihalyi and his colleagues also found that teens with little family support spent large amounts of time with peers instead of working on their talents and subsequently failed to develop their abilities, suggesting the need for careful parental monitoring of talent development. They also found that children must first be recognized as talented to develop that talent, and therefore must have skills considered useful in their cultures. These researchers also learned that talents can be developed if the process produces optimal, enjoyable experiences, and if the memories of peak moments will continue to motivate students. Csikszentmihalyi's subsequent work (2014) continued to investigate flow, and particularly, his finding that people were most creative, productive, and happiest when they are in a state of flow. His interviews with athletes, musicians, artists and other creative individuals enabled him to better understand when they experienced the most optimal performance levels, how they felt during these experiences, and what piqued their creativity, especially in the workplace, and how creativity lead to more productivity.

Subotnik and colleagues (2011, 2012) recently developed a new definition of giftedness as well as a comprehensive organizational scheme that integrates giftedness, talent, expertise, and eminence, conceptualizing them as a sequence of stages, that under the best of circumstances, leads to eminence. They acknowledge that the abilities of individuals do matter, particularly their abilities in specific talent domains, which have different developmental trajectories that vary. The variation in domains include aspects such as when they start, peak, and end, and in the opportunities provided by society that are crucial at every point in the talent-development process. The authors believe that society must strive to promote these opportunities, but that individuals with talent also must assume responsibility for their own growth and development. Subotnik and colleagues (2011, 2012) also believe that preparing young people for outstanding achievement or eminence should be the chief goal of gifted education, a point of view that Borland (2012) criticized as perhaps leading to unfavorable educational outcomes, such as fewer educational programs for talented youth, for as Borland indicates, many talent-development activities lie outside of the mission of schools.

Renzulli and Reis (2014) have consistently advocated that schools should be places for talent development and that educators will not know which students have the best chance to develop their talents until they have had the opportunities, resources, and encouragement to do so. They believe that some enrichment opportunities should be developed for all students to know which students should continue to receive more advanced opportunities to develop their talents. They also believe that talents are most likely to emerge in students with above-average intellectual abilities, in whom creativity and task commitment are more likely to be developed. Across all of these scholars, it seems that a clear trend exists in perceptions that giftedness is and can be developed, that there is a need for talent development both in schools and in home environments, and that one important aim for high aptitude or potential in students is the development of outstanding achievement, creativity, or eminence in life, resulting in more creative and intellectual work in all fields.

Students with Learning Disabilities Who also Have High Potential: 2E

Across the past few decades, a more comprehensive understanding has emerged of twice-exceptional (2E) students (Reis, Baum, & Burke, 2014), those who are highly intelligent but who also have significant learning difficulties and special education needs. These students have been found to have unique learning patterns, requiring different strategies both to identify them as twice exceptional and therefore eligible for various types of programming, talent development opportunities, and instructional strategies to develop their potential. As many as 70,000 students in elementary and secondary schools in the United States are currently identified as 2E among school districts that voluntarily track and report this data, consistent with estimates that 2 to 5 percent of the gifted population have some type of learning disability and that up to 5 percent of students with learning disabilities also have intellectual gifts and talents (Baum, Schader, & Hébert, 2014). This number will most likely increase as more school personnel become aware of the needs of twice-exceptional students, and as more districts collect these data. The reported numbers of 2E students who are highly intelligent but who also have autism, for example, has continued to increase over the past decade (Gelbar, 2017).

Researchers who have studied 2E students encourage teachers to create nurturing environments in which these exceptional students have opportunities for talent development (Baum, 2008; Baum et al., 2014; Renzulli & Reis, 2014) and the time to understand their interests and learning styles and the ways they learn best (Reis et al., 2014). Teachers should enable students to compensate for their student deficits and be given accommodations using technology and students' interests, enabling students to self-select high-interest topics within the curriculum for more in-depth study. Both Baum (2008) and Renzulli and Reis (2014) suggest strength-based specific teaching strategies that provide the structure necessary for student success but also include open-ended challenges requiring divergent thinking. They also suggest enabling students to work in their preferred learning styles, interest, and strengths areas, by providing advanced opportunities for students to investigate real-world problems for real audiences (Renzulli & Reis, 2014, Type III projects). Emotional and social support for these students may also be necessary as lowered self-concept and self-esteem, as well as fear of failure, negative interactions with teachers, and poor peer relations have been found to occur in these students (Reis et al., 1997). The use of compensation strategies to enable 2E students to use their strengths to compensate for weaknesses is also important. These strategies include breaking down difficult tasks into more manageable chunks, teaching the use of self-regulation techniques, establishing goals, learning time management skills, and mastering the use of all types of technology aides, such as speech-recognition software.

Fixed versus Malleable Traits: Carol Dweck

Other work, which has been widely discussed and has become influential in many schools, especially with educators who have embraced a mindset view of intelligence and giftedness, has recently been questioned by other psychologists. Carol Dweck's (2006) theory of an entity view of intelligence as opposed to an incremental (malleable) view of intelligence has been considered a major contributor to our understanding of why some high-potential students are more willing than others to expend effort to succeed. If a student believes that intelligence is a fixed trait (e.g., I can't do this because I am not smart enough), they may fail or even refuse to try to complete a challenging task simply because they believe they do not have the capacity to succeed. If the same person believes that their abilities can improve, that is, that they are malleable, they will have more of a chance at being successful. In other words, a belief that one's performance can improve is a key to success on cognitive tasks. Dweck's research about how beliefs influence cognitive ability and whether or not a student's view of intelligence is a fixed or malleable ability is an interesting addition to current research on intellectual giftedness and is ubiquitous in education. Li and Bates (2017), however, attempted to replicate Dweck's findings over time without success, finding no support for the theory that fixed beliefs about basic ability are harmful, or that implicit theories of intelligence play any significant role in development of cognitive ability, response to challenge, or educational attainment.

Where Things Stand Today

In the past few decades, a consensus seems to have been reached that giftedness cannot be expressed in a unitary manner, suggesting a wider acceptance of more multifaceted approaches to intellectual giftedness and to the development of talent in children. More contemporary research has provided support for multiple components of intellectual giftedness, as summarized in two different volumes related to conceptions of giftedness by Sternberg and Davidson (1986, 2005) and a recent comprehensive handbook on giftedness and talent (Pfeiffer et al., 2018). The majority of current researchers define giftedness in terms of multiple qualities and extend definitions beyond unitary views of intellectual giftedness. Most also believe that IQ scores alone are inadequate measures of intellectual giftedness, and that motivation, high self-concept, and creativity are key qualities in many of these broader conceptions of giftedness (Sternberg & Davidson, 1986, 2005).

The realization that many students demonstrate traits of intellectual giftedness and still fail to achieve in school or life is also an increasing concern for parents, psychologists, and educators. Why, for example, do some extremely smart children fail to realize their promise and potential (Reis & McCoach, 2000; Renzulli & Park, 2002)? Why is it that some prodigies grow up to be average performers in the very fields in which they showed such promise when they were children (Feldman & Goldsmith, 1991; Winner, 1996)? Why do other traits, described by Renzulli (2002) as co-cognitive traits, appear to be so important in the process of talent development and intellectual giftedness? This chapter has summarized some pertinent research about intellectually gifted and talented individuals but much remains to be learned. Some researchers who have studied talent development have identified trends and findings that can help us as we consider the types of experiences needed to maximize any developmental considerations related to intellectual giftedness. However, a consensus has not and probably will not be reached about how to develop intellectual giftedness because of the very diversity of how we define giftedness. This lack of consensus may be completely appropriate, as the complexities surrounding this construct continue to both intrigue and challenge researchers.

Current Federal Definition

In 1993, a task force of psychologists, educational psychologists, educational researchers, and teachers developed a new federal definition; healthy debate and discussion resulted. The federal definition, still widely used by many states and school districts, that emerged from this committee follows:

> Children and youth with outstanding talent perform or show the potential for performing at remarkably high levels of accomplishment when compared with others of their age, experience, or environment. These children and youth exhibit high performance capability in intellectual, creative, and/or artistic areas, possess an unusual leadership capacity, or excel in specific academic fields. They require services or activities not ordinarily provided by the schools. Outstanding talents are present in children and youth from all cultural groups, across all economic strata, and in all areas of human endeavor. (US Department of Education, Office of Educational Research and Improvement, 1993, p. 26)

In 2002, under the No Child Left Behind legislation, gifted students were defined as, “Students, children, or youth who give evidence of high achievement capability in areas such as intellectual, creative, artistic, or leadership capacity, or in specific academic fields, and who need services and activities not ordinarily provided by the school in order to fully develop those capabilities."[1]

Characteristics of Individuals with High Intellectual Ability or Potential

Some consensus also exists about the characteristics of high-potential students. In an extensive review of research about identified gifted and high-potential students from diverse backgrounds, Frasier and Passow (1994) identified "general/common attributes of giftedness" – traits, aptitudes, and behaviors consistently identified by researchers as common to all gifted students. They found that the following basic elements of giftedness are similar across cultures (though each is not displayed by every student): motivation, advanced interests, communication skills, problem-solving ability, well-developed memory, inquiry, insight, reasoning, imagination/creativity, sense of humor, and an advanced ability to deal with symbol systems. Each of these common characteristics may be manifested in different ways in different students and we should be especially careful in attempting to identify these characteristics in students from diverse backgrounds since behavioral manifestations of the characteristics may vary with context. By this we mean that motivation may be manifested differently by a Hispanic urban student who speaks English as a second language than by a student who lives in an upper-socioeconomic neighborhood and is from a majority culture.

Lack of Progress for Culturally, Linguistically, and/or Economically Diverse Gifted Students

One fact remains clear in the research literature on gifted education. Students of some racial and ethnic backgrounds (e.g., African American, Hispanic or Latino, and Native American), as well as students from lower-income and high-poverty backgrounds, are disproportionally underrepresented in gifted and talented programs (Davis, 2010; Ford, Grantham, & Whiting, 2008; Hamilton et al., 2018). These students are less likely to be identified as gifted and talented in early elementary school, and less likely to participate in programs for gifted and talented students. This lack of identification and subsequent participation is due to a number of factors, including inappropriate identification methods used in districts and states, bias in instrumentation, and problems with referrals for gifted identification, as students from diverse and high-poverty populations are generally much less likely to be recommended by their teachers. The effects of poverty and undiagnosed learning disabilities may also negatively affect assessment, as well as have a negative impact on student performance. Challenges in areas such as nutrition, health, and parents' time and availability to be involved in their children's education and provide

[1] No Child Left Behind Act of 2001, Pub. L. 107–110, 115 Stat. 1595, codified as amended at 20 U.S.C. §1425–2094.

enrichment are all factors that influence the identification and development of giftedness and talent potential. Without an understanding of the needs and characteristics of high-poverty, culturally and linguistically diverse learners, their talents will continue to fail to be identified.

Interventions and Programs for Gifted and High-Potential Students

The need for and types of interventions required by high-potential and gifted and talented students suggest several important findings. Research has consistently demonstrated that the needs of these students are generally not met in American classrooms, as the focus is often on the deficits of struggling learners and most classroom teachers have not had the training necessary to meet the needs of gifted students (Archambault et al., 1993; Reis et al., 2004; Renzulli & Reis, 2014; Westberg et al., 1993). Research also documents the benefits of cluster grouping gifted students together for differentiated and advanced instruction in order to increase achievement for gifted students, and in some cases, also for students who are achieving at average and below-average levels (Gentry & Owen, 1999). A strong, decades-old research base also demonstrates that the use of acceleration results in higher achievement for gifted and talented learners (Colangelo, Assouline, & Gross, 2004). Acceleration of various types is described in *A Nation Deceived* (Colangelo et al., 2004), a widely cited publication that illuminated all types of false misinformation about acceleration that has persisted over decades. Acceleration includes grade skipping, accelerated content such as giving fifth-grade reading to an advanced third-grade reader, and is usually warranted when students are very high academic achievers who require advanced content to keep them engaged and challenged. It also includes summer programs taken by talented youth in top-notch universities such as Johns Hopkins, Duke, Iowa, Northwestern, and Vanderbilt, where thousands of intellectually advanced adolescents annually qualify to participate in fast-paced (accelerated) educational opportunities and many receive credit for a full high school course after three weeks. Approximately 200,000 seventh- and eighth-graders take college entrance exams to learn about their abilities and to qualify for these types of educational programs annually (Olszewski-Kubilius, 2015) and many who participate receive credit for a full high school course (Assouline, Colangelo, & VanTassel-Baska, 2015a). Recently updated, the newest *A Nation Empowered* summarizes progress about more recent acceleration practices (Assouline et al., 2015a; Assouline et al., 2015b).

Enrichment, including interest-based projects, opportunities for independent study, and work on topics of interest that extend beyond the regular curriculum, should also be considered for gifted and high-ability students, and for other students with advanced interests or creativity (Renzulli & Reis, 2014). Based on current research discussed in this chapter, it appears that whenever possible, a combination of enrichment and acceleration is needed to engage and challenge gifted and high-potential students. Research on the use of enrichment and curriculum enhancement resulted in higher achievement for gifted and talented learners as well as other students (Gavin et al., 2007; Gentry & Owen, 1999; Gubbins et al., 2007; Reis et al., 2007; Tieso, 2002). Gifted programs and strategies are effective at serving gifted and high-ability students in a variety of educational settings (Colangelo et al., 2004; Gavin et al., 2007; Reis et al., 2007). There is also a trend to identify and provide enrichment to more 2E students, those with high ability and learning disabilities (Baum et al., 2014), as well as more students who attend schools that serve diverse ethnic and socioeconomic populations (Renzulli & Reis, 2014). Enrichment has also been used to reverse underachievement (Baum, Renzulli, & Hébert, 1995). Gifted education programs and enrichment strategies have also been found to benefit gifted and talented students longitudinally, helping students increase aspirations for college and careers, determine postsecondary and career plans, develop creativity and motivation that is applied to later work, and achieve more advanced degrees (Colangelo et al., 2004; Hébert, 1993; Lubinski et al., 2001; Westberg, 2012).

To challenge academically talented and high-potential learners, a need exists for educators to develop a continuum of services in each school, as suggested by the most popular enrichment approach in gifted education, the Schoolwide Enrichment Model (SEM; Renzulli & Reis, 1997, 2003, 2014). This continuum of services should incorporate both enrichment and acceleration to challenge the diverse learning and affective needs of gifted and talented students. Services should be provided for gifted and high-potential students across all grade levels, and a broad range of services should be defined to ensure that children have access to areas such as curriculum and instructional differentiation. A broad range of enrichment and acceleration opportunities should be offered to meet the needs of rapid, advanced learners; opportunities for advanced content should be delivered so that students can continue to make progress in all content areas; and opportunities should be made available for individualized research for students who are highly creative and want the chance to pursue appropriate interests. For students who are underachieving or who have gifts and talents but also learning disabilities, counseling and other services are recommended to address these special affective needs.

Conclusions

Our nation and the schools within it must be cautious not to squander the intellectual potential of students and ensure that we do not contribute to the underachievement of our most academically able students across their life spans. As many as half of our urban high-poverty gifted and talented students underachieve by the time they reach high school (Reis et al., 1995), and although psychologists differ on exactly how we should define giftedness, a consensus exists that we must expend greater efforts to develop gifts and talents in a broad population of youth by understanding how personal variables, family influences, and school and other environmental factors can be enhanced to enable students to be able to make the effort to develop their talents in the directions and areas they choose.

Decades ago, Borland (1989) made the distinction between two views of gifted students and gifted education, portraying two concomitant rationales for the existence of the field. One was characterized as the *national-resource approach*, in which gifted students are thought of as a vast, largely untapped resource that needs to be identified and developed for the common good. He labeled this a future-oriented approach, as educators work with gifted students who have the greatest potential to become creative-productive adults (Renzulli, 1978) or develop eminence (Subotnik et al., 2011). He labeled the second approach for the goal of gifted education the *special-education approach*, not with a the goal to produce eminent adults but rather to make education appropriate for high-ability students. In this approach, gifted students are considered exceptional learners who require special educational provisions if they are to receive the education to which they are entitled. We believe that a distinction does not have to be drawn between these two views, and that, rather, we can and should provide an education that all gifted and high potential students need and deserve, and also, simultaneously we should expose these young people to experiences and pedagogy, that with their own development of focus and effort, enables these students to emerge as the next generation of creative adults who can achieve eminence.

References

Archambault, F. X., Jr., Westberg, K. L., Brown, S., Hallmark, B. W., Emmons, C., & Zhang, W. (1993). Regular classroom practices with gifted students: Results of a national survey of classroom teachers. Research Monograph 93102. National Research Center on the Gifted and Talented, University of Connecticut, Storrs, CT.

Assouline, S. G., Colangelo, N., & VanTassel-Baska, J. (2015a). *A nation empowered: Evidence trumps the excuses that hold back America's brightest students*, vol. 1. Iowa City: Belin-Blank Center for Gifted Education and Talent Development, University of Iowa.

Assouline, S. G., Colangelo, N., VanTassel-Baska, J., & Lupkowski-Shoplik, A. (2015). *A nation empowered: Evidence trumps the excuses that hold back America's brightest students*, vol. 2. Iowa City: Belin-Blank Center for Gifted Education and Talent Development, University of Iowa.

Baum, S. M. (2008). Talent centered model for twice exceptional students. In J. S. Renzulli, E. J. Gubbins, K. S. McMillen, R. D. Eckert, & C. A. Little (Eds.), *Systems and models in gifted education* (pp. 17–48). Mansfield Center, CT: Creative Learning Press.

Baum, S. M., Renzulli, J. S., & Hébert, T. P. (1995). Reversing underachievement: Creative productivity as a systematic intervention. *Gifted Child Quarterly*, *39*, 224–235.

Baum, S., Schader, R. M., & Hébert, T. P. (2014). Through a different lens: Reflecting on a strengths-based, talent-focused approach for twice-exceptional learners. *Gifted Child Quarterly*, *58*, 311–327.

Bloom, B. S. (Ed.) (1985). *Developing talent in young people*. New York: Ballantine Books.

Borland, J. H. (1989). *Planning and implementing programs for the gifted*. New York: Teachers College Press.

Borland, J. H. (2012). A landmark monograph in gifted education, and why I disagree with its major conclusion. The Creativity Post. www.creativitypost.com/education/a_landmark_monograph_in_gifted_education_and_why_i_disagree_with_its_major

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Ceci, S. J., & Papierno, P. B. (2005). The rhetoric and reality of gap closing: When the "have nots" gain but the "haves" gain even more. *American Psychologist*, *60*, 149–160.

Colangelo, N., Assouline, S., & Gross, M. (Eds.) (2004). *A nation deceived: How schools hold back America's brightest students*. Iowa City: University of Iowa.

Csikszentmihalyi, M. (2014). *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi*. New York: Springer.

Csikszentmihalyi, M., Rathunde, K., & Whalen, S. (1993). *Talented teenagers: A longitudinal study of their development*. New York: Cambridge University Press.

Dai, D. Y. (2010). *The nature and nurture of giftedness: A new framework for understanding gifted education*. New York: Teachers College Press.

Dai, D. Y., & Renzulli, J. S. (2008). Snowflakes, living systems, and the mystery of giftedness. *Gifted Child Quarterly*, *52*, 114–130.

Davis, J. L. (2010). *Bright, talented, and Black: A guide for families of African-American gifted learners*. Scottsdale, AZ: Great Potential Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Feldman, D. H., & Goldsmith, L. T. (1991). *Nature's gambit: Child prodigies and the development of human potential*. New York: Teachers College Press.

Ford, D. Y., Grantham, T. C., & Whiting, G. W. (2008). Culturally and linguistically diverse students in gifted education: Recruitment and retention issues. *Exceptional Children*, *74*, 289–306.

Frasier, M. M., & Passow, A. H. (1994). Toward a new paradigm for identifying talent potential. Research Monograph 94112. National Research Center on the Gifted and Talented, University of Connecticut, Storrs, CT.

Gagne, F. (2000). Understanding the complex choreography of talent development. In K. A. Heller, F. J. Monks, R. J. Sternberg, & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (pp. 67–79). Amsterdam: Elsevier.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Needham Heights, MA: Allyn & Bacon.

Gardner, H. (1995a). Reflections on multiple intelligences: Myths and messages. *Phi Delta Kappan*, *77*(3), 200–209.

Gardner, H., with the collaboration of Laskin, E. (1995b). *Leading minds: An anatomy of leadership*. New York: Basic Books.

Gardner, H. (2006). *Multiple intelligences: New horizons*. New York: Basic Books.

Gardner, H. (2016a). Multiple intelligences: Prelude, theory, and aftermath. In R. J. Sternberg, S. T. Fiske, & D. J. Foss (Eds.), *Scientists making a difference: One hundred eminent behavioral and brain scientists talk about their most important contributions* (pp. 167–170). New York: Cambridge University Press.

Gardner, H. (2016b). Three frames for good work. In M. Levinson & J. Fay (Eds.), *Dilemmas of educational ethics: Cases and commentaries* (pp. 97–100). Cambridge, MA: Harvard Education Press.

Gavin, M. K., Casa, T. M., Adelson, J. L., Carroll, S. R., Sheffield, L. J., & Spinelli, A. M. (2007). Project M3: Mentoring mathematical minds: Challenging curriculum for talented elementary students. *Journal of Advanced Academics*, *18*, 566–585.

Gelbar, N. W. (Ed.) (2017). *Adolescents with autism spectrum disorder: A clinical handbook*. New York: Oxford University Press.

Gentry, M. L., & Owen, S. V. (1999). An investigation of the effects of total school flexible cluster grouping on identification, achievement, and classroom practices. *Gifted Child Quarterly*, *43*, 224–243.

Gubbins, E. J., Housand, B., Oliver, M., Schader, R., De Wet, C., Moon, T. et al. (2007). Unclogging the mathematics pipeline through access to algebraic understanding. Research Monograph 08236. National Research Center on the Gifted and Talented, University of Connecticut, Storrs, CT.

Gustafsson, J., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186–242). New York: Macmillan.

Hamilton, R., McCoach, D. B., Tutwiler, M. S., Siegle, D., Gubbins, E. J., Callahan, C. M., Brodersen, A. V., et al. (2018). Disentangling the roles of institutional and individual poverty in the identification of gifted students. *Gifted Child Quarterly*, *62*, 6–24.

Hébert, T. P. (1993). Reflections at graduation: The long-term impact of elementary school experiences in creative productivity. *Roeper Review*, *16*, 22–28.

Housand, A., & Reis, S. M. (2009). Self-regulated learning in reading: Gifted pedagogy and instructional settings. *Journal of Advanced Academics*, *20*, 108–136.

Li, Y., & Bates, T. C. (2017). Does growth mindset improve children's IQ, educational attainment or response to setbacks? Active-control interventions and data on children's own mindsets. osf.io/preprints/socarxiv/tsdwy

Lohman, D. F., Gambrell, J., & Lakin, J. (2008). The commonality of extreme discrepancies in the ability profiles of academically gifted students. *Psychology Science Quarterly*, *50*, 269–282.

Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science*, *25*, 2217–2232. https://doi.org/10.1177/0956797614551371

Lubinski, D., Webb, R. M., Morelock, M. J., & Benbow, C. P. (2001). Top 1 in 10,000: A 10 year follow-up of the profoundly gifted. *Journal of Applied Psychology*, *4*, 718–729.

Olszewski-Kubilius, P. (2015). Talent searches and accelerated programming for gifted students. In S. G. Assouline, N. Colangelo, J. Vantassel-Baska, & A. Lupkowski-Shoplik (Eds.), *A nation empowered* (vol. 2, pp. 111–121). Iowa City, IA: Belin-Blank Center.

Pfeiffer, S. I., Shaunessy-Dedrick, E., & Foley-Nicpon, M. (Eds.) (2018). *APA handbook of giftedness and talent* (APA handbooks in psychology). Washington: American Psychological Association.

Phillipson, S. N., & McCann, M. (2007). *Conceptions of giftedness: Sociocultural perspectives*. Mahwah, NJ: Lawrence Erlbaum.

Reis, S. M. (1998). *Work left undone*. Mansfield Center, CT: Creative Learning Press.

Reis, S. M. (2002). Toward a theory of creativity in diverse creative women. *Creativity Research Journal*, *14*, 305–316.

Reis, S. M. (2005). Feminist perspectives on talent development: A research based conception of giftedness in women. In R. J. Sternberg & J. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 217–245). Boston: Cambridge University Press.

Reis, S. M., Baum, S. M., & Burke, E. (2014). An operational definition of twice-exceptional learners: Implications and applications. *Gifted Child Quarterly*, *58*, 317–230.

Reis, S. M., Gubbins, E. J., Briggs, C., Schreber, F. R., Richards, S., Jacobs, J., et al. (2004). Reading instruction for talented readers: Case studies documenting few opportunities for continuous progress. *Gifted Child Quarterly*, 48, 309–338.

Reis, S. M., Hébert, T. P., Díaz, E. I., Maxfield, L. R., & Ratley, M. E. (1995). Case studies of talented students who achieve and underachieve in an urban high school. Research Monograph No. 95120. National Research Center on the Gifted and Talented, University of Connecticut, Storrs, CT.

Reis, S. M., & McCoach, D. B. (2000). The underachievement of gifted students: What do we know and where do we go? *Gifted Child Quarterly*, *44*, 152–170.

Reis, S. M., McCoach, D. B., Coyne, M., Schreiber, F. J., Eckert, R. D., & Gubbins, E. J. (2007). Using planned enrichment strategies with direct instruction to improve reading fluency, comprehension, and attitude toward reading: An evidence-based study. *Elementary School Journal*, *108*, 3–24.

Reis, S. M., Neu, T. W., & McGuire, J. M. (1997). Case studies of high ability students with learning disabilities who have achieved. *Exceptional Children*, *63*, 463–479.

Renzulli, J. S. (1978). What makes giftedness: Reexamining a definition. *Phi Delta Kappan*, *60*, 180–184.

Renzulli, J. S. (1986). The three ring conception of giftedness: A developmental model for creative productivity. In R. J. Sternberg & J. Davidson (Eds.), *Conceptions of giftedness* (246–279). New York: Cambridge University Press.

Renzulli, J. S. (2002). Expanding the conception of giftedness to include co-cognitive traits and to promote social capital. *Phi Delta Kappan*, *84*(1), 33–40, 57–58.

Renzulli, J. S. (2005). The three-ring conception of giftedness: A developmental model for promoting creative productivity. In R. J. Sternberg & J. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 217–245). Boston: Cambridge University Press.

Renzulli, J. S. (2012). Reexamining the role of gifted education and talent development for the 21st century: A four-part theoretical approach. *Gifted Child Quarterly*, *56*, 150–159.

Renzulli, J. S., & Park, S. (2002). Giftedness and high school dropouts: Personal, family, and school-related factors. Research Monograph No. 02168. National Research Center on the Gifted and Talented, University of Connecticut, Storrs, CT.

Renzulli, J. S., & Reis, S. M. (1997). *The school-wide enrichment model: A comprehensive plan for educational excellence*. Mansfield Center, CT: Creative Learning Press.

Renzulli, J. S., & Reis, S. M. (2003). Conception of giftedness and its relation to the development of social capital. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (3rd ed., pp. 75–87). Boston: Allyn & Bacon.

Renzulli, J. S., & Reis, S. M. (2014). *The Schoolwide Enrichment Model: A how-to guide for talent development* (3rd ed.). Waco, TX: Prufrock Press.

Shavinia, L. V. (2001). Beyond IQ: A new perspective on the psychological assessment of intellectual abilities. *New Ideas in Psychology*, *19*(1), 27–47.

Simonton, D. K. (1998). Creativity, genius, and talent development. *Roeper Review*, *21*(1), 86–87.

Simonton, D. K. (2005). Genetics of giftedness: The implications of an emergenic-epigenetic model of giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 312–326). Boston: Cambridge University Press.

Simonton, D. K. (2008). Scientific talent, training, and performance: Intellect, personality, and genetic endowment. *Review of General Psychology*, *12*, 28–46

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (1990). Wisdom and its relations to intelligence and creativity. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 142–159). New York: Cambridge University Press.

Sternberg, R. J. (1996). *Successful intelligence: How practical and creative intelligence determine success in life*. New York: Simon & Schuster.

Sternberg, R. J. (1998). A balance theory of wisdom. *Review of General Psychology*, *2*, 347–365.

Sternberg, R. J. (2001). Why schools should teach for wisdom: The balance theory of wisdom in educational settings. *Educational Psychologist*, *36*, 227–245.

Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist*, *59*, 325–338.

Sternberg, R. J. (2005). The WISC model of giftedness. In R. J. Sternberg & J. Davidson (Eds.), *Conceptions of giftedness* (2nd ed., pp. 327–342). Boston: Cambridge University Press.

Sternberg, R. J. (2015). Successful intelligence: A new model for testing intelligence beyond IQ tests. *European Journal of Education and Psychology*, *8*, 76–84.

Sternberg, R. J., & Davidson, J. (Eds.) (1986). *Conceptions of giftedness*. New York: Cambridge University Press.

Sternberg, R. J., & Davidson, J. (Eds.) (2005). *Conceptions of giftedness* (2nd ed.). Boston: Cambridge University Press.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York: Free Press.

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest*, *12*, 3–54.

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2012). A proposed direction forward for gifted education based on psychological science. *Gifted Child Quarterly*, *56*, 176–188.

Tannenbaum, A. J. (1991). The social psychology of giftedness. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 27–44). Boston: Allyn & Bacon.

Terman, L. M. (1925–1959). *Genetic studies of genius* (5 vols.). Stanford: Stanford University Press.

Tieso, C. L. (2002). The effects of grouping and curricular practices on intermediate students' math achievement. Research Monograph 02154. National Research Center on the Gifted and Talented, University of Connecticut, Storrs, CT.

United States Department of Education, Office of Educational Research and Improvement. (1993). *National excellence: A case for developing America's talent*. Washington: US Government Printing Office.

Westberg, K. L. (2012). Young creative producers: twenty years later. *Gifted Education International*, *26*, 261–270.

Westberg, K. L., Archambault, F. X., Jr., Dobyns, S. M., & Salvin, T. J. (1993). An observational study of instructional and curricular practices used with gifted and talented students in regular classrooms. Research Monograph 93104. National Research Center on the Gifted and Talented, University of Connecticut, Storrs, CT.

Winner, E. (1996). *Gifted children: Myths and realities*. New York: Basic Books.

14 Sex Differences in Intelligence

Diane F. Halpern and Jonathan Wai

Questions about whether, why, and how much females and males differ in intelligence have engendered heated debates in contemporary psychology. The way researchers answer these questions has implications for public policy decisions as well as the way people think about education, career choices, and "natural" roles for males and females. For example, more than two decades ago, research was released proclaiming that girls were being "shortchanged" in schools (e.g., American Association of University Women, 1992; Sadker & Sadker, 1995). This conclusion was soon met with counterclaims that schools were biased against boys (Sommers, 2000). Claims about biases for and against girls and boys in school were interpreted in the context of international comparisons that document the overall low achievement of both boys and girls in the United States, relative to students in other countries, especially in science and math (National Science Foundation, 2014) and low high school graduation rates for both sexes, but especially for low-income males (US Department of Education, 2017). These proclamations about biases in education soon took on a political tone about the causes, correlates, and cures for sex differences in intelligence.

More recently, an entire issue of the prestigious *Journal of Neuroscience Research* was devoted to research on sex differences in the brain. It was titled, "An issue whose time has come" (Cahill, 2017). In the same month, a popular book titled *Testosterone Rex* was published and received a prestigious book award from the British Royal Society. Its main thesis is that research on sex differences in the brain is flawed (Fine, 2017). These sorts of controversies continue unabated with no signs of weakening or of either side calling for a truce.

Although most education pundits agree that education in the United States is in need of serious reform, some politicians and educators used the available data to argue that girls and boys learn differently and thus need single-sex (SS) schooling that would cater to these differences. The No Child Left Behind Act of 2001

Some prefer to use the term "sex" to refer to biological variables and the term "gender" to refer to psychosocial variables. We believe that a biopsychosocial model that includes the effect of these two categories of variables on each other is more accurate. Brain differences and hormone levels are changed by life experiences, and life experiences, in turn, are influenced by brain structures and functions and hormones. Thus, we use both terms somewhat interchangeably in this chapter. We use "sex differences" more often because we are most often discussing differences between two groups of people who are categorized by their genitals, but this is not meant to imply that biologically related variables are more important than psychosocial ones. We use "gender" less frequently, usually where it is the typical term in the literature being discussed.

authorized school districts to use funding to offer single-sex schools and classrooms at public expense, as long as this arrangement was consistent with applicable laws. An October 2006 amendment to Title IX, which mandates that educational institutions not discriminate on the basis of sex, was reinterpreted to allow single-sex schooling at public expense. Advocates for single-sex schooling maintain the position that single-sex education has benefits even though an extensive review conducted by the US Department of Education found that the majority of studies comparing single-sex with coeducational schooling report either no difference or mixed results (US Department of Education, 2005). Additionally, a more recent synthesis published in *Science* confirms the lack of evidence and suggests single-sex schooling might even exaggerate sexism and gender stereotyping (Halpern et al., 2011). Other reviews report a host of negative consequences associated with single-sex education, including increased sex-role stereotyping, which harms both boys and girls (Karpiak et al., 2007). A meta-analysis based on 184 studies with 1.6 million students compared single-sex education with coeducation and found "results from the highest quality studies, then, do not support the view that SS schooling provides benefits compared with CE schooling" (Pahlke, Hyde, & Allison, 2014, p. 1042). Other studies show that any benefits for single-sex education can be explained by selection effects – students who are academically more advanced are often selected for single-sex educational experiences (Hayes, Pahlke, & Bigler, 2011). But advocates of SS schooling seem unmoved by the extensive body of research that shows no benefits and suggests possible harm such as increased sex-role stereotyping caused by SS classrooms. Rather, they prefer to rely on personal opinions and anecdotes instead of analyses based on over one million participants. Challenges to the reinterpretation of Title IX to allow single-sex classes (in public education) are moving from the laboratory to the courthouse, where research findings are scrutinized by lawyers, judges, news reporters, and the general public, all of whom are asking these questions: What are the sex differences in intelligence? Are the brains of females and males so dissimilar that they justify the conclusion that males and females need separate educational experiences tailored to "the way they learn"? Should empirical research inform political decisions about how to educate boys and girls?

In this chapter, we explore the ways in which the sexes are similar and different in their cognitive abilities. Obviously, there are differences in the relative roles that men and women play in reproduction, but these have few, if any, implications for intellectual functioning. In this chapter, we present a balanced overview of the current findings in the research literature on sex differences in intelligence, describe the bases of controversies, and interpret these findings in a biopsychosocial context.

The Smarter Sex

Which is the smarter sex – males or females? This may seem like an easy question to answer because it would be a simple task to compare the average scores of large samples of females and males on intelligence tests. However, this obvious strategy will not work because tests of intelligence are carefully written so that there will be no average overall difference between the sexes (Brody, 1992; Loewen, Rosser, & Katzman, 1988). Questions that favor either sex are either eliminated from the test or matched with questions that favor the other sex to the same degree. Although some researchers report a small advantage for males on tests that were standardized to show no sex differences (Nyborg, 2005), most studies do not (Colom et al., 2000; Spinath, Spinath, & Plomin, 2008). In a review of this question, Dykiert, Gale, and Deary (2008) found that reported sex differences on intelligence tests can be explained by the use of samples that are not representative of females and males, in general, and thus reflect errors in the methods used to study this question. This conclusion was confirmed by Hunt and Madhyastha (2008), who provided a model of the subject-selection problem that occurred in studies that report sex differences in intelligence.

Researchers vary in the extent to which they stress either similarities or differences – sometimes called minimizers or maximizers. In comprehensive reviews of the sex-differences literature, Hyde (2005, 2014) concluded that males and females are more similar than different. By contrast, Irwing and Lynn (2005) focused their discourse on differences. The reality is far more nuanced, with some tests and measurements showing consistent findings that favor one sex over the other and many others that show little or no differences.

One set of findings that has been replicated many times is that females, on average, score higher on some tests of verbal abilities, especially those that require rapid access to and use of phonological and semantic information in long-term memory, production and comprehension of complex prose, perceptual speed, spelling, selective attention, and language comprehension (Hedges & Nowell, 1995; Paus et al., 2017; Torres et al., 2006). On average, girls outperform boys on reading assessments, with the size of the effect depending on the nature of the reading task. In one large-scale international assessment of 1.5 million children's reading scores, girls outperformed boys in all seventy-five countries included in the analysis (Stoet & Geary, 2013). The size of the sex difference in reading was moderately large (between 0.36 and 0.65 standard deviation units) in 55 percent of the cases, and there is some evidence that it may be increasing over time. The largest differences in reading were found at the lowest achievement levels. One of the largest sex differences in cognition is found with tests of writing (effect sizes over 0.5 standard deviation units) and has remained unchanged since the beginning of the twenty-first century (Reynolds et al., 2015).

Males, on the other hand, score higher on some tasks that require transformations in visual-spatial working memory, motor skills involved in aiming, spatiotemporal responding, and fluid reasoning, especially in abstract mathematical and scientific domains (Hedges & Nowell, 1995; Paus et al., 2017; Torres et al., 2006). Results with tasks that require generating an image and maintaining it in memory while "working" on it vary depending on the complexity of the image to be generated and the specific nature of the task, with observed differences favoring males that range between $d = 0.63$ and $d = 0.77$ (Loring-Meier & Halpern, 1999). Kaufman (2007) investigated whether sex differences in visual-spatial ability could result from differences in spatial working memory. He found sex differences favoring males on spatial working memory. These differences could explain a portion of the sex differences in mental rotation and other spatial tasks.

Jensen (1998) addressed the question of female-male differences in intelligence by analyzing tests that "load heavily on *g*" (which stands for general intelligence), but were not normed to eliminate sex differences. He concluded, "No evidence was found for sex differences in the mean level of *g* or in the variability of *g* . . . Males, on average, excel on some factors; females on others" (pp. 531–532). The distinction among cognitive tasks that favor either females or males has led to a recent model of intelligence that comprises three subcomponents – verbal, perceptual, and visual-spatial – with females showing an advantage for verbal and perceptual and males showing an advantage for visual-spatial (Johnson & Bouchard, 2006). Because much of the research literature has focused on sex differences in these and other components of intelligence, we frequently use the term "cognitive abilities" instead of the more global term "intelligence" when discussing sex differences.

Although sex differences in mathematics have received widespread attention as a possible reason for the underrepresentation of women in math-intensive careers (see Williams & Ceci, 2014, for a review), these differences depend on the portion of the distribution examined and the data that are used to support a particular conclusion. There are many more males with intellectual disabilities than females, reflecting an X-linked genetic locus for many categories of mental retardation. Estimates of the percentage of males and females with intellectual disabilities vary somewhat across studies, over time (probably because of changes in diagnostic criteria) and by type of disorder, but virtually every study reports a higher incidence of males than females (Braun et al., 2015). Autism spectrum disorder, for example, is diagnosed in 1 in 40 males and 1 in 182 females (Braun et al., 2015). Some tests of quantitative and visual-spatial abilities also show more males at the high end of the distribution and miss the greater number of males at the low end because the intellectually disabled are rarely included in tests that are administered in school settings. These data support the generally accepted conclusion that males are more variable in quantitative and visual-spatial abilities, with more males at both high- and low-ability ends of test scores. In a large-scale study of sex differences in variability, Johnson, Carothers, and Deary (2008) found that males are more variable, with greater variability at the low end of the distribution than at the high end, which reflects a greater incidence of intellectual disability among males. These authors concluded that sex differences at the high end of the distribution of intelligence scores cannot account for sex differences in high-level achievement.

Intelligence is a multidimensional construct, with different theorists focusing on different components. Salovey and Mayer (1990) have made a strong case for a type of intelligence involved in perceiving the emotions of others and regulating one's own emotions. Although we recognize the importance of emotional intelligence, we are limiting our discussion to cognitive aspects of intelligence because of space limitations.

Sex differences in variability in intelligence emerge in individuals as young as three years of age, even though girls obtain higher mean scores and it is girls who are overrepresented at the high-ability tail at ages two, three, and four (Arden & Plomin, 2006). By age ten boys are overrepresented at the high-ability tail, as would be expected given their greater variability. These data suggest that sex differences in variability emerge before preschool and are not shaped by educational experiences. Talent search data from Johns Hopkins University and Duke University can help us understand the fact that more boys achieve scores at the high end of the distribution on tests that presumably reflect mathematical ability. In the early 1980s, Benbow and Stanley (1983) observed sex differences in mathematical reasoning ability among tens of thousands of intellectually talented twelve- to fourteen-year-olds who had taken the SAT several years before the typical age achieved by high school seniors. Among this elite group, the math section revealed a large sex difference favoring boys. There were twice as many boys as girls with math scores of 500 or higher (out of a possible score of 800), four times as many boys with scores of at least 600, and thirteen times as many boys with scores of at least 700 (putting these test-takers in the top 0.01% of the 12- to 14-year-olds nationwide). However, changes occurred among these junior math wizards from the early 1980s to the mid-1990s: The relative number of girls among them soared during this period. A recent analysis based on the 1.6 million seventh-grade students who took the SAT and ACT as part of the screening process to identify academically precocious youth found that the ratio of boys to girls in the high-ability tail of the math and science portions of these exams has remained steady at between 3:1 to 4:1 since the early 1990s (Makel et al., 2016; Wai et al., 2010). The time period during which the number of girls has risen among the ranks of the mathematically precocious coincides with a trend of special programs and mentoring to encourage girls to take higher-level math and science courses, and with girls participating in high school calculus classes at approximately the same rate as boys (Snyder, Dillow, & Hoffman, 2009, p. 220, Table 149). However, other findings looking at quantitative reasoning abilities among males and females at the high end have actually shown that the ratio of males to females increases rather than decreasing or remaining stable (Lakin, 2013). Additionally, boy math wizards tend to have an ability “tilt” or pattern of abilities which favors math ability over verbal ability, and the ratio of boys to girls with this tilt has remained steady for the past thirty-five years (Wai, Makel, & Hodges, 2017). Differences in these findings could have been due to issues of sample stability, measurement, and/or real changes over time in populations, among other factors. These trends should be tracked across multiple samples and measures as we cannot know whether they will change or stay the same in the future.

Sex Differences across the Life Span

Sex differences in cognitive abilities vary throughout the life span. For example, among young children (ages 4 to 10 years), girls and boys perform similarly on tests of primary mathematical reasoning abilities (Spelke, 2005). During or shortly after elementary school, however, when quantitative tests become more complex and more visual-spatial in nature, sex differences emerge and continue to grow thereafter. By the end of their secondary schooling (12th grade), males demonstrate significantly higher achievement than females in the areas of number properties and operations as well as measurement and geometry (Rampey, Dion, & Donahue, 2009). This trend has remained steady since 1972 (College Board, 2015). Interestingly, females get higher grades than males in school in all subjects, including math, at all grade levels (Kimball, 1989; Snyder et al., 2009; Voyer & Voyer, 2014; Willingham & Cole, 1997). When males and females are compared on tests that reflect content learned in school, such as statewide assessment tests and in-class examinations, the differences disappear. However, it should be noted that these tests tend to evaluate lower-level skills and leave open the possibility of sex differences if higher-order skills were assessed (Hyde et al., 2008). Differences in mathematics favoring males are larger and more commonly found on tests that are not directly tied to the curriculum, such as the SATs, which may reflect novel problem–solving skills. On average, males taking the SATs have consistently scored about a third of a standard deviation higher than girls over the last forty years (data from College Board, 2015; see Halpern et al., 2007, for a review; Perry, 2015). Data suggest that these differences are not because girls take fewer mathematics courses in school, so they remain difficult to interpret (College Board, 2015). However, these data can be misleading because many more females than males take the SATs; lower average scores for females may therefore reflect the greater range of levels of female abilities, especially toward the lower region of the distribution (Hyde et al., 2008). The largest difference is found at the right-hand tail of the distribution, with small differences for most females and males who score near the means.

Spatial abilities are often categorized into three broad areas – spatial perception (ability to determine spatial relationships with respect to the orientation of one's own body, such as indicating the water level in a tilted glass); spatial visualization (ability to engage in multistep manipulations of spatial information, such as finding figures embedded in borders of larger figures; and mental rotation (ability to imagine what a complex figure would look like if it were in another orientation). Sex differences are smaller for visual-spatial working memory (d = 0.09 to 0.22; Voyer, Voyer, & Saint-Aubin, 2016), spatial perception (d = 0.04 to 0.84), and spatial visualization (d = 0.24 to 0.50) than for mental rotation (d = 0.50 to 0.96; Linn & Petersen, 1985). Given these results, most of the research in cognitive sex differences has focused on mental rotation tasks. For mental rotation, a visual-spatial skill that is related to some types of mathematics, such as geometry and topology, males demonstrate an advantage across the life span, especially when figures are three-dimensional. A male advantage in mental rotation, a task that requires participants to imagine what a complex figure would look like if it were rotated in three-dimensional space, is found as early as three to five months of age (Moore & Johnson, 2008, 2011; Quinn & Liben, 2008, 2014), but although several studies have documented these differences in the mental rotation abilities of infants, some have not (e.g., Frick & Mohring, 2013).

In a review of the preschool literature on sex differences in spatial skills, researchers found that, on average, preschool boys are more accurate than girls at spatial tasks that measure accuracy of spatial transformations (d = 0.31) and score higher on the Mazes subtest of the Wechsler Preschool and Primary Scale of Intelligence (d = 0.30; Levine et al., 1999). More recent data confirm these findings in a 2015 sample with older children from the United States and Bahrain, with higher average scores for boys on the

Block Design and Mazes subtests of spatial ability and higher average scores for girls on Coding, and no overall sex difference in Verbal, Performance, and Full Scale IQ (Bakhiet & Lynn, 2015). Other studies show a male advantage on mental rotation for eight-year-old children, but only on measures of accuracy and not speed (Heil & Jansen-Osmann, 2008). Although this very early difference in the ability to visualize an object that is rotated in space suggests a strong biological basis for the large sex differences in mental rotation, there is also evidence for a large sociocultural/learning contribution. Experience with crawling and manipulating objects can affect mental rotation scores (Schwarzer et al., 2013). In one study, female and male college students were trained with computer games that required the use of spatial visualization skills (with appropriate controls for prior experience and other types of games; Feng, Spence, & Pratt, 2007). As the researchers predicted, this intervention reduced the gap between male and female performance; however, it was not completely eliminated.

Sex differences in mental rotation have been studied for decades and findings have been summarized in several meta-analytic reviews. A review of the sex-differences literature on mental rotation found that male performance exceeds that of females across all age ranges, with the size of the between-sex difference ranging between $d =$ 0.52 and 1.49, and the size of the difference increasing slightly across the life span (Geiser, Lehmann, & Eid, 2008).

Girls begin talking somewhat earlier than boys and have a greater vocabulary at two years of age (Lutchamaya, Baron-Cohen, & Raggatt, 2002). Girls also show better language skills in preschool (C. Blair, Granger, & Razzam, 2005). A recent study confirmed these findings with toddlers between eight and thirty-six months of age (Simonsen et al., 2014). These researchers found that boys lagged behind girls in vocabulary production and comprehension and in grammatical complexity. Based on a review of twenty-four large data sets (including several large representative samples of US students, working adults, and military personnel), Willingham and Cole (1997) concluded that differences are small in the elementary school grades, with only writing, language use, and reading favoring females at fourth grade, $d > 0.2$. In the United States, by the end of high school, the largest differences, again favoring females, are found for writing (d between 0.5 and 0.6) and language usage (d between 0.4 and 0.5). Another report on writing proficiency for children in grades 4, 8, and 11 in 1984, 1988, and 1990 showed that girls were better writers in each of the nine comparison groups (US Department of Education, 1997). After a comprehensive review of the literature on writing skills, Hedges and Nowell (1995) concluded, "The large sex differences in writing . . . are alarming. These data imply that males are, on average, at a rather profound disadvantage in the performance of this basic skill" (p. 45). A recent analysis of adolescents confirms that girls outperform boys on tasks of writing (Reynolds et al., 2015). Females also exhibit an early advantage over males in writing among the most gifted writers who score in the right-hand tail of the distribution (Wai et al., 2010).

We know that sex differences on many cognitive tests are reliably found by adolescence. Researchers looked at whether the sex differences found at adolescence would differ as a function of sex hormones or puberty development for girls and boys between the ages of twelve and fourteen (Herlitz et al., 2013). They found the usual sex differences in cognitive tasks for these young adolescents (girls performing better on a verbal episodic memory task – better memory for words presented orally and visually; girls better at face recognition and verbal fluency, and boys performing better on mental rotation), with no effect for hormone levels or self-rated puberty maturation.

In a study of sex differences across the adult life span, Maitland and colleagues analyzed data from the Seattle Longitudinal Study (Maitland et al., 2000). These researchers grouped participants into three age categories at the start of the study: younger (22–49), middle-aged (50–63), and older (64–87). They then tracked their performance on six cognitive ability tests over seven years. Women in the younger and middle-aged groups performed better than men on processing speed. Across all age groups, women performed better than men on verbal recall and men performed better than women on spatial orientation. There were no sex differences in inductive reasoning, verbal comprehension, or numerical facility. Research that looks at elderly populations generally finds that all cognitive abilities decline with age (e.g., Gerstorf, Herlitz, & Smith, 2006; Ritchie et al., 2016). A recent longitudinal study of clinically normal adults found a female advantage in fluency tasks (e.g., name as many words as you can that begin with a selected letter or have a similar meaning to another word) and higher performance for males on some tests of visual-spatial abilities, with a steeper decline for men over time (McCarrey et al., 2016).

Sex Differences over Time

There has been speculation over the possibility that sex differences in cognitive abilities are decreasing, possibly as a result of decreased pressure to conform to sex-role stereotypes (e.g., Baker & Jones, 1992; Corbett, Hill, & St. Rose, 2008). In an extensive meta-analytic review of tests of reading, writing, math, and science, Hedges and Nowell (1995) concluded, "Contrary to the findings of small scale studies, these average differences do not appear to be decreasing, but are relatively stable across the 32-year period investigated" (p. 45). Often the basis of claims that sex differences are decreasing over time comes from evidence of more flexible sex-role stereotypes and socialization practices. A recent study of gender stereotypes in advertising found the biggest change over time in the way men are depicted in commercials (Grau & Zotos, 2016). Men are now shown in "softer" and more egalitarian roles, but there is also a growing number of commercials that depict "empowered" women. It is a long leap to go from female and male depictions in advertisements to cognitive abilities, but there is clear evidence that gender roles have changed over the past several decades and probably will continue to evolve.

Why?

Evolutionary Perspectives

For evolutionary psychologists, the answer to the "why" questions of sex differences lies in the division of labor in hunter-gatherer societies (Buss, 1995; Geary, 2007).

Proponents of this perspective base their claims on evidence that males in early human societies roamed over large areas in their hunt for the animals that provided protein for the community, whereas females gathered crops and traveled shorter distances because much of their adult lives were spent in pregnancy, nursing, and child care. Through the evolutionary pressures of adaptations, males developed brain structures that supported the cognitive and motor skills needed in navigating large areas and killing animals.

The underlying idea is that traits that influence mating success and everyday life in hunter-gatherer societies become enhanced, which results in the emergence of sex differences (Geary, 2017). Geary (1996) made a distinction between those skills that are primary – skills that were shaped by evolutionary pressures and therefore would be found across cultures and developed universally in children's play – and those that are secondary, skills found only in technologically complex societies (i.e., skills such as reading and spelling that are important in school, but would not have evolved in hunter-gatherer societies). Most of the cognitive skills that we can observe today are thought to be built on earlier adaptive solutions for functioning in a specific cultural context rather than directly resulting from evolution (Geary, 2007).

Although theories that posit evolutionary origins for complex human behaviors offer interesting alternatives to nature-nurture dichotomies, they are untestable and ignore large bodies of data that do not conform to these explanatory frameworks. Virtually any finding can be explained by hypothesizing how that difference might have been advantageous to hunter-gatherers. For example, evolutionary theorists criticized Hyde's (2005) analysis of the relationship between psychosocial variables and sex differences for not considering the larger picture. They also used her findings as evidence for their own theories by arguing that social mores exert selection pressures for sex-typed traits, resulting in observed sex differences (e.g., A. P. C. Davies & Sheckelford, 2006). Evolutionary theories ignore the fact that women have always engaged in spatial tasks and they have often had to travel long distances to gather food because plants ripen in different locations in different seasons. Additionally, there is archaeological evidence that women played significant roles in hunting and warfare (Adler, 1993). Typical "women's work" like basket weaving and cloth- and shelter-making are spatial tasks that were very important to the survival of a community because success at gathering depended on the quantity and strength of the baskets, and the protections afforded by clothing and shelters was critical. Furthermore, the visual-spatial tasks that show the largest sex differences favoring males, such as mental rotation, are performed in small arenas of functioning (paper-and-pencil tasks or on a computer screen), which are qualitatively different from finding one's way over miles of territory.

Biological Perspectives

Researchers have identified three mutually influencing biological systems that could account for cognitive sex differences: (1) chromosomal or genetic determinants of sex; (2) sex hormones secreted from endocrine glands and other systems; and (3) structure, organization, and function of the brain (Halpern, 2011). Each of these systems and its effects are the topic of large bodies of research and introduce a few of the possibilities for sex differences as a result of biological processes. First, it is important to note that because these systems are interrelated in most individuals, it is difficult to isolate the relative influence of each. For example, chromosomes determine the type of sex hormones that are secreted. Sex hormones then influence brain development and the development of internal reproductive organs and external genitalia (Halpern, 2011). Research in the area of biological underpinnings of sex differences in intelligence is increasing at a rapid rate, as seen in the dedication of an entire issue of a mainstream neuroscience journal that is dedicated to these questions (*Journal of Neuroscience Research*, Cahill, 2017).

Genes, Hormones, and Brains

Genetic theories emphasize that males and females both inherit intelligence (Schmidt & Hunter, 2004) and possess separate mental capacities related to verbal and spatial abilities. Genetic studies of sex differences in intelligence seek out links between the X and Y chromosomes (males are XY, females are XX) and cognitive abilities. It is well established that some types of intellectual disabilities are linked to the sex chromosomes, which explains the disproportionate numbers of males who are mentally retarded (Skuse, 2005). Johnson, Carothers, and Deary (2009) proposed an X-linked basis for high intelligence. The hypothesized relationship between genes that are responsible for high intelligence and their location on the sex chromosomes is purely speculative, with good evidence supporting the notion that high intelligence must result from the simultaneous influences of many, perhaps hundreds, of genes that are located on many chromosomes (Turkheimer & Halpern, 2009).

Three sex hormones – estrogen, progesterone, and testosterone – have primarily been investigated with respect to their influence on sex differences in cognitive abilities (e.g., Sherwin, 2003). Sex hormones act throughout the brain via hormone receptors within neurons. Females, in general, possess much higher concentrations of estrogen and progesterone, whereas males possess higher concentrations of androgens, the most common of which is testosterone. In addition, these hormones convert from one to another via chemical processes in the brain. At various stages of life, sex hormones play an important role in brain development and subsequent cognition and behavior (e.g., Halpern & Tan, 2001), but adult-level fluctuations in hormones most likely have a minute effect on intelligence.

Normally in humans, the genetic code determines whether the undifferentiated gonads will become ovaries or testes. If development is in the male direction, approximately seven weeks after conception, the newly formed testes will secrete androgens, primarily testosterone and dihydrotestosterone. If ovaries are formed, they will develop approximately twelve weeks following conception and secrete estrogens (e.g., estradiol) and progestins (e.g., progesterone).

Although these hormones are commonly referred to as male and female hormones, all three are found in both females and males (Collaer & Hines, 1995). As these hormones circulate through the bloodstream, they are converted by enzymes into chemical structures that are important in the formation of the brain and internal and external sex organs.

Brain structure, organization, and function are complicated and greatly influenced by hormones. Broadly, there is some evidence that different areas of the brain are activated for males and females during cognitive tasks, and that the overall size and shape of some portions of the brain are different between the sexes (Giedd et al., 1997; Ruigrok et al., 2014). Among the many areas that show on average sex differences, researchers found that males have larger gray-matter volume in the amygdala, hippocampi, and temporal areas, and females have larger volumes at the right frontal area, thalami, and portions of the parahippocampus and occipital cortex. But these data are "messy," with many variations in measurement and in the procedures used to assess them.

There is an abundance of data showing that there are several areas of the brain that differ, on average, between males and females. We believe that most, if not all, neuroscientists would agree with this statement. But the more controversial and theoretically important questions are (1) whether there is a distinctly recognizable male or female brain or whether it is more accurate to think of these differences as areas of overlap along a continuum, and (2) does it matter? It is easier to find some areas that differ (on average) between female and male brains, and much more difficult to link any of these differences to intelligence.

One brain area where there has been intense disagreement among researchers is whether females and males differ in the size and/or shape of some portions of the corpus callosum. The reason why so much effort has been centered on this very large tract of fibers is that a larger size for females would imply better connectivity between the two cerebral hemispheres, on average (Innocenti, 1994), and would also support the theory that female brains are more bilaterally organized in their representation of cognitive functions (Jancke & Steinmetz, 1994). One reason why researchers often disagree about their findings is that the corpus callosum has a highly irregular shape and there are several different ways to measure each of its several portions. One recent study concluded that there are statistically significant, but "subtle" differences (Bjornholm et al., 2017). But a meta-analysis of research on the corpus callosum found the sex differences were eliminated when corrections were made for overall brain size (Tan et al., 2016).

The largest single-sample study of structural and functional sex differences in the human brain to date recently found that males had higher cortical and subcortical volumes, cortical surface areas, and white matter diffusion directionality, whereas females had thicker cortices and higher white matter tract complexity (Ritchie et al., 2018). But again, we caution readers to ask the big questions – knowing that there are some average differences between the sexes in some areas of the brain does not tell us much (if anything at all) about the importance of these differences in human intelligence. Nor do these data imply that females or males have a "better" brain.

Exciting advances in brain imagery have shown that there are also different patterns of activity in male and female brains when they are engaged in some cognitive tasks (Haier et al., 2005). We conclude that for the most part, studies have shown that the brains of males and females are not very different – there are "mosaics" of features, some more common in females, and some more common in males, with considerable overlap (Joel et al., 2015). Thus, there is no distinctly different female or male brain, but a continuum of differences, with much overlap between the sexes. Because the brain reflects learning and other experiences, it is possible that many of the sex differences in the brain that are reported in the literature are influenced by the differences in life experiences that are typical for women and men.

Causal links between prenatal hormones and sex differences in brain structures and organization have been determined in several different ways, including experimental manipulations with nonhuman mammals (e.g., administering testosterone, estrogens, or both, prenatally and perinatally, and removing naturally occurring hormones from the prenatal and perinatal environment).

Individuals with various diseases that cause over- or underproduction of gonadal hormones either prenatally or later in life show cognitive patterns that are in the direction predicted by the data from individuals with normal hormone levels. For example, girls exposed to high levels of prenatal androgens (congenital adrenal hyperplasia) are raised as girls from birth and have normal female hormones starting at birth, yet they tend to show male-typical cognitive patterns and other male-typical behaviors such as preferences for "boys' toys" and rough play, and an increased incidence of sexual orientation toward females (Berenbaum, Korman, & Leveroni, 1995). Females exposed to high levels of prenatal androgens perform at high levels on visual-spatial tasks; their performance is comparable to that of same-aged males and better than the performance of control females (Mueller et al., 2008). A meta-analysis of nine studies found that females with congenital adrenal hyperplasia (CAH) had better spatial performance than control females (Puts et al., 2008). This finding was replicated in a study that compared boys and girls with CAH to their unaffected same-sex siblings (Berenbaum, Bryk, & Beltz, 2012). Girls with CAH (high prenatal androgen) performed better than their unaffected sisters on tests of three-dimensional mental rotation, geography, and mechanical knowledge. But these results also show that it is not simply a matter that more androgens in prenatal life correlate with better spatial performance, because CAH boys scored lower on these tests than their unaffected brothers. There are several possibilities for the failure to find positive effects for CAH boys, including the likelihood that they are less likely to be identified early in life and that they may actually have lower levels of androgens in early infancy because of negative feedback loops.

Other evidence supports the idea that prenatal hormones can influence cognition. Females with fraternal male twins, who would have had higher levels of prenatal androgens than females with a female twin or singletons, had better mental rotation performance than did controls (Tapp, Maybery, & Whitehouse, 2011). These findings show that prenatal sex hormones manifest long-lasting changes in cognitive functioning. Imperato-McGinley and colleagues (1991) compared individuals with complete androgen insensitivity syndrome (AI) to control male and female family members on the Wechsler Adult Intelligence Scale (WAIS). Results showed that control males and females performed better than their androgen-insensitive counterparts on visual-spatial subtests, but that males, overall, still performed better on these tests than females; however, there were no group differences in Full Scale IQ.

Although research has shown effects for prenatal hormones and those that surge early in infancy and again at adolescence, the evidence is mixed, showing that testosterone and estrogen continue to play critical roles in sex-typical cognitive abilities throughout the life span in normal populations. Highly publicized studies have shown that women's cognitive abilities and fine motor skills fluctuate in a reciprocal fashion across the menstrual cycle, but the effects, if real, are most likely extremely small (Hampson, 1990; Hampson & Kimura, 1988).

With an aging population in the United States, there has been much interest in the role of steroidal hormones throughout adulthood and well into old age. Large numbers of postmenopausal women and comparably aged men are treated with various sex hormones for a wide range of possible benefits, including better sexual responsivity and cognitive enhancement. Studies published during the end of the twentieth century and early in the twenty-first century seemed to support this idea (e.g., LeBlanc et al., 2001; Ryan et al., 2009) and some suggested that hormone replacement therapy might postpone dementia for some women, but other researchers have found negative effects for hormone replacement therapy, with at least one study reporting an increased risk of dementia (see Low & Ansley, 2006, for a review). It is likely that the effects of hormone therapies on cognition depend on multiple variables, including age, type and dosage of hormones, timing of hormone therapy (i.e., soon after menopause or decades after menopause), and different cognitive assessments (Luine, 2008). Much more research is needed to untangle the multiple variables that determine the effect of hormone therapy on intelligence. The results for hormone therapies in older males are also mixed. It is common for older men to take testosterone supplements, but the data on whether such therapies can improve cognitive function are still mixed and much more research is needed (Yeap, 2014). Hormone levels also respond to environmental factors, which blurs the distinction between biological and environmental variables.

Another area of the brain that has been closely linked to human intelligence is the prefrontal cortex (Levy & Goldman-Rakic, 2000; Snow, 2016). As studies on this region of the brain have increased, so have reports of sex differences in the structure and function of the prefrontal cortex (e.g., Locklear et al., 2016). Areas of the prefrontal cortex underline some aspects of working memory, which is the ability to keep information in mind (actively using the information) while performing tasks such as reasoning and decision-making (Spaak et al., 2017). This is an emerging area of research, but like the other reports of (on average) sex differences in different areas of brains, we are still left with many questions, including whether these differences play any substantive role in intelligence differences as a function of sex. So far, the evidence is scant.

Intensive exercise, stress, disease, nutrition, and many other variables cause changes in hormones, which in turn affect behavior and emotions, creating continuous feedback loops between hormone levels and life events. Brain structures also change over the life span in response to both hormonal and environmental events, and the response properties of neurons are modified through experience, even in adulthood (Weaver, 2014).

Sociocultural Perspectives

"Math class is tough"; "I love dressing up"; "Do you want to braid my hair?" (Teen-Talk Barbie's first words). "Attack the Cobra Squad with heavy fire power"; "When I give the orders, listen or get captured" (GI Joe, as cited in Viner, 1994). The massive and decades-old literature on observational learning (Bandura, 1977), social reinforcement (Lott & Maluso, 1993), and the ubiquitous influence of sex-role stereotypes (United Nations Human Rights, 2014) shows that males and females still receive sex-differentiated messages, models, rewards, and punishments. From this perspective, it is the sex-typed practices of the socializing community that are most important in creating and understanding nonreproductive differences between the sexes.

Social Learning Theories and Educational Interventions

Social learning theories are more difficult to test than those involving hormone chemistry and brain structures because the experimental control needed to infer causality is virtually impossible to achieve. There is also the problem of causal-arrow ambiguity when psychologists study messy, real-world variables. Consider, for example, the finding that participation in spatial activities is important in the development of spatial activities, and females engage in fewer spatial activities than males (Kotsopoulos, Zambrzycka, & Makosz, 2017). This sort of finding still leaves open the question of why females engage in fewer spatial activities. It could be because they have been socialized to participate in other activities or because they have less spatial ability than males, on average, and therefore less interest. Of course, both are possible. In this case, an initially small sex difference could be widened by societal practices that magnify differences through differential experiences. Dickens and Flynn (2001) devised a mathematical model that can explain how events in the environment interact with heritability to produce large changes in intelligence.

It is also possible that differences are reduced by education and training. In an experimental test of these possibilities, Sorby and Baartmans (1996) targeted improvement in visual-spatial skills. All first-year engineering students at their university with low scores on a test of visual-spatial ability were encouraged to enroll in a course designed to teach these skills. Enrollment resulted in improved performance in subsequent graphics courses by these students and better retention in engineering programs, which suggests that the effects persisted over time and were of at least some practical significance for both women and men. These spatial training effects have been extended to gifted undergraduates in the sciences, technology, engineering, and mathematics fields (STEM: Miller & Halpern, 2012).

Terlecki (2005) examined the impact of training and practice on performance on mental rotation tasks and found that both men and women improved. Training produced more improvement than simple repetition of the task. However, her findings show that neither practice nor training was enough to reduce gender gaps in mental rotation, as both men and women improved equally. Cherney (2008) measured the effect of training using three-dimensional and two-dimensional computer games on tests of mental rotation. She found that training in general improved mental rotation scores, but women's gains were much greater than men's in this study.

A meta-analysis of the spatial training literature which included a wide variety of spatial training interventions concluded that such training is effective (Uttal et al., 2013). Virtually everyone can improve on cognitive tests if they receive appropriate instruction. These are all learnable skills. Education is one of the most potent variables in predicting level of achievement in a cognitive domain (assuming at least an educable range of mental functioning), and adolescence is a period where brain structures are developing and may be a sensitive period for educational interventions (e.g., Dumontheil, 2016). However, even if spatial training is likely to improve performance on tests of spatial ability, that does not necessarily mean that people's spatial ability is enhanced (e.g., see discussion in Shipstead, Redick, & Engle, 2012). Therefore it remains unclear to what extent and how such training impacts sex differences in cognition, especially in relation to real-world long-term impacts of such cognitive training interventions.

Values, Attitudes, and Interests

There are substantial differences in the values, attitudes, and interests of contemporary males and females, which may help to explain cognitive sex differences. This conclusion is based on studies that used the Allport-Vernon-Lindzey Study of Values (1970) assessment instrument (Lubinski, Schmidt, & Benbow, 1996) over many decades. "Masculine-typical" and "feminine-typical" patterns emerge from the Study of Values instrument, even when intelligence is held constant. Concepts of masculinity and femininity vary along the "people–things" dimension, which compares the extent to which individuals prefer working with people as compared to working with "things," with a higher percentage of females preferring "people" and a higher percentage of males preferring "things" (Su, Rounds, & Armstrong, 2009). In a meta-analysis using data from forty-seven interest inventories, Su, Rounds and Armstrong found that males are significantly more likely to be interested in careers that are realistic (d = 0.84) and investigative (d = 0.26), and women were significantly more likely to be interested in careers that are artistic (d = 0.35), social (d = 0.68), and conventional (d = 0.33). When considering preferences and interests it is not possible to determine the causal factors – individuals may prefer areas where they have better skills or skills may improve because individuals engage in activities in which they are interested. Most likely it is some combination of the two.

One of the most successful models of social learning has incorporated expectancies and motivation as a means for understanding the life choices that people make

(Eccles, 1987). The attributions that people make for their successes and failures, expectations of success, individual aptitudes, strategies, and socialized beliefs work in concert to determine how hard they are willing to work at certain tasks and which tasks they select from the environment. This is a popular theory for examining the relationships among expectancy of success (for example in science), how much someone values success in that area, and related influences on educational and career choices. A 2017 study, for example, tested several possible models that included parent education, math aptitude, self-concept about math ability, interest in math, career plans, and career attainment in math-related fields fifteen years after high school graduation (Lauermann, Tsai, & Eccles, 2017). The researchers found that sex-related differences in beliefs about expectancy for success and how much the students valued math were predictive of career goals. Similarly, another set of researchers hypothesized that the level of control and values would affect gender differences in emotions related to mathematics, even when controlling for prior achievement (Frenzel, Pekrun, & Goetz, 2007). These authors found that even though girls and boys had received similar grades in mathematics, girls reported significantly less enjoyment and pride than boys. They explain their findings in that the emotions described by the females could be attributed to the girls' low competence beliefs and domain value of mathematics, combined with the finding that girls value achievement in mathematics. This is a strong model that links values to achievement-related outcomes. It opens many educational routes for changing the status quo.

Nonconscious Influences: Stereotype Threat and Automatic Activation of Stereotypes

Two approaches to studying the effects of stereotypes have been proposed over the past two and a half decades. The significance of these paradigms lies in the way they demonstrate the unconscious, automatic, and powerful influences that stereotypes have on thought and performance. Steele and Aronson (1995) investigated stereotype threat in African Americans. Their study was based on the notion that "when negative stereotypes targeting a social identity provide a framework for interpreting behavior in a given domain, the risk of being judged by, or treated in terms of, those negative stereotypes can evoke a disruptive state among stigmatized individuals" (P. G. Davies, Spencer, & Steele, 2005, p. 276). In their studies, they manipulated testing conditions so that instructions described a college entrance–type test as either a test of intelligence or an investigation of a research problem. When African Americans were told that their intelligence was being tested, they performed significantly worse than when they were given other instructions. This difference was not found for the White students.

Steele and Aronson's (1995) findings regarding stereotypes of African Americans easily translate to a wide range of stereotypes and were confirmed in a study of female and male differences on a difficult math test (Steele, 1997). Females scored more poorly on a test of mathematics when they were told that the test produced gender differences than when the test was described as being insensitive to gender differences. The participants were not conscious of the effect of these instructions on their performance, but activating their knowledge of negative stereotypes prior to the tests had a substantial negative effect. In another study, women's attitudes toward the sex-stereotyped domains of the arts and mathematics were manipulated through subtle reminders of their gender identity. In both cases, those who were primed of their standing as female demonstrated more negative attitudes toward math and more positive attitudes toward the arts than females in the control condition (Steele & Ambady, 2006).

Stereotype threat paradigms have become a cottage industry in the study of cognitive sex differences. One reason for their popularity is that the premises that underlie this paradigm are firmly rooted in sociocultural messages, expectations, experiences, and gender roles for different areas of cognition. With the large number of individual studies, we need to look at meta-analyses. Doyle and Voyer (2016) examined 224 effect sizes from eighty-six separate studies of stereotype threat on math and spatial cognition. They found a small, but significant negative effect for stereotype threat on mathematics for females. In a meta-analysis of stereotype threat studies with children, Flore and Wicherts (2015) also found a small deleterious effect for girls in math, but they caution that the studies show large variations in outcomes and that the small average effect size calculated "is most likely inflated due to publication bias." There are other strong critiques of the stereotype threat literature in relation to sex differences in performance. Stoet and Geary (2012) reviewed the evidence for stereotype threat as an explanation of the achievement gap in math between the sexes, and concluded that stereotype threat research has many methodological problems (e.g., often no control group) and the evidence is rather weak. Additionally, recent meta-analyses suggest that once these problems are corrected for, and failures to replicate and publication bias have been taken into account, the most likely true effect size for stereotype threat is near zero (e.g., Flore & Wicherts, 2015; Ganley et al., 2013). Thus, we conclude that stereotype threat may be a real experimental phenomenon that, at least in some situations, harms the performance of females in mathematics, but the effect is small and there are other variables that are contributing to these differences.

Banaji and colleagues (Banaji & Hardin, 1996; I. V. Blair & Banaji, 1996; Greenwald & Banaji, 1995) used a different experimental paradigm that also revealed strong effects for stereotype knowledge on how people think. Banaji was primarily interested in understanding the automatic activation of sex-role stereotypes that underlie society's thoughts about females and males. The experimental procedure was varied, but all used tasks in which a prime word was flashed on a screen very quickly (for about 0.25 seconds) followed by a target word. Participants had to respond quickly and accurately in making a judgment about the target word. The prime and target words were either consistent with regard to sex-role stereotypes (e.g., soft–woman), inconsistent with sex-role stereotypes (e.g., soft–man), or neutral. In general, participants responded more quickly and accurately when the target was consistent with the prime than when it was not. Sex-role stereotypes were affecting how the participants decoded simple words, yet the participants were unaware of this powerful influence. Together, these two new types of investigations show that expectancies and group-level beliefs can have effects that are unknown even to the participants. A study of female undergraduates enrolled in a college-level calculus class examined the effects of gender identification and implicit and explicit stereotypes on a mathematical aptitude test (Kiefer & Sekaquaptewa, 2007). These authors found that women with low gender identification and low implicit stereotyping scored best on the mathematical aptitude test and women who scored high on both measures were least inclined to pursue careers in mathematics. Recent findings suggest that having a female teacher may make a difference when girls are confronted with the stereotype of boys doing better than girls in STEM classes (Master, Cheryan, & Meltzoff, 2014) and that by diversifying stereotypes this might reduce their impact as gatekeepers and help interest girls in STEM fields (Cheryan, Master, & Meltzoff, 2015).

An international study of implicit stereotypes that associate science and math abilities with being male has found a linear relationship between implicit stereotyping and the size of the male-female gap in science performance in the countries that participated in the Third International Math and Science Study (TIMSS; Nosek et al., 2009). Explicitly stated stereotypes were unrelated to the gender gap across countries. These data suggest that implicit stereotypes can exert powerful effects on the achievement of girls and boys in multiple countries. More recent findings also linked such stereotypes to women's representation in science (Miller, Eagly, & Linn, 2015).

Cultural Gender Equity

There are large differences among countries on various measures of gender equity, such as the percentage of women employed in research and serving on corporate boards, educational levels of males and females, income, and so on. One prediction is that sex differences on cognitive tasks would be smallest or nonexistent in the most gender-equal countries and cultures. As predicted, sex differences in mathematics are smallest in more gender-equal cultures, but the female advantage in reading grows larger in these cultures (Guiso et al., 2008). The male advantage in spatial tasks also grows larger in countries that are more gender-equal (Lippa, Collaer, & Peters, 2010). Thus, if readers were expecting a simple answer to the multifaceted questions about sex and intelligence, we must disabuse you of this possibility. If you are thinking that these results are a byproduct of some error in measurement or conceptualization of sex differences in intelligence, other areas of psychology report similar results. Sex differences in personality traits also grow larger in egalitarian countries (Schmitt et al., 2015).

It is difficult to reconcile the sex-related cognitive findings in societies that vary along the gender equity dimension. One way of understanding these findings is to consider gender equity as multidimensional (Else-Quest & Grabe, 2012; Miller & Halpern, 2014). For example, in Sweden there is more parity in the salaries of women and men than in other countries, but men and women still tend to pursue different types of occupations (Charles & Bradley, 2009). A more differentiated notion of gender equity may help to solve this puzzle.

Biopsychosocial Model

A biopsychosocial model based on the inextricable links between the biological bases of intelligence and environmental events is an alternative to the nature-nurture dichotomy. Research and debate about the origins of sex differences are grounded in the belief that the nonreproductive differences between men and women originate from sex-differentiated biological mechanisms (nature; e.g., "sex" hormones), socialization practices (nurture; e.g., girls are expected to perform poorly on tests of advanced mathematics), and their interaction. A biopsychosocial model offers an alternative conceptualization: It is based on the idea that some variables are both biological and social and therefore cannot be classified into one of these two categories. Consider, for example, the role of learning in creating and maintaining an average difference between females and males. Learning is both a socially mediated event and a biological process. Individuals are predisposed to learn some topics more readily than others. A predisposition to learn some behaviors or concepts more easily than others is determined by prior learning experiences, the neurochemical processes that allow learning to occur (release of neurotransmitters), and change in response to learning (e.g., long-term potentiation and changes in areas of the brain that are active during performance of a task; Posner & Raichle, 1994). Thus, learning depends on what is already known and on the neural structures and processes that undergird learning. Of course, psychological variables such as interest, motivation, and expectancy are also important in determining how readily information is learned, but interest, motivation, and expectancy are also affected by prior learning. The biopsychosocial model is predicated on an integral conceptualization of nature and nurture that cannot be broken into nature or nurture subcomponents. Neural structures change in response to environmental events; environmental events are selected from the environment on the basis of, in part, predilections and expectancies; and the biological and socially mediated underpinnings of learning help to create the predilections and expectancies that guide future learning. This model is depicted in Figure 14.1.

It is true that multiple psychological and social factors play a part in determining career direction. People's individual expectations for success are shaped by their perception of their own skills. One factor in forming our self-perception is how authority figures such as teachers perceive and respond to males and females. A study of London cab drivers found that they had enlarged portions of their right posterior hippocampus relative to a control group of adults. The cab drivers demonstrated a positive correlation between the size of the hippocampus that is activated during recall of complex routes and the number of years they had worked in this occupation, thus showing a "dose-size relationship" that is indicative of environmental influences (Maguire, Frackowiak, & Frith, 1997; Maguire et al., 2000).

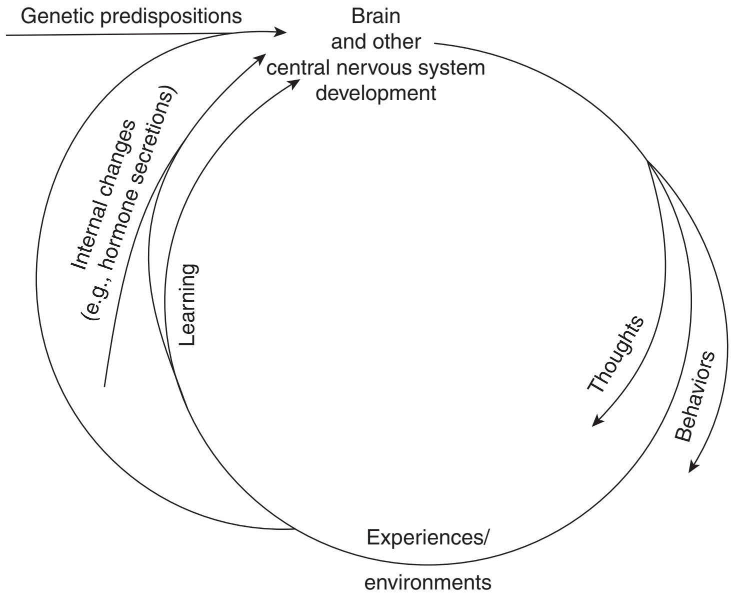

Figure 14.1 *Biopsychosocial model.*

Where We Go from Here

Understanding sex differences in intelligence is crucial to understanding cognition in general and the joint effects of nature and nurture on cognition. The truth about sex differences in intelligence depends on the nature of the cognitive task being assessed, the range of ability that is tested, the age and education of the participants, and numerous other modifying variables. There are intellectual areas in which females, on average, excel and others in which males, on average, excel. Psychological, social, and biological factors explain these differences. However, it does not seem that biology is limiting intelligence in any way because biology alone cannot explain the vast improvement of female performance on certain measures such as the increasing numbers of females scoring at the highest end on the SAT math test (Blackburn, 2004; Wai et al., 2010).

Data showing differences between men and women in intelligence do not support the notion of a smarter sex, nor do they imply that the differences are immutable. There is direct evidence showing that specifically targeted training on cognitive tasks boosts performance for both men and women over the short term. Thus, the application of good learning principles in education can improve intellectual performance for all students. There are no cognitive reasons to support sex-segregated education, especially given the large amount of overlap in test scores for girls and boys on all tests of cognitive ability. The finding that girls get higher grades in school has been linked, at least in part, to better self-regulation and self-discipline, which allows them to delay gratification and behave in ways that are rewarded in classrooms (Duckworth & Seligman, 2006). Self-discipline has been used to explain many outcomes in life because it is critical to learning, especially when the material is complex and requires extended effort. Thus, the ability to self-regulate is rewarded in school grades and necessary for advanced learning. The fact that girls get better grades in every subject in school shows that they are learning at least as well as boys, and the fact that boys score higher on some standardized measures of achievement shows that they are learning at least as well as girls. For those concerned with increasing the number of females in math and science, the problem lies in convincing more females that "math counts" and making academic and career choices that are "math-wise."

The data on intelligence show that both sexes, on average, have their strengths and weaknesses. Nevertheless, the research argues that much can be done to try to help more women excel in science and encourage them to choose it as a profession. We also advocate for programs that will attract more men to what have traditionally been seen as female professions – teaching, nursing, social work, and other areas where males are underrepresented. The challenges are many, requiring innovations in education, targeted mentoring and career guidance, and a commitment to uncover and root out bias, discrimination, and inequality. In the end, tackling these issues will benefit women, men, the economy, and science itself (National Science Board, 2010).

References

Adler, L. L. (Ed.) (1993). *International handbook on gender roles*. Westport, CT: Greenwood.

Allport, G. W., Vernon, P. E., & Lindzey, G. (1970). *Manual for the study of values* (3rd ed.). Boston: Houghton Mifflin.

American Association of University Women. (1992). *The AAUW report: How schools shortchange girls*. New York: Marlowe.

Arden, R., & Plomin, R. (2006). Sex differences in variance of intelligence across childhood. *Personality and Individual Differences*, *41*, 39–48.

Baker, D. P., & Jones, D. P. (1992). *Opportunity and performance: A sociological explanation for gender differences in academic mathematics*. In J. Wrigley (Ed.), Education and gender equality (pp. 193–206). London: Falmer Press.

Bakhiet, S. F. A., & Lynn, R. (2015). Sex differences on the Wechsler Intelligence Scale for Children – III in Bahrain and the United States. *Psychological Reports*, *117*, 794–798.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science*, *7*, 136–141.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science*, *222*, 1029–1030.

Berenbaum, S. A., Bryk, K. K., & Beltz, A. M. (2012). Early androgen effects on spatial and mechanical abilities: Evidence from congenital adrenal hyperplasia. *Behavioral Neuroscience*, *126*, 86–96. https://doi.org/10.1037/a0026652

Berenbaum, S. A., Korman, K., & Leveroni, C. (1995). Early hormones and sex differences in cognitive abilities. Special issue, *Learning and Individual Differences*, *7*, 303–321.

Bjornholm, L., Nikkinen, J., Kiviniemi, V., Nordstrom, T., Niemela, S., Drakesmith, M., et al. (2017). Structural properties of the human corpus callosum: Multimodal assessment and sex differences. *Neuroimage*, *152*, 108–118.

Blackburn, C. C. (2004). Developing exceptional talent: Descriptive characteristics of highly precocious mathematical and verbal reasoners. Paper presented at the Seventh

Biennial Henry B. and Jocelyn Wallace National Research Symposium on Talent Development, University of Iowa, Iowa City, May.
Blair, C., Granger, D., & Razzam, R. P. (2005). Cortisol reactivity is positively related to executive function in preschool children attending Head Start. *Child Development*, *76*, 554–567.
Blair, I. V., & Banaji, M. R. (1996). Automatic controlled processes in stereotype priming. *Journal of Personality and Social Psychology*, *70*, 1142–1163.
Braun, K. V. N., Christensen, D., Doernberg, N., Schieve, L., Rice, C., Wiggins, L., et al. (2015). Trends in the prevalence of autism spectrum disorder, cerebral palsy, hearing loss, intellectual disability, and vision impairment, Metropolitan Atlanta, 1991–2010. *PLoS One*. https://doi.org/10.1371/journal.pone.0124120
Brody, N. (1992). *Intelligence* (2nd ed.). New York: Academic Press.
Buss, D. M. (1995). Psychological sex differences: Origins through sexual selection. *American Psychologist*, *50*, 164–168.
Cahill, L. (Ed.) (2017). An issue whose time has come: Sex/gender influence on nervous system function. *Journal of Neuroscience Research*, *95*, 1–791.
Charles, M., & Bradley, K. (2009). Indulging our gendered selves? Sex segregation by field of study in 44 countries. *American Journal of Sociology*, *114*, 924–976.
Cherney, I. D. (2008). Mom, let me play more computer games: They improve my mental rotation skills. *Sex Roles*, *59*, 776–786.
Cheryan, S., Master, A., & Meltzoff, A. N. (2015). Cultural stereotypes as gatekeepers: Increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in Psychology*, *6*. http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00049/full
Collaer, M. L., & Hines, M. (1995). Human behavioral sex differences: A role for gonadal hormones during early development? *Psychological Bulletin*, *118*, 55–107.
College Board. (2015). *2015 college-bound seniors: Total group profile report*. media.collegeboard.org/digitalServices/pdf/sat/total-group-2015.pdf
Colom, R., Juan-Espinosa, M., Abad, F., & García, L. F. (2000). Negligible sex differences in general intelligence. *Intelligence*, *28*, 57–68.
Corbett, C., Hill, C., & St. Rose, A. (2008). *Where the girls are: The facts about gender equity in education*. Washington: American Association of University Women.
Davies, A. P. C., & Sheckelford, T. K. (2006). An evolutionary psychological perspective on gender similarities and differences. *American Psychologist*, *61*, 641–642.
Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology*, *88*, 276–287.
Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox. *Psychological Review*, *108*, 346–369.
Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences*, *47*, 103–116.
Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, *98*, 198–208.
Dumontheil, I. (2016). Adolescent brain development. *Current Opinion in Behavioral Sciences*, *10*, 39–44.
Dykiert, D., Gale, C. R., & Deary, I. J. (2008). Are apparent sex differences in mean IQ scores created in part by sample restriction and increased male variance? *Intelligence*, *37*, 42–47.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly*, *11*, 135–172.

Else-Quest, N. M., & Grabe, S. (2012). The political is personal: Measurement and application of national-level indicators of gender equity in psychological research. *Psychology of Women Quarterly*, *36*, 131–144.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, *18*, 850–855.

Fine, C. (2017). *Testosterone rex: Myths of sex, science, and society*. New York: Norton.

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology*, *53*, 25–44.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics – A "hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, *22*, 497–514.

Frick, A., & Mohring, W. (2013). Mental object rotation and motor development in 8- and 10-month-old infants. *Journal of Experimental Child Psychology*, *115*, 708–720.

Ganley, C. M., Mingle, L. A., Ryan, A. M., Ryan, K., Vasilyeva, M., & Perry, M. (2013). An examination of stereotype threat effects on girls' mathematics performance. *Developmental Psychology*, *49*, 1886–1897.

Geary, D. C. (1996). Sexual selection and sex differences in mathematical abilities. *Behavioral and Brain Sciences*, *19*, 229–284.

Geary, D. C. (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Educating the evolved mind* (pp. 1–100). Greenwich, CT: Information Age.

Geary, D. C. (2017). Evolutionary framework for identifying sex- and species-specific vulnerabilities in brain development and functions. *Journal of Neuroscience Research*, *95*, 355–361.

Geiser, C., Lehmann, W., & Eid, M. (2008). A note on sex differences in mental rotation in different age groups. *Intelligence*, *36*, 556–563.

Gerstorf, D., Herlitz, A., & Smith, J. (2006). Stability of sex differences in cognition in advanced old age: The role of education and attrition. *Journal of Gerontology: Psychological Sciences and Social Sciences*, *61*, 245–249.

Giedd, J. N., Castellanos, F. X., Rajapakse, J. C., Vaituzis, A. C., & Rapoport, J. L. (1997). Sexual dimorphism of the developing human brain. *Progress in Neuropsychopharmacology and Biological Psychiatry*, *21*, 1185–1901.

Grau, S. L., & Zotos, Y. C. (2016). Gender stereotypes in advertising: A review of current research. *International Journal of Marketing*, *35*, 761–770.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science*, *320*, 1164–1165.

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2005). The neuroanatomy of general intelligence: Sex matters. *NeuroImage*, *25*, 320–327.

Halpern, D. F. (2011). *Sex differences in cognitive abilities* (4th ed.). New York: Psychology Press.

Halpern, D. F., Benbow, C., Geary, D., Gur, D., Hyde, J., & Gernsbacher, M. A. (2007). The science of sex-differences in science and mathematics. *Psychological Science in the Public Interest*, *8*, 1–52.

Halpern, D. F., Eliot, L., Bigler, R. S., Fabes, R. A., Hanish, L. D., Hyde, J., et al. (2011). The pseudoscience of single-sex schooling. *Science*, *333*, 1706–1707.

Halpern, D. F., & Tan, U. (2001). Stereotypes and steroids: Using a psychobiosocial model to understand cognitive sex differences. *Brain and Cognition*, *45*, 392–414.

Hampson, E. (1990). Estrogen-related variations in human spatial and articulatory-motor skills. *Psychoneuroendocrinology*, *15*, 97–111.

Hampson, E., & Kimura, D. (1988). Reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. *Behavioral Neuroscience*, *102*, 456–459.

Hayes, A. R., Pahlke, E. E., & Bigler, R. S. (2011). The efficacy of single-sex education: Testing for selection and peer quality effects. *Sex Roles*, *65*, 693–703.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*, 41–45.

Heil, M., & Jansen-Osmann, P. (2008). Gender differences in math and mental rotation accuracy but not in mental rotation speed in 8-years-old children. *International Journal of Developmental Science*, *2*, 190–196.

Herlitz, A., Reuterskiöld, L., Lovén, J., Thilers, P. P., & Rehnman, J. (2013). Cognitive sex differences are not magnified as a function of age, sex hormones, or puberty development during early adolescence. *Developmental Neuropsychology*, *38*, 167–179. https://doi.org/10.1080/87565641.2012.759580

Hunt, E., & Madhyastha, T. (2008). Recruitment modeling: An analysis and an application to the study of male-female differences in intelligence. *Intelligence*, *36*, 653–663.

Hyde, J. S. (2005). The gender similarity hypothesis. *American Psychologist*, *60*, 581–592.

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, *65*, 373–398.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*, 494–495.

Imperato-McGinley, J., Pichardo, M., Gautier, T., Voyer, D., & Bryden, M. P. (1991). Cognitive abilities in androgen insensitive subjects – Comparison with control males and females from the same kindred. *Clinical Endocrinology*, *34*, 341–347.

Innocenti, G. M. (1994). Some new trends in the study of the corpus callosum. *Behavioral and Brain Research*, *64*, 1–8.

Irwing, P., & Lynn, R. (2005). Intelligence: Is there a difference in IQ scores? *Nature*, *438*, 31–32.

Jancke, L., & Steinmetz, H. (1994). Interhemispheric-transfer time and corpus callosum size. *Neuroreport*, *5*, 2385–2388.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. New York: Praeger.

Joel, D., Berman, Z., Tavor, I., Wexler, N., Gaber, O., Stein, Y., et al. (2015). Sex beyond the genitalia: The human brain mosaic. *Proceedings of the National Academy of Sciences*, *112*, 15468–15473.

Johnson, W., & Bouchard, T. J. (2006). Sex differences in mental abilities: g masks the dimensions on which they lie. *Intelligence*, *35*, 23–59.

Johnson, W., Carothers, A., & Deary, I. J. (2008). Sex differences in variability in general intelligence: A new look at an old question. *Perspectives on Psychological Science*, *3*, 518–531.

Johnson, W., Carothers, A., & Deary, I. J. (2009). A role for the X chromosome in sex differences in variability in general intelligence? *Perspectives in Psychological Science*, *4*, 598–611.

Karpiak, C. P., Buchanan, J. P., Hosey, M., & Smith., A. (2007). University students from single-sex and coeducational high schools: Differences in majors and attitudes at a Catholic university. *Psychology of Women Quarterly*, *31*, 282–289.

Kaufman, S. B. (2007). Sex differences in mental rotation and spatial visualization ability: Can they be accounted for by differences in working memory capacity? *Intelligence*, *35*, 211–223.

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes, gender identification, and math-related outcomes: A prospective study of female college students. *Psychological Science*, *18*, 13–18.

Kimball, M. M. (1989). A new perspective on women's mathematics achievement. *Psychological Bulletin*, *105*, 198–214.

Kotsopoulos, D., Zambrzycka, J., & Makosz, S. (2017). Gender differences in toddlers' visual-spatial skills. *Mathematical Thinking and Learning*, *19*, 167–180.

Lakin, J. M. (2013). Sex differences in reasoning abilities: Surprising evidence that male-female ratios in the tails of the quantitative reasoning distribution have increased. *Intelligence*, *41*, 263–274.

Lauermann, F., Tsai, Y., & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy-value theory of achievement-related behaviors. *Developmental Psychology*, *53*, 1540–1559.

LeBlanc, E. S., Janowsky, J., Chan, B. K., & Nelson, H. D. (2001). Hormone replacement therapy and cognition: Systematic review and meta-analysis. *Journal of American Medical Association*, *285*, 1489–1499.

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, *35*, 940–949.

Levy, R., & Goldman-Rakic, P. S. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. *Experimental Brain Research*, *133*, 23–32.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*, 1479–1498.

Lippa, R. A., Collaer, M. L., & Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Archives of Sexual Behavior*, 39(4), 990–997. https://doi.org/10.1007/s10508-008-9460-8

Locklear, M. N., Cohen, A. B., Jone, A., & Kritzer, M. F. (2016). Sex differences distinguish intracortical glutamate receptor-mediated regulation of extracellular dopamine levels in the prefrontal cortex of adult rats. *Cerebral Cortex*, *26*, 599–610.

Loewen, J. W., Rosser, P., & Katzman, J. (1988). Gender bias in SAT items. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, April 5–9.

Loring-Meier, S., & Halpern, D. F. (1999). Sex differences in visual-spatial working memory: Components of cognitive processing. *Psychonomic Bulletin and Review*, *6*, 464–471.

Lott, B., & Maluso, D. (1993). The social learning of gender. In A. E. Beall & R. J. Sternberg (Eds.), *The psychology of gender* (pp. 99–123). New York: Guilford Press.

Low, L.-F., & Ansley, K. J. (2006). Hormone replacement therapy and cognitive performance in postmenopausal women – a review by cognitive domain. *Neuroscience and Biobehavioral Reviews*, *30*, 66–84.

Lubinski, D., Schmidt, D. B., & Benbow, C. P. (1996). A 20-year stability analysis of the Study of Values for intellectually gifted individuals from adolescence to adulthood. *Journal of Applied Psychology*, *81*, 443–451.

Luine, V. N. (2008). Sex steroids and cognitive function. *Journal of Neuroendocrinology*, *20*, 866–872.

Lutchamaya, S., Baron-Cohen, S., & Raggatt, P. (2002). Foetal testosterone and vocabulary size in 18- and 24-month-old infants. *Infant Behavior and Development*, *24*, 418–424.
Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1997). Recalling routes around London: Activation of the right hippocampus in taxi drivers. *Journal of Neuroscience*, *17*, 7103–7110.
Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Ashburner, C. D., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Science*, *97*, 4398–4403.
Maitland, S. B., Intrieri, R. C., Schaie, K. W., & Willis, S. L. (2000). Gender differences and changes in cognitive abilities across the adult life span. *Aging, Neuropsychology, and Cognition*, *7*, 32–53.
Makel, M. C., Wai, J., Peairs, K. F., & Putallaz, M. (2016). Sex differences in the right tail of cognitive abilities: An update and cross cultural extension. *Intelligence*, *59*, 8–15.
Master, A., Cheryan, S., & Meltzoff, A. N. (2014). Reducing adolescent girls' concerns about STEM stereotypes: When do female teachers matter? *International Review of Social Psychology*, *27*, 79–102.
McCarrey, A. C., An, Y., Kitner-Triolo, M. H., Ferrucci, L., & Resnick, S. M. (2016). Sex differences in cognitive trajectories in clinically normal older adults. *Psychology and Aging*, *31*, 166–175.
Miller, D. I., Eagly, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology*, *107*, 631–644.
Miller, D. I., & Halpern, D. F. (2012). Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learning and Individual Differences*, *26*, 141–152.
Miller, D. I., & Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends in Cognitive Science*, *18*, 37–45.
Moore, D. S., & Johnson, S. P. (2008). Mental rotation in human infants: A sex difference. *Psychological Science*, *19*, 1063–1066.
Moore, D. S., & Johnson, S. P. (2011). Mental rotation of dynamic, three-dimensional stimuli by 3-month-old infants. *Infancy*, *16*, 435–445.
Mueller, S. C., Temple, V., Oh, E., VanRyzin, C., Williams, A., Cornwell, B., et al. (2008). Early androgen exposure modulates spatial cognition in congenital adrenal hyperplasia. *Psychoneuroendocrinology*, *33*, 973–980.
National Science Board. (2010). *Preparing the next generation of STEM innovators: Identifying and developing our nation's human capital*. www.nsf.gov/nsb/publications/2010/nsb1033.pdf
National Science Foundation. (2014). STEM education data and trends 2014. How do US 15-year-olds compare with students from other countries in math and science? www.nsf.gov/nsb/sei/edTool/data/highschool-08.html
Nosek, B. A., Smyth, F. L., Sriram, N., Linder, N. M., Devos, T. A., Thierry, A., et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Science*, *106*, 10593–10597.
Nyborg, H. (2005). Sex-related differences in general intelligence g, brain size, and social status. *Personality and Individual Differences*, *39*, 497–509.
Pahlke, E., Hyde, J. S., & Allison, C. M. (2014). The effects of single-sex compared with coeducational schooling on students' performance and attitudes: A meta-analysis. *Psychological Bulletin*, *140*, 1042–1072.

Paus, T., Wong, A. P-Y., Syme, C., & Pausova, Z. (2017). Sex differences in the adolescent brain and body: Findings from the Saguenay youth study. *Journal of Neuroscience Research*, *95*, 362–370.

Perry, M. J. (2015). 2015 SAT test results confirm pattern that's persisted for 40+ years – high school boys are better at math than girls. AEIdeas. www.aei.org/publication/2015-sat-test-results-confirm-pattern-thats-persisted-for-40-years-high-school-boys-are-better-at-math-than-girls

Posner, M. I., & Raichle, M. E. (1994). *Images of mind*. New York: Freeman.

Puts, D. A., McDaniel, M. A., Jordan, C. L., & Breedlove, S. M. (2008). Spatial ability and prenatal androgens: Meta-analyses of congenital adrenal hyperplasia and digit ratio (2D:4D) studies. *Archives of Sexual Behavior*, *37*, 100–111.

Quinn, P. C., & Liben, L. S. (2008). A sex difference in mental rotation in young infants. *Psychological Science*, *19*, 1067–1070.

Quinn, P. C., & Liben, L. S. (2014). A sex difference in mental rotation in infants: Convergent evidence. *Infancy*, *19*, 103–116.

Rampey, B. D., Dion, G. S., & Donahue, P. L. (2009). NAEP trends in academic progress (NCES 2009–479). National Center for Education Statistics, Institute of Education Sciences, US Department of Education, Washington.

Reynolds, M. R., Scheiber, C., Hajovsky, D. B., Schwartz, B., & Kaufman, A. S. (2015). Gender differences in academic achievement: Is writing an exception to the gender similarities hypothesis? *Journal of Genetic Psychology*, *176*, 211–234.

Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., et al. (2018). Sex differences in the adult human brain: Evidence from 5,216 UK Biobank participants. *Cerebral Cortex*, *28*(8), 2959–2975.

Ritchie, S. J., Tucker-Drob, E. M., Cox, S. R., Corley, J., Dykiert, D., Redmond, P., et al. (2016). Predictors of ageing-related decline across multiple cognitive functions. *Intelligence*, *59*, 115–126. https://doi.org/10.1016/j.intell.2016.08.007

Ruigrok, A. N. V., Salimi-Khorshidi, G., Lai, M-C., Baron-Cohen, S., Lombardo, M. V., Tait, R. J., & Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioral Reviews*, *39*, 34–50.

Ryan, J., Carriere, I., Scali, J., Ritchie, K., & Ancelin, M-L. (2009). Life-time estrogen exposure and cognitive functioning in later life. *Psychoneuroendocrinology*, *34*, 287–298.

Sadker, M., & Sadker, D. (1995). *Failing at fairness: How our schools cheat girls*. New York: Touchstone.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, *9*, 185–211.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162–173.

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2015). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, 94, 168–182. https://doi.org/10.1037/0022-3514.94.1.168.

Schwarzer, G., Freitag, C., Buckel, R., & Lofruthe, A. (2013). Crawling is associated with mental rotation ability by 9-month-old infants. *Infancy*, *18*, 432–441.

Sherwin, B. (2003). Estrogen and cognitive functioning in women. *Endocrine Reviews*, *24*, 133–151.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628–654.

Simonsen, H. G., Kristoffersen, K. E., Bleses, D., Wehberg, S., & Jørgensen, R. N. (2014). The Norwegian Communicative Development Inventories: Reliability, main developmental trends and gender differences. *First Language*, *34*(1), 3–23. https://doi.org/10.1177/0142723713510997

Skuse, D. (2005). X-linked genes and mental functioning. *Human Molecular Genetics*, *14*, R27–R32.

Snow, P. J. (2016). The structural and functional organization of cognition. *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2016.00501

Snyder, T. D., Dillow, S. A., & Hoffman, C. M. (2009). *Digest of Education Statistics 2008* (NCES 2009–020). National Center for Education Statistics, Institute of Educational Sciences, US Department of Education, Washington.

Sommers, C. H. (2000). The war against boys. *The Atlantic*. www.theatlantic.com/magazine/archive/2000/05/the-war-against-boys/304659

Sorby, S. J., & Baartmans, B. J. (1996). The development and assessment of a course for enhancing the 3-D spatial visualization skills of first year engineering students. *Engineering Design Graphics Journal*, *60*, 13–20.

Spaak, E., Watanabe, K., Funahashi, S., & Stokes, M. G. (2017). Stable and dynamic coding for working memory in primate prefrontal cortex. *Journal of Neuroscience*, *37*, 6503–6516.

Spelke, E. S. (2005). Sex difference in intrinsic aptitude for mathematics and science? A critical review. *American Psychologist*, *60*, 950–958.

Spinath, F. M., Spinath, B., & Plomin, R. (2008). The nature and nurture of intelligence and motivation in the origins of sex differences in elementary school achievement. *European Journal of Personality*, *22*, 211–229.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613–629.

Steele, J. R., & Ambady, N. (2006). “Math is hard!” The effect of gender priming on women’s attitudes. *Journal of Experimental Social Psychology*, *42*, 428–436. http://dx.doi.org.ccl.idm.oclc.org/10.1016/j.jesp.2005.06.003

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811.

Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology*, *16*, 93–102.

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within and across-nation assessment of 10 years of PISA data. *PLoS One*, *8*, e57988.

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, *135*, 859–884.

Tan, A., Ma, W., Vira, A., Marwha, D., & Eliot, L. (2016). The human hippocampus is not sexually-dimorphic: Meta-analysis of structural MRI volumes. *Neuroimage*, *124* (Part A). 350–366. https://doi.org/10.1016/j.neuroimage

Tapp, A. L., Maybery, M. T., & Whitehouse, A. O. (2011). Evaluating the twin testosterone transfer hypothesis: A review of the empirical evidence. *Hormones and Behavior*, *60*, 713–722.

Terlecki, M. S. (2005). The effects of long-term practice and training on mental rotation. *Dissertation Abstracts International*, *65*(10–B), 5434.

Torres, A., Gómez-Gil, E., Vidal, A., Puig, O., Boget, T., & Salamero, M. (2006). Gender differences in cognitive functions and the influence of sex hormones. *Actas Españolas de Psyiquiatria*, *34*, 408–415.

Turkheimer, E., & Halpern, D. F. (2009). Sex differences in variability for cognitive measures: Do the ends justify the genes? *Perspectives on Psychological Science*, *4*, 612–614.

United Nations Human Rights. (2014). The harms of gender stereotyping. www.ohchr.org/EN/NewsEvents/Pages/GenderStereotyping.aspx

US Department of Education. (1997). *National assessment of educational progress (Indicator 32: Writing Proficiency; prepared by the Educational Testing Service)*. Washington: US Department of Education.

US Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Research. (2005). *Single-sex versus coeducational schooling: A systematic review*. Washington: Author.

US Department of Education, National Center for Education Statistics. (2017). The Condition of Education 2017 (NCES 2017–144). https://nces.ed.gov/pubs2017/2017144.pdf

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., et al. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*, 352–402.

Viner, K. (1994). Live Issues. *Cosmopolitan*, p. 105.

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*, 1174–1204.

Voyer, D., Voyer, S. D., & Saint-Aubin, J. (2016). Sex differences in visual-spatial working memory: A meta-analysis. *Psychonomic Bulletin and Review*, *24*, 307–334.

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30-year examination. *Intelligence*, *38*, 412–423.

Wai, J., Makel, M. C., & Hodges, J. (2018). Sex differences in ability tilt in the right tail of cognitive abilities: A 35-year examination. *Intelligence*, *67*, 76–83.

Weaver, J. (2014). How brain circuits adapt to changes in sensory experience. *PLoS Biol*, *12*, e1001802.

Williams, W. M., & Ceci, S. J. (2014). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*, *112*, 5360–5365.

Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment*. Mahwah, NJ: Erlbaum.

Yeap, B. B. (2014). Hormonal changes and their impact on cognition and mental health of ageing men. *Maturitas*, *79*, 227–235.

15 Racial and Ethnic Group Differences in Intelligence in the United States

Multicultural Perspectives

Lisa A. Suzuki, Dylan Larson-Konar, Ellen L. Short, and Christina S. Lee

The road to understanding the relationship between culture and intelligence has been a bumpy one, impacted by the sheer complexity of the concept and plagued by controversy, especially when applied to members of marginalized racial and ethnic groups (Armour-Thomas & Suzuki, 2016). Debates regarding the usage of intelligence measures with minority groups have continued for over a century. Thousands of articles have been published to illuminate the linkage between cultural factors (e.g., acculturation, racial identity, ethnic identity, cultural intelligence) and measures of intelligence.

Understanding the relationship between culture and intelligence has real-world implications for members of the racially and ethnically diverse communities that reside in the United States and abroad. Intelligence tests have been touted as the "gold standard" in psychological and educational testing arenas. Scores on mainstream intelligence tests are moderately predictive of school grades, work performance, and other indicators of life success (Nisbett et al., 2012). While many view intelligence as a stable and fixed entity, others have argued for its malleability (Suzuki & Aronson, 2005). This chapter will address multicultural perspectives of intelligence in the United States. We will focus our attention on the following: definitions of relevant concepts; fairness in testing; environment, social location, and cultural context; measures of intelligence; and outcome implications in testing ethnocultural populations.

Defining the Relevant Concepts

Multiple definitions of culture and intelligence have emerged in the literature. In the following sections we highlight the definitions of terms that will serve as the foundation of our discussion in this chapter. We are aware that in selecting a limited set of definitions we exclude other perspectives.

Constructs Related to Culture, Context, and Intelligence

While hundreds of definitions of culture are found in the literature (Kroeber & Kluckhohn, 1963), one of the most frequently cited definitions in the social sciences literature comes from Geertz's (1973) text *The Interpretation of Cultures* (1973):

> [Culture] denotes a historically transmitted pattern of meanings embodied in symbols, a system of inherited conceptions expressed in symbolic forms by means of which men communicate, perpetuate, and develop their knowledge about and attitudes toward life. (p. 89)

Serpell (2000) elaborates by stating:

> Culture consists of a set of practices (constituted by a particular pattern of recurrent activities with associated artifacts) that are informed by a system of meanings (encoded in language and other symbols) and maintained by a set of institutions over time. (p. 549)

The dynamic and complex nature of culture makes measurement of the construct challenging (e.g., López & Guarnaccia, 2000). Individuals often belong to different cultures simultaneously and possess multiple intersecting identities over their lifetime. For example, Goldberger and Veroff (1995) define culture as a common set of experiences related to numerous variables, including geographic boundaries, language, race, ethnicity, religious belief, social class, gender, sexual orientation, age, and ability status. In 2017, the American Psychological Association (APA) adopted new multicultural guidelines emphasizing context, identity, and intersectionality from an ecological approach. These guidelines address cultural competence from an individual perspective that considers: the fluid and complex nature of identity and self-definition; conceptualizations reflecting categorical assumptions and biases; language and communication; social and physical environments; experiences with power, privilege, and oppression; culturally adaptive interventions and advocacy; professional practices in an international context; developmental stages and transitions intersecting with biosociocultural context; culturally appropriate informed research and practice (citing assessment), and that builds on a strengths-based approach. Each of the guidelines can be applied to the testing and assessment process that forms a major cornerstone of psychological practice.

An essential point reiterated in the latest formulation of the multicultural guidelines and through the general literature is that culture provides a context in which people develop and learn. Therefore, it is difficult to define intelligence without first understanding the individual's sociocultural context.

Intelligence

All definitions of intelligence contain reference to cognitively based abilities such as abstract thinking, reasoning, problem-solving, and acquisition of knowledge (Snyderman & Rothman, 1988). It is critical to integrate these traditional characteristics into a more nuanced conceptualization of intelligence that accounts for the pervasive role of culture. The question of what determines intelligence is best considered within a cultural context, where certain strengths may be more or less valued in specific environments, but not in others (Cole, 2017; Rogoff et al., 2017, Sternberg, 2017). What follows is an argument that "true" intelligence and "measured" intelligence are not necessarily equivalent (Whitaker, 2015), especially in environments that emphasize skills that are less valued by conventional standardized testing.

Accounting for the importance of understanding cognition within a cultural context, Sternberg (1996) describes "successful intelligence" as "The ability to choose and sometimes re-choose a life course that is prosocial, personally meaningful, and self-fulfilling, and that enables one to capitalize on one's strengths and to compensate for or correct one's weaknesses, in order to adapt to, shape, and select environments" (p. 3). It is important to note that there are a number of intelligences (e.g., Ceci, 1996; Gardner, 1983; Sternberg, 1996), among which conventionally measured cognitive abilities and skills are only one component. Definitions of intelligence are "value laden," given their focus on "concepts of appropriateness, competence, and potential" (Serpell, 2000, p. 549). Cultural intelligence includes skills that enable an individual to operate socially in multiple cultural contexts by effectively transferring the skills learned in one context to another (Brislin, Worthley, & Macnab, 2006).

Fagan and Holland (2006) investigated definitions of intelligence based on information processing. They theorized that an IQ score was a measure of an individual's knowledge based upon the person's information-processing ability and the information available to the individual in the cultural context. They question whether racial differences in intelligence were due to variations in individuals' cognitive abilities or to disparities in their exposure to information. The authors suggest that all individuals have not had equal opportunity of exposure to information presented in standardized tests of intelligence, and they find that observed racial differences in intelligence can be accounted for by differential access to information found in conventional IQ tests.

Heritability

One of the most heated debates about intelligence and race exists at the intersection of genetics, heritability, and culture. Heritability itself is an elusive construct and estimates of this construct "can vary in different populations or at different times" (Rushton & Jensen, 2005, p. 239): "Heritability describes what is the genetic contribution to individual differences in a particular population at a particular time, not what could be. If either the genetic or the environmental influences change (e.g., due to migration, greater educational opportunity, better nutrition), then the relative impact of genes and environment will change" (p. 239).

The heritability of intelligence is heavily influenced by the characteristics of a particular sample. For example, it is likely stronger in older populations, weaker in children, and potentially close to zero at the lowest socioeconomic status levels, where environmental impact is hugely unstable (Mackintosh, 2011; Nisbett et al., 2012).

Rushton and Jensen (2005) published a review of thirty years of research on racial differences in cognitive ability. After discussion of research underlying ten categories of evidence, they conclude that a "genetic component" exists underlying the differences between Blacks and Whites. Since this article's publication, a number of scholars have critiqued their conclusions that favor a hereditarian explanation that they identify as 50 percent genetic and 50 percent environmental. For example, Rushton and Jensen (2005) cite decades of research on high correlations of intelligence test scores between identical twins reared apart to support their hereditarian perspective. Nisbett (2009) provides a contrasting argument, noting that the high correlation among twins reared apart "reflects not just the fact that their genes are identical but also the fact that their environments are highly similar" (p. 26). Thus, it is unlikely that the similarity in scores is totally due to heredity.

Poverty, Home Environment, and Other Contextual Factors

Culture and environment are intimately linked as culture impacts the meaning assigned to the perception of one's environment. The culture of poverty produces a number of environmental (e.g., home environment) and contextual factors that have been related to lower intelligence.

Poverty generally connotes inadequate financial resources and purchasing power. However, poverty must also be viewed in the broader context of a social class framework of oppression, or classism, which serves to limit access to socially valued services and resources, such as education, health care, mainstream opportunities and experiences, and societal representation (Smith, 2010; Sue & Sue, 2016). Conditions of poverty also impact the way one views their environment and may create lived experiences of cultural invisibility and social exclusion (Sue & Sue, 2016).

The *home environment* can be impacted by poverty for multiple generations, resulting in what can be described as a culture of poverty. Valencia and Suzuki (2001) detail the ways in which socioeconomic status indirectly impacts performance on intelligence measures through a myriad of contextual factors. Parental expectations, access to quality education and other resources, and risk of neighborhood violence are all influenced by socioeconomic status and home environment. The authors note that socioeconomic status is linked to racial group membership as members of marginalized groups are often found in lower SES groups.

May, Azar, and Matthews (2018) studied the relationships between *neighborhood*-concentrated disadvantage, residential instability, and home physical and learning environments for pre-school-age children in low-income families. They hypothesized that living in a disadvantaged neighborhood is associated with heightened risk of poor school readiness and health outcomes in early childhood. Their findings partially corroborated this hypothesis, indicating that concentrated disadvantage (e.g., overall sanitation, furnishings, and safety in the home) was negatively associated with quality of home physical environments. Residential instability was not associated with the home learning environment through mothers' perceptions of neighborhood disorder and depressive symptoms. The effects of concentrated disadvantage were buffered by mothers' neighborhood social embeddedness.

Hadd and Rodgers (2017) conducted two studies focusing on intelligence, income, and *education* as potential influences on a child's home environment, using a maternal sibling-comparison design. The authors note that the creation of a child's home environment has primarily been explored via a focus on parental influences, and that only a few studies have considered that children participate in constructing their own environments. Their study reverses the usual direction of evaluation to consider the home environment as an outcome of child-maternal characteristics instead of a predictor of child outcomes. Thus, the researchers posed the question "How is a child's home environment constructed?" (p. 1286). They concluded that maternal characteristics, such as intelligence, are useful predictors but have diminished utility in within-family settings (e.g., as opposed to between-family settings). Child intelligence remained a meaningful predictor of the quality of home environments even within families, particularly in late childhood and early adolescence.

Chiu and DiMarco (2010) noted that homelessness and poverty can cause developmental delays in children, but they conclude that children with developmental delays who receive early interventions may benefit in critical areas of learning. They cited developmental screening as crucial for homeless children due to increased risks related to poverty and homelessness.

Herbers and colleagues (2014) studied resilience as related to the culture of poverty, home environment, and parental relationships. Participants were primary caregivers who were African American, multiracial, American Indian, White, and other races. They concluded that positive parent-child coregulation (relationship processes of guiding and responding to children's behavior) among a racially diverse sampling of families (including African American, American Indian, multiracial and White) supports competence and resilience in young children who experience high levels of adversity (e.g., homelessness). Positive parent-child coregulation was predictive of peer acceptance at school and was found to be related to executive functioning and intelligence. Additionally, Anderson (2018) highlighted the importance of differentiating levels of stress among impoverished Black families. Families with less conflictual parent-child relationships had more optimal school readiness than families with less financial strain but higher levels of conflict. Thus, children's school readiness may be impacted by familial resilience characterized by adaptive coping skills for counteracting intersectional challenges of race, poverty, and discrimination.

Educational engagement and academic achievement are often promoted as pathways to transcend poverty and prejudice among at-risk students. However, emerging evidence questions that important assumption. First, educational disparities in the US reveal that families who have the greatest need to escape poverty are often relegated to the least adequate educational resources (Sue & Sue, 2016). Second, number of years of education is known to be related to intelligence test performance. Yet the quality of education, even when controlling for amount of education, is a potentially pervasive source of cognitive test performance differences by race/ethnicity (Shuttleworth-Edwards, 2016).

Moreover, the school-to-prison pipeline (STPP), which refers to a path from the education system to the juvenile or adult criminal justice system, is another example of one of the consequences of inadequate educational resources for families living in poverty (McCarter, 2017). McCarter (2017) observed that the change in discipline in the public schools in America has contributed to disparate outcomes for students of color. Similarly, Morris (2016) outlines how an increasing number of girls of color come into contact with the criminal and juvenile justice system. In the past two decades, public schools have increased punitive responses to expressions of dissent among Black and/or Latina, lesbian, gay, bisexual, transgender, queer/questioning (LGBTQ), or gender non-conforming students, increased surveillance in homes, schools, and communities, and increased reliance on exclusionary discipline (e.g., suspensions and expulsions). This situation has led to increasing numbers becoming involved with the criminal and juvenile justice systems. Morris identifies this process as criminalized education.

Fluency in English may impact intelligence test scores, as familiarity with the dominant culture upon which the test is based impacts performance. Discrepancies continue to be noted between children with limited English proficiency and those students who have mastery of English (Puente & Puente, 2009). An inverse relationship between English language proficiency (ELP) and performance on cognitive measures that require higher levels of English language development and mainstream cultural knowledge has been identified (Sotelo-Dynega et al., 2013). Authors note the importance of accurately assessing an examinees' level of developmental language proficiency and cultural knowledge acquisition in order to determine whether scores on a cognitive measure accurately reflect the abilities of the individual being assessed. Lakin and Lai (2012) found that cognitive ability tests that assess multiple content domains (e.g., verbal, quantitative, nonverbal) provide more reliable information regarding the strengths and limitations of students who are English language–learners when compared to unidimensional reasoning measures.

Acculturation is often linked to contextual variables such as language proficiency and familiarity with a testing situation, which in turn influence performance on intelligence tests (Mpofu & Ortiz, 2009). Acculturation is a "dynamic process of change and adaptation that individuals undergo as a result of contact with members of different cultures" (Rivera, 2008, p. 76). Acculturative stress is characterized by stressors related to the process of acculturation involving cultural, social, and psychological variables contextualized between the host and the immigrant or refugee (Short et al., 2010). Alberg and Castro-Olivo (2014) concluded that acculturative stress and symptoms of internalizing mental health problems had a significant inverse association with Latino students' academic performance; acculturative stress was found to be related to symptoms of depression, anxiety, and poor academic behaviors. The relationship between acculturation and achievement is complex given what has come to be called the immigrant paradox (Garcia Coll & Marks, 2012), indicating that the developmental and educational achievement of assimilated children of immigrants may be diminished in comparison to those who are not assimilated.

Testing Constructs in Relation to Culture and Intelligence

Measuring intelligence via standardized assessment is challenging and subject to certain assumptions and error, especially among individuals who do not fit standardized sampling characteristics (Bowden, Saklofske, & Weiss, 2011; Whitaker, 2015). While "test use is universal" (Oakland, 2009, p. 2), most test development occurs "in countries that emphasize individualism and favor meritocracy (i.e., the belief that persons should be rewarded based upon their accomplishments) rather than collectivism and egalitarianism (i.e., the belief that all people are equal and should have equal access to resources and opportunities)" (Oakland, 2009, p. 4). In addition, "conventional tests are not designed to measure adaptation to any environment, but rather, to environments in which Western schooling dominate" (Sternberg, 2017, p. 23). In other words, people coming from cultures where achievement on standardized tests is not a valued or prioritized method of assessment may not perform as well on these measures.

The American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME) endorsed a new set of standards for educational and psychological testing (2014), citing three foundations of testing – the traditional psychometric properties of validity and reliability, and a third pillar which impacts the use of intelligence tests with culturally diverse groups, that is, *fairness in testing*. Fairness emphasizes the importance of context in relation to testing, noting threats to validity that include limited English proficiency and differences due to educational and cultural background. The *Standards* state that there are threats to fair and valid interpretation of test scores, including test content, test context, test response, and opportunity to learn. Specific attention is noted with respect to subgroup mean differences that "do not in and of themselves indicate a lack of fairness, but such differences should trigger follow-up studies, where feasible to identify potential causes of such differences" (AERA, APA, & NCME, 2014, p. 65).

g-factor. In 1927, Spearman hypothesized that intelligence consists of a general factor (*g*) and specific ability factors. His work in the development of factor analysis led to the operationalization of *g* as the first unrotated factor of an orthogonal factor analysis. Tests with high *g* loadings included those focusing on "reasoning, comprehension, deductive operations, eduction of relations (determining the relationship between or among two or more ideas), eduction of correlates (finding a second idea associated with a previously stated one), and hypothesis-testing tasks" (Valencia & Suzuki, 2001, p. 31). In contrast, tests with low *g* loadings are those that focus on visual-motor ability, speed, recognition, and recall. Tests with high *g* loading indicate greater cognitive complexity.

Some researchers have made note that the Black/White IQ test performance gap is greater for subtests and items with higher *g* loadings and weaker for those with smaller *g* loadings (Rushton & Jensen, 2005). Rushton (2012) applies these findings to support a genetic hypothesis on the origin of the racial IQ gap; he argues that a genetic versus an environmental hypothesis better explains this pattern of larger differences for more heritable, more cognitively complex items. A meta-analysis comparing Whites with Amerindians noted that group differences were strongly related to general intelligence (*g*), particularly on verbal measures (te Nijenhuis, van den Hoek, & Armstrong, 2015). Other scholars have refuted this argument, noting how environmental factors aptly explain the pattern of the IQ racial gap. They argue that environmental disadvantage likely impacts performance on cognitively complex tasks (Nisbett et al., 2012). Despite this controversy surrounding the general intelligence factor, most intelligence tests provide an indicator of the level of *g* associated with various subtests that comprise the measure.

Test bias. Test bias often refers to the existence of systematic error in the measurement of a construct or variable, in this case, intelligence: "The discussion of bias in psychological testing as a scientific issue should concern only the statistical meaning: whether or not there is systematic error in the measurement of a psychological attribute as a function of membership in one or another cultural or racial subgroup" (Reynolds, 1982a,1982b, cited in Reynolds & Lowe, 2009, p. 333).

Reynolds and Lowe (2009) report the following as possible sources of test bias: inappropriate content, inappropriate standardization samples, examiner bias, language bias, inequitable social consequences, measurement of different constructs, differential predictive validity, and qualitatively distinct minority and majority aptitude and personality. Serpell (2000) cites work distinguishing among various forms of bias, including outcome bias, predictive bias, and sampling bias. Predictive bias focuses on intelligence tests as they predict "future performance in educational settings" (Serpell, 2000, p. 563). For example, in a meta-analysis drawing samples from 1960 to 2010, the correlation between cognitive ability test scores and performance (in educational admissions, civilian employment, and military settings) was significantly greater for Whites ($r = 0.34$) than it was for Blacks ($r = 0.24$) or Hispanics ($r = 0.30$) (Berry, Clark & McClure, 2011), leading researchers to conclude that cognitive test scores may differentially predict future performance in minorities and Whites. Sampling bias occurs when a standardized test of intelligence is "biased in favor of a range of skills, styles, and attitudes valued by the majority culture and promoted within the developmental niche that it informs" (Serpell, 2000, p. 563). Helms (2004) cites problems with existing definitions of test bias: "evidence of test-score validity and lack of bias, as those terms are currently construed in the literature, does not mean that test scores are fair for African American test takers and other people of color" (p. 481). Ford and Helms (2012) question the use of conventional test scores to determine African Americans' true intelligence in response to the fact that "test bias and unfairness abound" (p. 186), citing the underrepresentation of African Americans in the actual process of test development and the fact that many traditional IQ tests were originally structured for a White, middle- to upper-class, monolingual, American test-taker.

Valencia, Suzuki, and Salinas (2001) note, "Test bias in the context of race/ethnicity often is referred to as cultural bias" (p. 115). In a review of sixty-two empirical studies of cultural bias with cognitive ability tests, the majority (71%) detected no significant evidence of bias, while the remainder (29%) indicated bias or mixed findings (Valencia et al., 2001). It appears that the findings on test bias with respect to cognitive ability testing remain inconclusive.

In order to address the potential of cultural (i.e., race/ethnicity) bias, most state-of-the-art intelligence tests are standardized based upon representative census data with respect to gender, race and/or ethnicity, region of the country, urban or rural status, parental occupation, socioeconomic status, and educational level (Valencia & Suzuki, 2001). In addition, test developers employ expert reviewers to examine item content and statisticians to perform analyses to determine differential item functioning. Numerous reliability (e.g., split-half, test-retest, internal consistency) and validity studies (e.g., factor-analytic studies, external validity) are often conducted and may employ the Rasch model of item response theory to assess the fit of subtest items to the ability area being assessed. Some test developers also engage in racial and ethnic oversampling to address potential test-bias issues.

One study finds that that African-American and White preschool children did not differ on overall cognitive ability as measured by the Stanford-Binet Intelligence Scale – Fifth Edition (SB5) (Dale et al., 2014). Tests for parallelism were also nonsignificant, indicating that patterns displayed similar highs and lows for both groups. Additionally, researchers displayed the invariance of factor structure of the Kaufman Assessment Battery for Children and Kaufman Test of Educational Achievement – Second Edition based on the Cattell-Horn-Carroll model of broad abilities for Black, Hispanic, and White schoolchildren (Scheiber, 2016).

Cultural loading refers to the degree of cultural specificity contained within a particular measure. All tests are culturally loaded, as their content and format reflect what is important in the cultural context of the community for which it was developed. Cultural loading has important implications for understanding cultural bias: "For an intelligence test to be deemed culturally biased, it must be culturally loaded. A culturally loaded test does not, however, necessarily mean that such a test is culturally biased. In other words, cultural loading on an intelligence test is a necessary, but not a sufficient, condition for the existence of cultural bias" (Valencia et al., 2001, p. 114).

The cultural loading of a particular test can favor certain groups over others. If there is a match or "congruence" between tasks required on an intelligence test and the cultural background of the individual test-taker, then the individual may possess an advantage. If there is "little or no congruence" between the content of the test and cultural background of the test-taker, then the test-taker may be at a disadvantage (Valencia et al., 2001, p. 114). Given that all forms of measurement are developed within a cultural context, it is difficult to ascertain a fundamental cognitive task that would not be impacted by cultural loading.

Cultural equivalence , that is, whether "interpretations of psychological measurements, assessments, and observations are similar if not equal across different ethnocultural populations" (Trimble, 2010, p. 316), is difficult to address with regard to intelligence testing. For example, there are issues related to contextual equivalence (e.g., whether intelligence can be measured in different contexts in which examinees reside) and conceptual equivalence (e.g., does the content of the test, e.g., stimuli materials, have the same meaning for examinees from different cultural backgrounds?).

Cultural equivalence, cultural bias, test fairness, and the impact of individual-difference variables and their relationship to the racial/ethnic group ordering of intelligence test scores have been a focus of the literature in the past three decades (Berry et al., 2011; Helms, 1992, 2004, 2006). The racial/ethnic hierarchy of intelligence refers to the ordering of different minority groups based upon their average intelligence test score. As noted earlier, test bias refers to systematic error in the measurement of intelligence for a particular group. Helms (2006) provides input into the complexity of addressing error that may be due to factors unrelated to intelligence (e.g., internalized racial or cultural experiences and environmental socialization). She hypothesizes that these factors may have a greater impact on the test performance of members of racial and ethnic minority groups relative to nonminority group members. Given the ubiquity of conventional intelligence testing, unpacking the differential impact of these factors is paramount (Shuttleworth-Edwards, 2016). For example, the underrepresentation of African Americans in gifted educational programs is directly linked to cognitive scores (Ford, Grantham, & Whiting, 2008).

Culture and Intelligence: Neuroscience Implications

Researchers have also looked to the neurosciences to explain racial and ethnic differences in cognitive assessment. Parkinson and Wheatley (2015) propose a process deemed "cultural repurposing" to explain how culturally invented ways of thinking and perceiving map onto pre-existing brain structures. Within the lifetime of an individual, neural plasticity can alter brain circuitry in response to adaptive demands. For example, we see this in the differing neurocognitive networks mediating the use of English and Chinese language (Chan et al., 2002). Chan and colleagues hypothesized that speaking and thinking in Chinese involved more bilateral brain areas than did speaking and thinking in English, which were more lateralized to the left brain hemisphere. This hypothesis suggests that early language experiences can influence how the brain processes information. Language structure can lead to cultural variations in performance on basic cognitive tasks (Cheung & Kemper, 1993; Chincotta & Underwood, 1997; Hedden et al., 2002).

Hwa-Froelich and Matsuo (2005) examined how quickly bilingual (Vietnamese-English) preschool children were able to "fast map," or learn the meaning of a new word by associating it with a sound or image, after hearing the word. They found that regardless of exposure to English or Vietnamese, children were more likely to produce sound patterns that were more familiar to them, even when the stimuli presented to them were new. This finding emphasizes the importance of cultural exposure to words and images in determining learning style and cognitive performance among new immigrants. Additionally, recent findings demonstrate that language interacts with cultural identity during semantic processing. Ellis and colleagues (2015) note that among a fluent, bilingual Welsh sample, culturally relevant information about Wales was integrated with more ease when it was presented in Welsh rather than English; however, when the information was not culturally relevant, there was no difference in processing ease as a function of language. Such findings emphasize the potential impact of acculturation and language status on cognitive test-taking.

Alternative Assessment Practices

A number of alternative assessment practices have emerged in recent years in part to address criticisms of the usage of intelligence tests with members of racial and ethnic minority groups. These assessments address concerns related to the limited impact of intelligence testing on actual instruction and intervention. We provide a brief discussion of the major areas and types of assessment that are currently used.

Nonverbal intelligence tests. Originally, nonverbal measures were believed to reduce the requirement of verbal instruction, feedback, and responses, thus minimizing "the impact of culturally based linguistic differences on assessment results and outcomes" (Harris, Reynolds, & Koegel, 1996, p. 223). Instructions were communicated through gestures and behavioral demonstrations (e.g., modeling). These nonverbal procedures have been used with students who have limited English reading and writing skills (i.e., English language–learners) in the United States. These tests, however, have been cited as "culturally reduced measures" as all tests involve some form of language and communication. Therefore, nonverbal tests "are not entirely devoid of cultural content" (Mpofu & Ortiz, 2009, p. 65). Nonverbal tests also assess a more limited range of ability areas, including "visual processing, short-term memory, and processing speed" (p. 65). Studies of the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV) and Wechsler Intelligence Scale for Children – Fifth Edition (WISC-V) indicate that consideration must be given to the linguistic demands of the oral directions on nonverbal measures. For example, Cormier, Wang, and Kennedy (2016) found that oral directions given for nonverbal subtests including Block Design, Letter-Number Sequencing, Cancellation, Picture Span, Visual Puzzles, and Figure Weights contain relatively high linguistic demands.

Dynamic assessment. A number of dynamic assessment procedures have been developed to provide more relevant data about students to inform educational planning. Dynamic assessment is an active form of informal assessment and often involves the examiner engaging in a test-teach-test procedure (Meller & Ohr, 1996). Dynamic assessment is characterized by "active interaction on the part of the evaluator to obtain an estimate of the child's current level of functioning, and then the use of non-standardized procedures to probe for the nascent cognitive potentials not revealed in the previous standardized assessments" (Armour-Thomas & Suzuki, 2016, p. 58).

The focus of the assessment is on the process. Dynamic assessment enables evaluators to observe the processes of learning for an individual as they provide feedback to the examinee to improve performance. This is an important assessment tool, as it provides opportunities for an individual to demonstrate learning of material that they may not have been exposed to in the past (Sternberg, 2004). The focus on process has implications for culturally diverse individuals, as they are provided with feedback and the opportunity to demonstrate learning.

Performance-based, authentic, and curriculum-based assessment (CBA) measures were designed to address concerns regarding norm-based measures like intelligence tests (Hintze, 2009) in response to claims that "published tests have played too large a role in educational and psychological decision making, not just with students from diverse backgrounds" (Shinn & Baker, 1996, p. 186). These assessment methods require students to solve "real-world tasks" incorporating what they have learned. These methods often include attention to the curriculum to which the individual has been exposed (e.g., informal reading inventories; Shinn & Baker, 1996). CBA examines behavior in a natural context, focuses on what is being taught in the classroom, leads to purposeful interventions in the classroom, and is useful in formative and idiographic (i.e., within-student) evaluation of progress (Hintze, 2009). CBA has been used in screening, determination of eligibility for special education, goal setting, program evaluation, and development of interventions (Hintze, 2009).

Response to intervention. Response to intervention (RTI) "is a data-based process to establish, implement, and evaluate interventions that are designed to improve human services outcomes" (Reschly & Bergstrom, 2009, p. 434). RTI involves a series of tiered interventions taking into consideration the prior knowledge of the individual learner. RTI includes attention to empirically validated instructional and behavioral programs and interventions (Reschly & Bergstrom, 2009). This approach has the potential of eliminating the use of tests that have been accused of being biased against particular racial and ethnic groups.

Think-aloud protocols (TAPs; van Someren, Barnard, & Sandberg, 1994). TAPs are designed to gain insight into an examinee's problem-solving process as they verbally narrate their thinking and as they construct a response to a given problem or task. This process enables the data to be obtained more directly from the examinee beyond just scoring a series of final answers. The process often involves the gathering of TAPs and then conducting a qualitative analysis to create a model of the cognitive processes used in problem-solving for the individual being assessed.

The G*f*-G*c* cross-battery assessment model (XBA; Flanagan, Ortiz, & Alfonso, 2007). The XBA is a method of intelligence assessment that enables evaluators to measure a wider range of cognitive abilities by selecting from a range of potential tests assessing broad and narrow ability areas (McGrew & Flanagan, 1998). Information regarding the cultural content, loading (e.g., cultural specificity), and linguistic demands (e.g., verbal vs. nonverbal, receptive language, expressive language requirements) of various measures are provided in the culture-language test classifications (C-LTC) and culture-language interpretive matrix (C-LIM). Classifications are based upon empirical data available on the particular test and expert consensus procedures in the absence of data. The matrix serves to assist clinicians in interpreting test score patterns. The C-LTC and the C-LIM were created to guide test selection and interpretation by taking into consideration the potential impact of acculturation and language proficiency (Ortiz & Ochoa, 2005). Unfortunately, the diagnostic utility of the C-LIM has been challenged, as data has not supported its usage in making decisions about services for students who are English language–learners (e.g., Styck & Watkins, 2013).

The Multidimensional Assessment Model for Bilingual Individuals (MAMBI; Ortiz & Ochoa, 2005). The MAMBI takes into consideration the unique features of each testing case based upon the designated referral question. The evaluator must make decisions regarding the methods and approaches to be used to assess the student to obtain the most relevant and accurate information from a collection of data sources guided by a systemic framework based upon the individual's cultural and linguistic history (Rhodes, Ochoa, & Ortiz, 2005).The MAMBI integrates language (e.g., preproduction, early production, speech emergence, intermediate fluency, and development of cognitive academic language proficiency) and instructional programming/history and considers how types of bilingual instruction impact cognitive and linguistic development and current grade level (with the assumption that the level of formal schooling impacts language development) to determine the most appropriate assessment (e.g., nonverbal assessment, assessment primarily in the native language, assessment primarily in English, and bilingual assessment).

Outcome Implications for Multicultural Populations

A number of controversies surround the use of intelligence measures centering on findings of a racial and ethnic group hierarchy of scores. Overall estimates of group scores based upon a mean of 100 and standard deviation of 15 have been reported as follows: Whites 100; Blacks (African Americans) 85; Hispanics, midway between Whites and Blacks; and Asians and Jews, above 100 (*Wall Street Journal*, 1994). Research indicates that American Indians score at approximately 90 (McShane, 1980). Lynn (2015) reports average IQ scores by race and location/country based upon his review of empirical literature. The following IQ scores for racial groups in the United States/North America are noted as follows: sub-Saharan Africans 85; Southeast Asians 93; Native Americans 86; Hispanics 89, Arctic Peoples 91; and Northeast Asians 101. Lynn acknowledges potential inaccuracies due to sampling and measurement error but overall the ordering of racial and ethnic groups by average intelligence test scores appears to be commensurate with earlier findings.

Despite these overall group differences, it is important to remember that within-group variability exceeds between-group variability on these psychological and educational measures (Reynolds & Lowe, 2009). Thus, intelligence tests like the Wechsler scales may provide psychometric information (e.g., subtest score means by racial/ethnic group) based upon racial/ethnic oversampling, but racial or ethnic group norms are not provided.

Tests as Gatekeepers

Despite the growing number of alternatives readily available to substitute for traditional intelligence tests, these measures continue to play a role in educational placement. In particular, intelligence tests play a critical role in admission to services like special education and giftedness programs).

Weiss and colleagues (2006) suggest that scores reflect societal differences tied to the current practices in test development – that is, stratified norming taking into consideration age, gender, region of the country, parental education, and socio-economic status. The authors note that the "sampling methodology accurately reflects each population as it exists in society, [but] it exaggerates the differences between the mean IQ of groups because the SES levels of the various racial/ethnic samples are not equal" (p. 31). If test developers equated the percentages for all groups, then the discrepancies between the groups would be minimized but not eliminated. Thus, SES level accounts only partially for group differences. Other variables may also play a role, for example, home environment factors, noted earlier in this chapter, which may differ even within comparable SES levels.

In addition to examining the impact of these stratification variables, Weiss and colleagues (2006) also reported that parental expectations were assessed by asking parents how likely they believed their child would be to get good grades, graduate from high school, attend college, and graduate from college. Interestingly, parental expectations accounted for approximately 31 percent of the variance in Full Scale IQ (FSIQ). Thus, the researchers conclude that parental expectations account for more variance than parental education and income combined.

What is most salient about this ordering is that it reflects the sociocultural contexts for particular racial and ethnic minority groups in the United States, and these scores have significant implications. Intelligence tests are used to determine eligibility for special services and classifications of learning disabilities and other intellectual impairments.

Black–White Intelligence Test Score Gap

"Differences between African Americans and Whites on IQ measures in the United States have received extensive investigation over the past 100 years" (Reynolds & Lowe, 2009, p. 333). Nisbett and colleagues (2012) noted that the IQ difference between Blacks and Whites had been reduced by 0.33 standard deviations (SD) from 1975 to 2008. In addition, they reported an average Black/White gap reduction of 0.57 SD for reading and 0.3 SD for math, averaging out to an IQ gain equivalent of 6.45 points. Dickens and Flynn (2006) similarly identified a 5.5 IQ–point gain in Black samples studied between 1972 and 2002.

In addition, when socioeconomic status is taken into account, the differences between groups are reduced. For example, the mean difference between Blacks and Whites in the United States drops from 1 standard deviation to 0.5 to 0.7 standard deviations when SES is held constant (Reynolds & Lowe, 2009). The relationship between the Black/White gap and SES gap has undergone dramatic change; fifty years ago the Black/White gap was significantly larger than the SES gap, and the opposite has been found in recent years (Reardon, 2011). Despite the lowered discrepancy between Black and White children on this standardized IQ test, and an understanding of the role of SES, researchers, scholars, and other professionals continue to struggle with the complexities inherent in the understanding of intelligence and racial difference.

Historically, the discussion of intelligence among Black/African-American populations has been ongoing in both educational and academic research environments. Franklin (2007) reviewed publications appearing in the *Journal of Negro Education* (*JNE*) since 1932 focusing on the intelligence testing of African Americans. He notes that social scientists contributing to the *JNE* have, for many decades, attempted to identify and clarify what the tests were measuring and to emphasize the culturally biased processes involved in the standardization of these measures (i.e., favoring White middle-class populations). The *JNE* "participated in laying the educational and legal ground work for the U.S. Supreme Court's *Brown v. Board of Education* decision" in 1954 and published literature concerning the impact of the Brown decision throughout the 1950s and 1960s (p. 224). Additionally, in the late 1960s, the Association of Black Psychologists (ABPsi) submitted a petition of concerns to the American Psychological Association calling for a "moratorium of testing of all Black children until appropriate and culturally sensitive tests were developed" (Franklin, 2007, p. 224). These calls for better assessment measures for African Americans also came in response to research that was conducted in the late 1960s and early 1970s by Jensen (1969), in which he focused on the heritability of intelligence.

Stereotype threat. Steele and Aronson's (1995) seminal article about the effect of stereotype threat on the test-taking performance of African-American students included a series of four experiments that revealed depressed standardized test performance among African-American participants relative to White participants, when the African-American students were made vulnerable to judgment by negative stereotypes. Stereotype threat has been defined as a phenomenon that occurs when an individual recognizes that negative stereotypes about a group to which they belong are applicable to themselves in a particular context or situation (Steele, 1998). When conditions were designed to alleviate stereotype threat, African-American participants' test performance improved. Steele and Aronson (1995) concluded that although stereotype threat was not the sole explanation for the gap in scores, it did appear to cause an "inefficiency of processing much like that caused by other evaluative pressures" among the African-American participants (Steele & Aronson, 1995, p. 809).

In the past twenty-four years since the publication of the Steele and Aronson (1995) article, there has been much debate about stereotype threat as an explanation for the Black-White test score gap. Critical analyses of the research conducted by Steele and Aronson (1995) have included concerns about internal validity of empirical studies of stereotype threat, specifically, perceptions of face validity and test-taking motivation among African-American participants (Whaley, 1998). Additional criticisms of the study identified alleged "misinterpretation of research" and questioned the generalizability of stereotype threat in applied testing sessions (Sackett, Hardison, & Cullen, 2004, p. 11). Relationships between stereotype threat and gender have also been explored (e.g., Spencer, Steele, & Quinn, 1999) and greater specificity in the construct has been identified in terms of stereotype-specific threat (e.g., threat that results directly from the testing environment) and stereotype-general threat (e.g., based on a global sense of threat that is pervasive in a variety of contexts/situations) (Mayer & Hanges, 2003). A number of studies have been conducted to address the level of contribution of stereotype threat to the test score gap (e.g., Brown & Day, 2006; Cohen & Sherman, 2005; Helms, 2005; Steele, 1998; Steele & Aronson, 2004). Borman and Pyne (2016) conclude that stereotype threat may account for as much as "one-quarter of the black-white achievement gap and that interventions to buffer students from the harm of stereotype threat can help close that fraction of the gap" (p. 181). Their review suggests that the effects of stereotype threat on the Black-White test score gap are associated with lower positive self-identity among Black students when they are placed in majority-White classrooms.

Racial identity. Helms's (1995) racial identity theory posits identity statuses, some of which are characterized by self-denial and others by self-affirmation regarding one's socioracial group. Each racial identity status is related to distinct affects, behaviors, and cognitions concerning one's understanding of race and racism. These statuses comprise individual-difference variables that have been linked to Black student performance on cognitive ability tests (Helms, 2002, 2004). Data indicate that higher levels of Black idealization (i.e., idealization of an individual's Blackness and Black culture; emic culture-specific focus) were associated with lower SAT scores, and higher SAT scores were related to lower levels of Black idealization (Helms, 2002).

Higher Intelligence Test Scores for Asians

Asians and Asian Americans have often obtained the highest group averages on standardized intelligence tests, with high scores reported in particular on subtests measuring numerical and spatial reasoning abilities (Suzuki, Mogami, & Kim, 2002). What accounts for this difference has been the focus of speculation for decades. Some believe that the higher scores are due to perseverance and not to innate intellectual aptitude. As Nisbett (2009) writes, "What is not in dispute is that Asian Americans achieve at a level far in excess of what their measured IQ suggests they would be likely to attain. Asian intellectual accomplishment is due more to sweat than to exceptional gray matter" (p. 154).

In a related vein, the "model minority myth" portrays Asian students as being, on average, more perfectionistic, self-controlled, cooperative, academically successful, and with fewer behavioral problems than other students (e.g., Chang & Sue, 2003; Loo & Rappaport, 1998). Chang and Demyan (2007) examined the content of teachers' race-related stereotypes. Their findings indicated that teachers were more likely to note Asian students to be more industrious and intelligent, and less athletic and sociable compared with African and European American students. The authors note that the implications of these findings are that real learning needs, such as weaknesses in math or science, may be overlooked.

Studies on the intelligence of Asian Americans note that there has been little published data on the reliability and validity of the most frequently used intelligence measures (e.g., the Wechsler scales) with Asian samples in the United States alone (Okazaki & Sue, 2000). Most of the published studies in the past decade have focusing on non-US Asian samples (e.g., Chinese internationals), and Nisbett and colleagues (2012) state that differences between Asian and Western samples have often not been a focus of study. Okazaki and Sue hypothesize that Asians were not often included in studies to standardize cognitive or personality assessments because of a lack of clinicians proficient in the native language of the particular Asian group, difficulties locating Asian participants (who may be more geographically scattered), and difficulties in recruitment as a result of Asian cultural attitudes toward testing. Asian Americans may be less likely to seek testing because of potential stigma associated with learning disabilities (Okazaki & Sue, 2000; Uba, 1994), especially in contrast to a community emphasis on achievement.

Research is being conducted to address measures of intelligence with particular Asian ethnic groups. One such study examined the performance of Hmong American students aged five to fourteen years on the Kaufman Assessment Battery for Children – Second Edition (KABC-II) and WISC-V (Romstad & Xiong, 2017). Findings indicate that Hmong students scored one SD below the national mean on both measures. The study also challenged the use of the C-LIM as potentially misrepresenting the intellectual and processing abilities of Hmong students.

Major intelligence tests like the Wechsler scales have been exported to other Asian countries, normed, and restandardized. The WAIS has been translated and standardized in China, Hong Kong, India, Japan, Korea, Taiwan, Thailand, and Vietnam (Cheung, Leong, & Ben-Porath, 2003). The restandardized norms developed in an Asian country may not be applicable to someone living in the United States. First, norms may be outdated. Second, US immigrant Asian groups are more heterogeneous relative to their overseas counterparts (Okazaki & Sue, 2000). Chinese immigrants in the United States, for example, may represent diverse population clusters from China and speak different dialects compared with a sample of Chinese individuals living in Hong Kong. Therefore, applying norms based on one Asian ethnic group to interpret the test results of an individual from a different ethnic group may be misleading. Yet another source of heterogeneity among US immigrants is that they are exposed to American values and are considered a minority group in the United States (Okazaki & Sue, 2000). Future research should compare the validity of overseas Asian norms to Asian Americans and vice versa, to determine whether US Asians need to be normed as a separate stand-alone sample.

Intelligence from an American Indian/Native American Perspective

There appear to be limited studies currently focusing on the application of intelligence tests with American Indian/Native American communities. A preliminary meta-analysis of Wechsler studies conducted between 1986 and 2003 on American Indian cognitive abilities yielded a total of sixty-three empirical studies that sampled a range of tribal groups (Suzuki et al., 2003). The most frequently cited groups were Navajo, Papago, Ojibwa, Inuit, and Eskimo. A consistent finding across studies was that American Indian samples scored relatively higher on nonverbal spatial reasoning measures with specific strengths noted on Object Assembly and Block Design in comparison to lower scores on Vocabulary and Information subtests. The average standard score difference between Verbal IQ and Performance IQ was 17 points (SD 8.92), range 3.4–31.3. Interpretation of these findings often focuses on the Verbal IQ as being lower due to linguistic and cultural factors rather than intelligence, with attention to the Performance IQ as more indicative of an individual's true abilities. It should be noted that contextual variables were often not reported (e.g., reservation and referral status). In addition, important demographic and health information was often not provided (e.g., socioeconomic status, presence of ear infections, primary language spoken in the home). A more recent meta-analysis of studies comparing *g* loadings of Amerindians and Whites found similar profile differences (i.e., higher performance vs. verbal scores) with data supporting that the IQ gap was largely based upon *g* (te Nijenhuis et al., 2015).

Test bias was examined on the Bayley Scales of Infant Development and the WISC-III, with findings indicating that the performance of American Indian students may be impacted by "poverty, remoteness, access to resources, and health care" (Hagie, Gallipo, & Svien, 2003, p. 15). Hagie and colleagues note that these widely used tests "in most areas fail to reflect the local and cultural experiences of American Indian students and, subsequently, present a skewed picture of their true ability and performance" (p. 23). Many American Indian children learn problem-solving through collaborative effort that is not represented in traditional testing practices.

Studies have explored the impact of including naturalistic cultural tools in gaining an understanding of intelligence among Zinacantec Maya children (Greenfield, Maynard, & Childs, 2003; Maynard, Subrahmanyam, & Greenfield, 2005). The studies took place during a historical transition from a subsistence and agricultural economy to one based on money and commerce. The findings highlight a change from interdependent to independent weaving practices. The researchers observed that cultural learning, abstract representation, and innovation were linked to an emphasis on independence away from "scaffolded guidance, detail-oriented representation, and imitative representational strategies" (Greenfield et al., 2003, p. 455). They conclude that "patterns of cognitive development of a new generation change in response to a changing world, but always respecting constant patterns of cognitive development" (p. 485).

Estimates of Hispanic and Latino/a Intelligence in Context

Obtaining an accurate reading of the intelligence of Hispanics/Latino/as involves a number of challenges. This diverse group is notably the "fastest growing and possibly the most disenfranchised group in the United States today" (Puente & Puente, 2009, p. 418). One must attend to issues of limited educational opportunities, low socioeconomic status, and language. A number of diverse subgroups comprise the category of Hispanics. Each group has different histories of immigration and cultural traditions. There is growing emphasis on the need to examine ethnic group differences instead of grouping individuals under one "Hispanic" label. Nevertheless, a large percentage of Hispanics have limited English proficiency and have not fared well in the American educational system (Puente & Puente, 2009).

Puente and Puente (2009) also note that most tests that are published in the United States do not have Spanish translations. Tests that have been translated into Spanish are often not normed on American samples but rather on samples from Spanish-speaking countries abroad. This is an issue because "subcultures of Hispanic heritage may be as dissimilar with each other as they are to the U.S. culture" (Puente & Puente, 2009, p. 424). The complexities of translation are also evident, given issues of equivalence; linguistic or language equivalence does not ensure cognitive equivalence, which focuses on meaning.

Fenollar-Cortes and Watkins (2018) highlight the complexities of these issues in their construct validity study of the WISC-V Spain, a revised and adapted version of the US WISC-V. The measure went through major revision including the addition of new subtests and changes to the content and instruction of all subtests. Particular items were deleted and new items were added to several subtests. Picture Concepts was omitted. The WISC-V Spain was standardized on a sample of 1,008 children aged six to sixteen years, stratified by age, sex, parent educational level, and geographic region and type (rural, suburban, urban). Findings indicate that "45% of the total variance of WISC-V Spain scores was due to error and specific variance and none of the group factor scores was sufficiently reliable for confident interpretation" (p. 10). The authors note that clinicians can be "reasonably confident in using the WISC-V Spain Full Scale IQ for clinical decisions" but concerns are noted regarding the reliable interpretation of factor index scores.

In comparison, the published WISC-V Spanish (Pearson, 2017) is normed on Spanish-speaking children who have attended school in the United States for less than five consecutive years. The normative sample of 2,200 children aged 6–16 was stratified to match US census data. Test items were validated to address potential cultural bias across different regions of origin. Instructions indicate that the test is to be administered in Spanish and children can respond in either Spanish or English. The average WISC-V Full Scale Index Quotient for the Hispanic sample was 94.4 with index scores ranging from 94.2 (Verbal Comprehension Index) to 98.3 (Processing Speed Index) (Weiss, Munoz, & Prifitera, 2015). Children with more-educated parents obtained higher test scores (FSIQ 105) compared to those with the least-educated parents (FSIQ 88). Approximately 18.8 percent of the variance in

Hispanic/White FSIQ score differences was explained by parental education level in combination with income.

Among Hispanics, most samples show a consistent discrepancy between lower verbal (Verbal Comprehension Index, Fluid Reasoning Index, Working Memory Index) and relatively higher performance (Visual Spatial Index and Processing Speed Index) abilities on the Wechsler scales (Weiss et al., 2015). While some have attributed this difference to the importance of language, Weiss, Munoz, and Prifitera note that discrepancies should be interpreted in the context of the child's overall ability, language dominance, parents' level of education, and years of education in English-speaking schools. They also note that a Verbal Comprehension Index that is 11 points below the student's own mean should be "considered unusual," because it was obtained by only "15% of the Hispanic sample" (p. 231).

With the growing numbers of Hispanic individuals in the United States, particularly school-age children and adolescents, the need to develop adequate instruments to address their cognitive skills is imperative. The task is not an easy one because of the linguistic and cultural complexities of this population.

Conclusions

Understanding intelligence through a multicultural lens is an arduous task. As presented in this chapter, difficulties of interpretation and operationalization of relevant constructs, complexities of environmental context (e.g., home and community), and availability of instruments and methods of assessment are only a few of the challenges.

In terms of environmental factors, the importance of parental expectations, support of academic pursuits in the home, higher socioeconomic status, and familiarity with testing procedures are just some of the variables that impact the measurement of intelligence. The evaluator is presented with a menu of potential instruments with which to assess cognitive functioning, some based upon relatively newer theories (e.g., information processing). In addition, a number of approaches to assessment have evolved focusing on the integration of cultural variables in the assessment process to minimize barriers to accurate assessments such as cultural loading and linguistic demands, and methods to guide the selection of the most appropriate intelligence tests.

Despite the availability of these alternative assessment practices, the concerns that have challenged the intelligence literature remain, such as the reasons why the racial/ethnic ordering of intelligence test scores has remained the same over time, though there is evidence that the discrepancies are growing smaller; Nisbett et al., 2012; Weiss et al., 2006). Given the major role that intelligence tests continue to play in the determination of special services in educational settings and diagnosis, there remains a need for continuing attention to the use of these measures with individuals from diverse cultural communities.

Intelligence test scores do little to inform instructional intervention and curriculum (Armour-Thomas & Suzuki, 2016). Thus, a number of alternative assessment procedures have been developed (e.g., dynamic assessment, curriculum-based assessment, response to intervention), moving the focus away from cultural bias and test fairness to learning curriculum models (e.g., test-teach-test). Despite their appearance on the assessment scene, instructional interventions have not been able to unseat the usage of intelligence tests in evaluation.

The stronghold of intelligence testing has persisted despite decades of blistering criticisms from members of marginalized and oppressed racial and ethnic communities. Indeed, the most popular tests have been transported, renormed, and restandardized globally. Navigating test administration and interpretation, as well as determining the appropriate role of intelligence tests in a myriad of cultural contexts, remains a challenging and controversial task.

References

Alberg, L. J., & Castro-Olivo, S. M. (2014). The relationship between mental health, acculturative stress, and academic performance in a Latino middle school sample. *Contemporary School Psychology*, 18, 178–186.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington: AERA.

American Psychological Association. (2017). Multicultural guidelines: An ecological approach to context, identity, and intersectionality. www.apa.org/about/policy/multicultural-guidelines

Anderson, R. E. (2018). And still WE rise: Parent-child relationships, resilience, and school readiness in low income urban black families. *Journal of Family Psychology*, *32*(1), 60–70.

Armour-Thomas, E., & Suzuki, L. (2016). Measuring and understanding intellectual and other abilities. In F. T. L. Leung, D. Bartram, F. M. Cheung, K. F. Geisinger, & D. Iliescu (Eds.), *The ITC International handbook of testing and assessment* (pp. 49–62). New York: Oxford University Press.

Berry, C. M., Clark, M. A., & McClure, T. K. (2011). Racial/ethnic differences in the criterion-related validity of cognitive ability tests: A qualitative and quantitative review. *Journal of Applied Psychology*, *96*(5), 881–906.

Borman, G. D., & Pyne, J. (2016). What if Coleman had known about stereotype threat? how social-psychological theory can help mitigate educational inequality. *Russell Sage Foundation Journal of the Social Sciences*, *2*(95), 164–185. www.jstor.org/stable/10.7758/rsf.2016.2.5.08

Bowden, S. C., Saklofske, D. H., & Weiss, L. G. (2011). Invariance of the measurement model underlying the Wechsler Adult Intelligence Scale-IV in the United States and Canada. *Educational and Psychological Measurement*, *71*(1), 186–199.

Brislin, R., Worthley, R., & Macnab, B. (2006). Cultural intelligence: Understanding behaviors that serve people's goals. *Group and Organizational Management*, *31*(1), 40–55.

Brown, R. P., & Day, E. (2006). The difference isn't Black and White: Stereotype threat and the race gap on Raven's Advanced Progressive Matrices. *Journal of Applied Psychology*, *91*(4), 979–985.

Ceci, S. J. (1996). *On intelligence . . . more or less: A bioecological treatise on intellectual development*. Englewood Cliffs, NJ: Prentice-Hall.

Chan, A., Yeung, D., Chan, Y. L., He, W. J., Cheung, M. C., Lam, J., et al. (2002). Different neurocognitive semantic processes for alphabetic and logographic languages. Abstract presented at the 30th Annual Meeting of the International Neuropsychological Society, Toronto, Canada, February.

Chang, D. F., & Demyan, A. (2007). Teachers' stereotypes of Asians, Blacks, and White students. *School Psychology Quarterly*, *22*(2), 91–114.

Chang, D. F., & Sue, S. (2003). The effects of race and problem type on teachers' assessment of student behavior. *Journal of Consulting and Clinical Psychology*, *71*, 235–242.

Cheung, H., & Kemper, S. (1993). Recall and articulation of English and Chinese words by Chinese English bilinguals. *Memory and Cognition*, *21*(5), 666–670.

Cheung, F. M., Leong, F. T. L., & Ben-Porath, Y. S. (2003). Psychological assessment in Asia: Introduction to the special section. *Psychological Assessment*, *15*, 243–247.

Chincotta, D., & Underwood, G. (1997). Digit span and articulatory suppression: A cross-linguistic comparison. *European Journal of Cognitive Psychology*, *9*(1), 89–96.

Chiu, S. H., & DiMarco, M. (2010). A pilot study comparing two developmental screening tools for use with homeless children. *Journal of Pediatric Health Care*, *24*(2), 73–80.

Cohen, G. L., & Sherman, D. K. (2005). Stereotype threat and the social and scientific contexts of the race achievement gap. *American Psychologist*, *60*(3), 270–271.

Cole, M. (2017). Idiocultural design as a tool of cultural psychology. *Perspectives on Psychological Science*, *12*(5), 772–781.

Cormier, D. C., Wang, K., & Kennedy, K. E. (2016). Linguistic demands of the oral directions for administering the WISC-IV and WISC-V. *Canadian Journal of School Psychology*, *31*(4), 290–304.

Dale, B. A., Finch, M. H., Mcintosh, D. E., Rothlisberg, B. A., & Finch, W. H. (2014). Utility of the Stanford-Binet Intelligence Scales, Fifth Edition, with ethnically diverse preschoolers. *Psychology in the Schools*, *51*(6), 581–590.

Dickens, W. T., & Flynn, J. R. (2006). Black Americans reduce the racial IQ gap: Evidence from standardization samples. *Psychological Science*, *17*, 913–920.

Ellis, C., Kuipers, J. R., Thierry, G., Lovett, V., Turnbull, O., & Jones, M. W. (2015). Language and culture modulate online semantic processing. *Social Cognitive and Affective Neuroscience*, *10*(10), 1392–1396.

Fagan, J. F., & Holland, C. R. (2006). Racial equality in intelligence: Predictions from a theory of intelligence as processing. *Intelligence*, *35*(4), 319–334.

Fenollar-Cortes, J., & Watkins, M. W. (2018). Construct validity of the Spanish version of the Wechsler Intelligence Scale for Children Fifth Edition (WISC-V Spain). *International Journal of School and Educational Psychology*. https://doi.org/10.1080/21683603.2017.1414006

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2007). *Essentials of cross-battery assessment* (2nd ed.). San Francisco: Wiley.

Ford, D. Y., Grantham, T. C., & Whiting, G. W. (2008). Culturally and linguistically diverse students in gifted education: Recruitment and retention issues. *Exceptional Children*, *74*(3), 289–306.

Ford, D. Y., & Helms, J. E. (2012). Overview and introduction: Testing and assessing African Americans: “Unbiased” tests are still unfair. *Journal of Negro Education*, *81*(3), 186–189.

Franklin, V. P. (2007). The tests are written for the dogs: The *Journal of Negro Education*, African American children, and the intelligence testing movement in historical perspective. *Journal of Negro Education*, *76*(3), 216–231.

Garcia Coll, C., & Marks, A. K. (Eds.) (2012). *The immigrant paradox in children and adolescents: Is becoming American a developmental risk?* Washington: American Psychological Association.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Geertz, C. (1973). *The interpretation of cultures: Selected essays by Clifford Geertz*. New York: Basic Books.

Goldberger, N. R., & Veroff, J. B. (Eds.) (1995). *The culture and psychology reader*. New York: New York University Press.

Greenfield, P. M., Maynard, A. E., & Childs, C. P. (2003). Historical change, cultural learning, and cognitive representation in Zinacantec Maya children. *Cognitive Development*, *18*, 455–487.

Hadd, A. R., & Rodgers, J. L. (2017). Intelligence, income, and education as potential influences on a child’s home environment: A (maternal) sibling-comparison design. *Developmental Psychology*, *53*(7), 1286–1299.

Hagie, M. U., Gallipo, P. L., & Svien, L. (2003). Traditional culture versus traditional assessment for American Indian students: An investigation of potential test item bias. *Assessment for Effective Intervention*, *29*(1), 15–25.

Harris, A. M., Reynolds, M. A., & Koegel, H. M. (1996). Nonverbal assessment: Multicultural perspectives. In L. A. Suzuki, P. J. Meller, & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (pp. 223–252). San Francisco: Jossey-Bass.

Hedden, T., Park, D. C., Nisbett, R., Ji, L. J., Jing, Q., & Jiao, S. (2002). Cultural variation in verbal versus spatial neuropsychological function across the life span. *Neuropsychology*, *16*(1), 65–73.

Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist*, *47*(9), 1083–1101.

Helms, J. E. (1995). An update of Helms’ White and people of color racial identity models. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 181–198). Thousand Oaks, CA: Sage.

Helms, J. E. (2002). A remedy for the Black-White score disparity. *American Psychologist*, *57* (4), 303–305.

Helms, J. E. (2004). The 2003 Leona Tyler Award Address: Making race a matter of individual differences within groups. *Counseling Psychologist*, *32*(3), 473–483.

Helms, J. E. (2005). Stereotype threat might explain the black-white test score difference. *American Psychologist*, *60*(3), 269–270.

Helms, J. E. (2006). Fairness is not validity or cultural bias in racial-group assessment: A quantitative perspective. *American Psychologist*, *61*(88), 845–859.

Herbers, J. E., Cutuli, J. J., Supkoff, L. M., Narayan, A. J., & Masten, A. S. (2014). Parenting and coregulation: Adaptive systems for competence in children experiencing homelessness. *American Journal of Orthopsychiatry*, *84*(4), 420–430.

Hintze, J. M. (2009). Curriculum-based assessment. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th ed., pp. 397–409). Hoboken, NJ: Wiley.

Hwa-Froelich, D. A., & Matsuo, H. (2005). Vietnamese children and language-based processing tasks. *Language, Speech and Hearing Services in Schools*, *36*(3), 230–243.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*, 1–123.

Kroeber, A. L., & Kluckhohn, C. (1963). *Culture: A critical review of concepts and definitions*. Cambridge, MA: Harvard University Press.

Lakin, J. M., & Lai, E. R. (2012). Multigroup generalizability analysis of verbal, quantitative, and nonverbal ability tests for culturally and linguistically diverse students. *Educational and Psychological Measurement*, *72*(1), 138–158.

Loo, S. K., & Rappaport, M. D. (1998). Ethnic variations in children's problem behaviors: A cross-sectional, developmental study of Hawaii school children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*, 567–575.

López, S. R., & Guarnaccia, P. J. J. (2000). Cultural psychopathology: Uncovering the social world of mental illness. *Annual Review of Psychology*, *51*, 571–598.

Lynn, R. (2015). *Race differences in intelligence: An evolutionary analysis* (2nd ed.). Arlington, VA: Washington Summit Publisher.

Mackintosh, N. J. (2011). *IQ and human intelligence* (2nd ed.). Oxford: Oxford University Press.

May, E. M., Azar, S. T., & Matthews, S. A. (2018). How does the neighborhood "come through the door?" Concentrated disadvantage, residential instability, and the home environment for preschoolers. *American Journal of Community Psychology*, 61, 218–228.

Mayer, D. M., & Hanges, P. J. (2003). Understanding stereotype threat effect with "culture-free" tests: An examination of its mediators and measurement. *Human Performance*, *16*(3), 207–230.

Maynard, A. E., Subrahmanyam, K., & Greenfield, P. M. (2005). Technology and the development of intelligence: From the loom to the computer. In R. J. Sternberg & D. D. Priess (Eds.), *The impact of tools on the nature and development of human abilities* (pp. 29–53). New York: Routledge.

McCarter, S. (2017). The school-to-prison pipeline: A primer for social workers. *Social Work*, *62*(1), 53–61.

McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR): Gf-Gc Cross Battery Assessment*. Boston: Allyn & Bacon.

McShane, D. (1980). A review of scores of American Indian children on the Wechsler Intelligence Scale. *White Cloud Journal*, *2*, 18–22.

Meller, P. J., & Ohr, P. S. (1996). The assessment of culturally diverse infants and preschoolers. In L. A. Suzuki, P. J. Meller, & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological and educational applications* (pp. 561–610). San Francisco: Jossey-Bass.

Morris, M. (2016). *PUSHOUT: The criminalization of black girls in schools*. New York: The New Press.

Mpofu, E., & Ortiz, S. O. (2009). Equitable assessment practices in diverse contexts. In E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 41–76). New York: Springer.

Nisbett, R. E. (2009). *Intelligence and how to get it: Why schools and cultures count*. New York: Norton.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., et al. (2012). Group differences in IQ are best understood as environmental in origin. *American Psychologist*, *67*(6), 503–504.

Oakland, T. (2009). How universal are test development and use? In E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 1–40). New York: Springer.

Okazaki, S., & Sue, S. (2000). Implications of test revisions for assessment with Asian Americans. *Psychological Assessment*, *12*(30), 272–280.

Ortiz, S. O., & Ochoa, S. H. (2005). Advances in cognitive assessment of culturally linguistically diverse individuals. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (2nd ed., pp. 234–250). New York: Guilford Press.

Parkinson, C., & Wheatley, T. (2015). The repurposed social brain. *Trends in Cognitive Sciences*, *19*(3), 133–141.

Pearson (2017). Wechsler Intelligence Scale for Children – Fifth Edition Spanish (WISC-V Spanish). www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Cognition-%26-Neuro/Wechsler-Intelligence-Scale-for-Children–Fifth-Edition-Spanish/p/100001936.html

Puente, A. E., & Puente, A. N. (2009). The challenge of meaning, abilities and competence in Hispanic/Latinos. In E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 417–441). New York: Springer.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. I. Duncan & R. J. Murname (Eds.), *Whither opportunity? Rising inequality, schools, and children's life changes* (pp. 91–116). New York: Russell Sage Foundation.

Reschly, D. J., & Bergstrom, M. K. (2009). Response to intervention. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th ed., pp. 434–460). Hoboken, NJ: Wiley.

Reynolds, C. R., & Lowe, P. A. (2009). The problem of bias in psychological assessment. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th ed., pp. 332–374). Hoboken, NJ: Wiley.

Rhodes, R. L., Ochoa, S. H., & Ortiz, S. O. (2005). *Assessing culturally and linguistically diverse students: A. practical guide*. New York: Guilford Press.

Rivera, L. M. (2008). Acculturation and multicultural assessment: Issues, trends, and practice. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment* (3rd ed., pp. 73–91). San Francisco: Jossey-Bass.

Rogoff, B., Coppens, A. D., Alcalá, L., Aceves-Azuara, I., Ruvalcaba, O., López, A., & Dayton, A. (2017). Noticing learners' strengths through cultural research. *Perspectives on Psychological Science*, *12*(5), 876–888.

Romstad, C., & Xiong, Z. B. (2017). Measuring formal intelligence in the informal learner: A case study among Hmong American students and cognitive assessment. *Hmong Studies Journal*, *18*, 1–31.

Rushton, J. P. (2012). No narrowing in mean Black-White IQ differences – predicted by heritable g. *American Psychologist*, *67*(6), 500–501.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*(2), 235–294.

Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On interpreting stereotype threat as accounting for African American–White differences on cognitive tests. *American Psychologist*, *59*(1), 7–13.

Scheiber, C. (2016). Do the Kaufman tests of cognitive ability and academic display construct bias across a representative sample of Black, Hispanic, and Caucasian school-age children in grades 1 through 12? *Psychological Assessment*, *28*(8), 942–952.

Serpell, R. (2000). Intelligence and culture. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 549–577). New York: Cambridge University Press.

Shinn, M. R., & Baker, S. K. (1996). The use of curriculum-based measurement with diverse learners. In L. A. Suzuki, P. J. Meller, & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (pp. 179–222). San Francisco: Jossey-Bass.

Short, E. L., Suzuki, L., Prendes-Lintel, M., Prendes-Lintel Furr, G., Madashushi, S., & Mapel, G. (2010). Counseling immigrants and refugees. In J. Ponterotto, M. Casas, L. Suzuki, & C. Alexander (Eds.), *The handbook of multicultural counseling* (3rd ed., pp. 201–211). Thousand Oaks, CA: Sage.

Shuttleworth-Edwards, A. B. (2016). Generally representative is representative of none: Commentary on the pitfalls of IQ test standardization in multicultural settings. *Clinical Neuropsychologist*, *30*(7), 975–998.

Smith, L. (2010). *Psychology, poverty, and the end of social exclusion*. New York: Teachers College Press.

Snyderman, M., & Rothman, S. (1988). *The IQ controversy: The media and public policy*. New Brunswick, NJ: Transaction Books.

Sotelo-Dynega, M., Ortiz, S. O., Flanagan, D. P., & Chaplin, W. F. (2013). English language proficiency and test performance: An evaluation of bilingual students with the Woodcock-Johnson III tests of cognitive abilities. *Psychology in the Schools*, *50* (8), 781–797.

Spearman, C. (1927). *The abilities of man*. New York: Macmillan.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*, 4–28.

Steele, C. M. (1998). Stereotyping and its threat are real. *American Psychologist*, *53*, 680–681.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 787–811.

Steele, C. M., & Aronson, J. (2004). Stereotype threat does not live by Steele and Aronson alone. *American Psychologist*, *59*(1), 47–48.

Sternberg, R. J. (1996). *Successful intelligence: How practical and creative intelligences determine success in life*. New York: Simon & Schuster.

Sternberg, R. J. (2004). APA presidential address: Culture and intelligence. *American Psychologist*, *59*(5), 325–338.

Sternberg, R. J. (2017). Context-sensitive cognitive and educational testing. *Educational Psychology Review*. https://link.springer.com/article/10.1007/s10648-017-9428-0

Styck, K. M., & Watkins, M. W. (2013). Diagnostic utility of the Culture Language Interpretive Matrix for the Wechsler Intelligence Scale for Children – Fourth Edition among referred students. *School Psychology Review*, *42*(4), 367–382.

Sue, D. W., & Sue, D. (2016). *Counseling the culturally diverse: Theory and practice*. Hoboken, NJ: John Wiley & Sons.

Suzuki, L. A., & Aronson, J. (2005). The cultural malleability of intelligence and its impact on the racial/ethnic hierarchy. *Psychology, Public Policy, and Law*, *11*(2), 320–327.

Suzuki, L. A., Jordan, T., Vraniak, D., Short, E. L., Aguiar, L., & Mogami, T. (2003). Meta-analysis of Wechsler studies conducted on American Indian cognitive abilities.

Poster session II presented at the 111th American Psychological Association Convention, Toronto, Canada, August.

Suzuki, L. A., Mogami, T., & Kim, E. S. (2002). Interpreting cultural variations in cognitive profiles. In K. Kurasaki, S. Okazaki, & S. Sue (Eds.), *Asian American mental health: Assessment, theory, and methods* (pp. 159–172). New York: Plenum Publishers.

te Nijenhuis, J., van den Hoek, M., & Armstrong, E. L. (2015). Spearman's hypothesis and Amerindians: A meta-analysis. *Intelligence*, *50*, 87–92.

Trimble, J. E. (2010). Cultural measurement equivalence. In C. S. Clauss-Ehlers (Ed.), *Encyclopedia of cross-cultural school psychology* (pp. 316–318). New York: Springer.

Uba, L. (1994). *Asian Americans: Personality patterns, identity, and mental health*. New York: Guilford Press.

Valencia, R. R., & Suzuki, L. A. (2001). *Intelligence testing and minority students: Foundations, performance factors, and assessment issues*. Thousand Oaks, CA: Sage.

Valencia, R. R., Suzuki, L. A., & Salinas, M. F. (2001). Test bias. In R. R. Valencia & L. A. Suzuki (Eds.), *Intelligence testing and minority students: Foundations, performance factors, and assessment issues* (pp. 111–150). Thousand Oaks, CA: Sage.

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The Think Aloud Method: A practical guide to modeling cognitive processes* (Knowledge-based systems). London: Academic Press.

Wall Street Journal (1994) Mainstream science on intelligence. December 13, p. A18.

Weiss, L. G., Harris, J. G., Prifitera, A., Courville, T., Rolthus, E., Saklofske, D. H., et al. (2006) WISC-IV interpretation in a societal context. In L. G. Weiss, D. H. Saklofske, A. Prifitera, & J. A. Holdnack (Eds.), *Wechsler Intelligence Scale for Children – IV: Advanced clinical interpretation* (pp. 1–56). San Diego, CA: Academic Press.

Weiss, L. G., Munoz, M. R., & Prifitera, A. (2015). Testing Hispanics with WISC-V and WISC-IV Spanish. In L. G. Weiss, D. H. Saklofske, J. A. Holdnack, & A. Prifitera (Eds.), *WISC-V assessment and interpretation: Scientist practitioner perspectives* (pp. 215–236). Cambridge, MA: Academic Press.

Whaley, A. L. (1998). Issues of validity in empirical tests of stereotype threat theory. *American Psychologist*, *53*, 679–680.

Whitaker, S. (2015). How accurate are modern IQ tests at categorizing people as having intellectual disability or not? *Clinical Psychology Forum*, *270*, 49–53.

16 Race and Intelligence

It's Not a Black and White Issue

Christine E. Daley and Anthony J. Onwuegbuzie

Post-Truth in Contemporary Discourse in Academia

Since the establishment in AD 859 by Princess Fatima al-Fihri, the daughter of a wealthy merchant of the Al-Karaouine mosque and university in Fes, the oldest degree-granting university in the world, based in Islamic tradition, wherein grammar, mathematics, astronomy, and medicine were taught (Glenday, 2013), higher education institutions have been deemed worldwide as representing a bastion of specialized knowledge and essential expertise, containing faculty members with high levels of intelligence who produce theory and research on issues that they believe serve the needs of various segments of society (Dorn, 2017; Geiger, 2015; Gleason, 2018). However, in recent years, the authority of higher education faculty members in general and their knowledge production in particular have been undermined and delegitimized in a contemporary period that is referred to in social and political discourse in the United States and beyond as the *post-truth era*. Indeed, in 2016, the Oxford Dictionaries selected *post-truth* as its word of the year, which the Oxford Dictionaries publisher defines as an adjective "relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief."[1] According to the Oxford Dictionaries, the first use of the term "post-truth" can be traced to a 1992 essay published in *The Nation* by the late US playwright Steve Tesich, who used this term to reflect on and to criticize events such as the 1991 US war against Iraq (Tesich, 1992). Interestingly, simply Googling "post-truth" reveals an array of news stories, essays, academic works, long-form articles, and Web 2.0 missives (e.g., blog posts, twitter posts, Facebook posts) that explain how and why we are now operating in a post-truth era.

I am honored to be able to co-author this updated chapter with the late Dr. Christine E. Daley, in spirit, as the second author. Not only was Christine an exceptional psychologist, counselor, and scholar, but also, even more importantly, my late best friend was an exceptional human being, who strived so hard through her practice and research to make the world a better place by addressing issues of social justice.

[1] Oxford Dictionaries, s.v. "Post-truth," accessed May 16, 2019, https://en.oxforddictionaries.com/definition/post-truth

This post-truth era, characterized by discourses that are no longer moored in T/truth,[2] has led at least some, if not many, academicians not only to rethink education policy and methodology (cf. Wolgemuth et al., 2018a), but also to rethink data, fact, evidence, and validity/legitimation in education policy-making, as well as the onto-ethico-epistemology of research ethics. Li and Koedel (2017) documented that Black and Hispanic faculty members are underrepresented relative to their US population shares. Thus, regardless of the validity/legitimation of knowledge claims made by academicians, if one is to assume that institutions of higher education house *intelligent* scholars who represent many different fields, then does this underrepresentation of Black and Hispanic faculty members provide justification for the claim of racial differences in intelligence? In other words, does this claim represent a *black and white* (i.e., absolute) issue? In an attempt to address these questions, it is to these claims of racial differences in intelligence that we now turn.

Racial Differences in Intelligence

The debate over racial differences in intelligence remains one of the most hotly contested issues in the social sciences today, with the preponderance of the literature and subsequent media attention focusing heavily on the alleged disparity between the cognitive abilities of Blacks and Whites. From the earliest suggestion of such discrepancies (e.g., Galton, 1892) to more sophisticated modern-day reviews and analyses such as those of Hunt and Carlson (2007a), Hocutt and Levin (1999), Sternberg, Grigorenko, and Kidd (2005), Wicherts, Borsboom, and Dolan (2010a, 2010b), and Pesta and Poznanski (2014) – including those works that have emerged during the post-truth era, such as Rindermann, Becker, and Coyle (2016) – the topic evokes no less emotional response. Indeed, if there were any doubt about the degree to which this controversy ignites sentiments in the scientific community to the point of absurdity one need only examine the case of James Watson.

Watson, one of the most famous scientists alive today, a Nobel laureate in biology whose pioneering work provided us with the molecular structure of DNA, in 2007 was pilloried by his peers and forced to resign his position as chair of Cold Spring Harbor Laboratory, because of unfortunate words uttered in his characteristically brash and uncensored style during a controversial interview that he delivered to the *Sunday Times Magazine* while on a book tour in the UK. The substance of his comments regarding contributory causes for slow economic development in southern Africa was the suggestion that social policies tend to be predicated on the assumption that Blacks and Whites are equal in intelligence, whereas testing suggests this is not the case (Ceci & Williams, 2009).

[2] Truth (i.e., upper case "T") refers to absolute Truth that represents what will be the *final opinion* perhaps at the end of history, whereas truth (i.e., lowercase "t") refers to the instrumental and provisional truths that we obtain and live by in the meantime and that are provided via experience and experimenting (cf. Johnson & Onwuegbuzie, 2004).

The firestorm surrounding the Black-White intelligence debate was elevated to particularly colossal heights following the 1994 publication of Herrnstein and Murray's controversial book, *The Bell Curve* (1994). What made this event such a sensation was the fact that the text was not limited in its distribution to the predominantly scientific community, but was released to the public in the popular press. Needless to say, this engendered fierce disputes in both the professional and lay populations, resulting in responses ranging from thoughtful consideration to acrimonious accusation. Indeed, even the present authors jumped into the fray (Arthur Jensen, personal communication, April, 1997; Onwuegbuzie & Daley, 1996, 2001). In essence, the book supported the hereditarian assumptions that intelligence is substantially genetic in origin, that the environment plays little or no role in its determination, and that IQ tests, which purport to measure it and yield a Black-White differential of fully one standard deviation, are equally valid across racial groups. Let us first examine the *fuzzy* constructs of race and intelligence.

Race as a Construct

The concept of race itself is intensely debated in the social and behavioral sciences, with some subscribing to the notion that it represents a biological fact. Those who hold this view believe that human beings can be divided into a specific number of genetically determined groups possessing similar physical characteristics such as skin color, facial features, and hair texture. For example, Rushton (2000) argues for the existence of distinct groups classified as Mongoloid (those whose ancestors were born in East Asia), Caucasoid (those of European ancestry), and Negroid (those whose origins can be traced to sub-Saharan Africa). There are a number of difficulties with this reasoning. First, most anthropologists abandoned the notion of race approximately half a century ago, arguing that *all* human beings belong to a single genus and species (i.e., Homo sapiens), and that we are all descended from an evolutionary line of humans originating in Africa approximately 200,000 years ago (Fish, 2002). Second, although there is little doubt that groups of people share common genetically transmitted physical traits, the biological perspective ignores the role of migration in the development of regional differences in these physical characteristics. Adding to the confusion is the considerable interbreeding among the so-called races in industrialized societies. According to Schaefer (1988), "Given frequent migration, exploration, and invasion, pure gene frequencies have not existed for some time, if they ever did" (p. 12). In fact, as noted by Pearson (1995), "The vast majority of blacks harbor some degree of European as well as black African ancestry, and 40 percent harbor Native American ancestry too (and some white Americans, southerners in particular, harbor black African ancestry), further complicating any attempt to draw a definitive correlation between race and intelligence" (pp. 166–167). Further, this racial intermixing compromises virtually every inferential statistical test that compares races because the samples cannot be considered independent (L. C. Wilson & Williams, 1998).

Third, there seems to be no rationale for the selection of certain physical features to determine race and not others. Why skin color and not eye color? Fourth, the fact that scientists have postulated the existence of anywhere from 3 to 200 races (Schaefer, 1988) sheds considerable light on the lack of agreement as to the criteria used to delineate categories. The reality is, there are more similarities than differences among groups and more differences within racial groups than among them (Littlefield, Lieberman, & Reynolds, 1982). In fact, in a comprehensive study, Rosenberg and colleagues (2002) found that 94 percent of the variation in the human genome is due not to population-specific genetic material, but to the variation among unrelated individuals within the *same* subgroup.

Throughout the literature, race is alternately defined as a biological feature; a local geographic population; a group linked by common descent or origin; a population connected by a shared history, nationality, or geographic distribution; a subspecies; and a social construct; and the term is used interchangeably with *ethnicity, ancestry, culture, color, national origin*, and even *religion* (Hoffman, 2006). The majority of anthropologists today contend that *race* is nothing more than a sociopolitical phenomenon (e.g., Onwuegbuzie & Daley, 2001; Smedley & Smedley, 2005), based on phenotypic differences and too often used to perpetuate caste-like stratification. Sternberg and colleagues (2005) concluded that “Race is a socially constructed concept, not a biological one. It derives from people’s desire to classify” (p. 49). Furthermore, subjective self-identification is probably the most common specification of race when it comes to classification of participants for scientific research. Yet, there are sometimes significant discrepancies between researcher identification and participant self-identification of race. For example, in one national study, 6 percent of self-identified African Americans, 29 percent of self-identified Asian Pacific Islanders, 62 percent of self-identified Native Americans, and 80 percent of participants who identified themselves with another race were categorized by the researcher as White (Massey, 1980) – representing a fatal flaw in terms of measurement. The fact is, there simply is no scientific basis for the concept of race (Sternberg et al., 2005); yet, being labeled as a member of a specific racial group has pervasive and indelible consequences psychologically, educationally, socially, and politically.

Intelligence as a Construct

As with race, there is no universally accepted definition of intelligence. Some examples include the following:

> judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting one’s self to circumstances. To judge well, to comprehend well, to reason well, these are the essential activities of intelligence. (Binet & Simon, 1916, pp. 42–43)
>
> the ability to undertake activities that are characterized by (1) difficulty, (2) complexity, (3) abstractness, (4) economy, (5) adaptiveness to a goal, (6) social value, and (7) the emergence of originals, and to maintain such activities under conditions that demand a concentration of energy and a resistance to emotional forces. (Stoddard, 1943, p. 4)

> The aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment. (Wechsler, 1958, p. 7)
>
> a human intellectual competence must entail a set of skills of problem solving – enabling the individual to resolve genuine problems or difficulties that he or she encounters, and, when appropriate, to create an effective product – and must also entail the potential for finding or creating problems – thereby laying the groundwork for the acquisition of new knowledge. (Gardner, 1983, pp. 60–61)

The question thus arises, "How does one purport to measure a construct for which there is no consensus explanation?" Despite the obvious conundrum, researchers and test publishers throughout the years have continued in their efforts to unearth the "Holy Grail" of assessment instruments capable of capturing the elusive concept of intelligence. The extent to which this undertaking has been successful depends on whether or not one is willing to accept as evidence the rather significant degree of correlation among scores generated by these assorted measures and, even more fundamentally, whether or not one is willing to accept the equivalency of *intelligence* and *IQ*.

Whence the term IQ? *IQ* was an expression coined in the early part of the twentieth century to refer to the quotient obtained when one multiplied the ratio of mental age (a concept developed by Alfred Binet and Theodore Simon in France in 1905) to chronological age by 100. Examples of early tests of mental ability, that is, IQ, included the US Army's Alpha and Beta tests used to classify and to assign large numbers of recruits prior to World War I. By 1916, Lewis M. Terman at Stanford University had adapted the work of Binet and Simon for use in the US school system, and within a few years, the term *IQ* had become part of the popular vernacular. It remains today – even during this post-truth era – a convenient, albeit unfortunate, synonym for intelligence, to which James Watson owes his demise.

Admittedly, intelligence testing has come a long way in the past 100 years. Developers of modern tests of cognitive ability have attempted to achieve culture neutrality and tap a broader spectrum of underlying skills, and IQ has become a far more psychometrically sophisticated concept. Examples include the Wechsler scales (Wechsler, 2002, 2003, 2008) and the Stanford-Binet Intelligence Scales (Roid, 2003), which yield a *Full Scale IQ*; the Kaufman Assessment Battery for Children (Kaufman & Kaufman, 2004), generating a *Mental Processing Index* (Luria model) or a *Fluid-Crystallized Index* (Cattell-Horn model); the Woodcock-Johnson Tests of Cognitive Abilities (Woodcock, McGrew, & Mather, 2007), yielding a *General Intellectual Ability Score*; and the Das-Naglieri Cognitive Assessment System (Naglieri & Das, 1997), producing a *Full Scale Standard Score*. Despite what one chooses to call them, however, what these summary scores capture at best is a narrow set of cognitive abilities represented by a unitary construct identified by researchers as Spearman's *g*, or simply *g*, and bearing little resemblance to the definitions of intelligence found throughout the literature. That is, although these measures come in many forms and comprise a variety of subtests evaluating, for instance, an individual's facility with verbal or symbolic reasoning, pattern recognition, detecting similarities or details, or processing information quickly, their scores tend to be highly intercorrelated, suggesting some overarching factor common to all of them but independent of their specific subject matter. This factor, *g*, is argued by some (e.g., Jensen, 1969, 1998) to represent the essence of human intellectual ability.

The validity of *g* as a singular estimator of intelligence has long been contested (e.g., Gould, 1996; Kamin, 1997). Critics of this view contend that key cognitive abilities are poorly evaluated or left entirely untapped by traditional intelligence tests. Sternberg (1997a), for example, has posited a triarchic model of intelligence in which analytical abilities (in essence, *g*) are equally weighted against practical abilities (pragmatic and social skills) and creativity (the ability to generate novel solutions to problems). Thus, *intelligence* becomes a *system* in which the *internal* and *external* worlds of the individual are mediated by their experiences (Sternberg, 1997b). An even broader approach is that taken by Gardner (2006), who proposed the existence of at least nine types of intelligence: linguistic, logical-mathematical, spatial, musical, bodily-kinesthetic, naturalistic, interpersonal, intrapersonal, and (at least provisionally) existential. According to Gardner, those who endorse the primacy of *g* confuse intelligence with a highly specific type of scholastic aptitude, what others (e.g., Fagan, 1992, 2000) contend is *knowledge* acquired in a cultural context. And herein lies the conundrum, for traditional *g*-saturated tests of intelligence have been found, in general, to be very good predictors of performance in the educational environment for both Blacks and Whites (e.g., McCardle, 1998; Rushton, Skuy, & Fridjohn, 2003). This phenomenon, referred to in the literature as *positive manifold*, derives from the observation that individuals who perform well on one domain measure will perform equally well on other measures in the same or similar domains (Neisser, 1998). According to Onwuegbuzie (2003), this presents a threat to validity such that the resulting correlations, in this case between scores on IQ tests and scores on measures of educational performance, might result in incorrect inferences.

Whereas there exist less data on IQ as a predictor of achievement in the workplace (Hunt & Carlson, 2007a), one must consider syllogistically that if IQ scores between Blacks and Whites differ on average by 15 points (i.e., one standard deviation), and if IQ scores are equally predictive of educational success for both Blacks and Whites (positive manifold), then Whites have a decided advantage in situations where ability scores are used to determine access to higher education. Proceeding with this logic, it follows that higher education would provide access to more prestigious and lucrative employment opportunities for Whites, including the world of academe (see, e.g., Li & Koedel, 2017). If one then considers the reported correlations between socioeconomic status (SES) and childhood IQ (e.g., Gottfried et al., 2003; Liaw & Brooks-Gunn, 1994; Smith, Brooks-Gunn, & Klebanov, 1997) and the fact that Blacks in the United States tend to be disproportionately represented in the lower socioeconomic classes, one encounters a classic example of circular reasoning. Or as Layzer (1995) observed, "intelligence is what is measured by tests that successfully predict success in enterprises whose success is commonly believed to depend strongly on what is measured by tests that successfully predict success in enterprises whose success is commonly believed to depend strongly on . . . " (p. 669).

Thus, it would appear that the practice of equating intelligence with an IQ score helps to perpetuate – and even exacerbate – the continuing disparity between success rates of Blacks and Whites in the United States. However, with all due respect to Boring (1923), intelligence is *not* simply whatever it is that IQ tests measure.

A Question of Validity

A particular difficulty with IQ instruments is that, historically, they have not been subjected to comprehensive and rigorous score validation. Onwuegbuzie, Daniel, and Collins (2009), in an extension of Messick's (1989, 1995) theory, have provided a comprehensive framework that can be used to assess the fidelity of IQ tests. This *meta-validation model*, presented in Table 16.1, suggests that content-, criterion-, and construct-related validity each can be further partitioned into validity subtypes.

Table 16.1 *Areas of validity evidence*

Validity type	Description
Criterion-related:	
Concurrent validity	Assesses the extent to which scores on an instrument are related to scores on another, already established instrument administered approximately simultaneously or to a measurement of some other criterion that is available at the same point in time as the scores on the instrument of interest
Predictive validity	Assesses the extent to which scores on an instrument are related to scores on another, already established instrument administered in the future or to a measurement of some other criterion that is available at a future point in time as the scores on the instrument of interest
Content-related:	
Face validity	Assesses the extent to which the items appear relevant, important, and interesting to the respondent
Item validity	Assesses the extent to which the specific items represent measurement in the intended content area
Sampling validity	Assesses the extent to which the full set of items sample the total content area
Construct-related:	
Substantive validity	Assesses evidence regarding the theoretical and empirical analysis of the knowledge, skills, and processes hypothesized to underlie respondents' scores
Structural validity	Assesses how well the scoring structure of the instrument corresponds to the construct domain
Convergent validity	Assesses the extent to which scores yielded from the instrument of interest being highly correlated with scores from other instruments that measure the same construct

Table 16.1 (*cont.*)

Validity type	Description
Discriminant validity	Assesses the extent to which scores generated from the instrument of interest are slightly but not significantly related to scores from instruments that measure concepts theoretically and empirically related to but not the same as the construct of interest
Divergent validity	Assesses the extent to which scores yielded from the instrument of interest are not correlated with measures of constructs antithetical to the construct of interest
Outcome validity	Assesses the meaning of scores and the intended and unintended consequences of using the instrument
Generalizability	Assesses the extent that meaning and use associated with a set of scores can be generalized to other populations

Adapted from Collins, Onwuegbuzie, and Sutton (2006). Reprinted with permission of Learning Disabilities Worldwide.

It might be argued that the validity evidence for IQ tests is at least reasonable with respect to criterion-related validity (i.e., both concurrent and predictive validity). For example, as noted previously, IQ scores have been found to forecast an array of educational, occupational, and financial outcomes. Further, it might be contended that at least moderate evidence has been documented for three elements of construct-related validity – namely, convergent validity, divergent validity, and structural validity.

Convergent validity appears to be the most strongly substantiated, with scores from the target intelligence scale often being highly correlated with scores from one or more other intelligence scales (e.g., Jazayeri & Poorshahbaz, 2003). Similarly, evidence of divergent validity is routinely provided for measures of IQ by demonstrating a low correlation with variables deemed to have an irrelevant relationship (e.g., Kolar, 2001). Evidence of structural validity has been provided by researchers who have documented the existence of *g* via exploratory factor analysis, although others have expressed concern about the instability of the extracted factors and the inconsistency in the number and nature of factors (e.g., Carroll, 1993; Caruso & Cliff, 1998; Frank, 1983; Geary & Whitworth, 1988; Kamphaus et al., 1994; O'Grady, 1989, 1990). Even if one accepts the existing support for structural validity, sufficient evidence appears to be lacking with regard to the remaining construct-related validity types: substantive validity, discriminant validity, outcome validity, and generalizability.

In the context of IQ tests, substantive validity refers to the extent that the nature of the IQ testing process is consistent with the construct (i.e., intelligence) being measured. Unfortunately, because knowledge is limited with regard to the range of cognitive processes involved as individuals respond to items on an IQ test, it is difficult to claim that researchers have provided sufficient evidence of substantive validity regarding IQ scores. Some have attempted to develop IQ measures based on tested models of cognitive processing – in particular, the Cognitive Assessment System (CAS; Naglieri & Das, 1997). However, as noted by Telzrow (1990), “the degree to which the CAS meets the authors’ stated objectives of providing diversity in content and mode of presentation varies among the PASS [planning, attention, simultaneous, and successive] domains” (p. 344). A further criticism of IQ tests relative to substantive validity is that they focus more on acquired knowledge than on the ability to learn (Kolar, 2001).

As noted in the section Intelligence as a Construct, discriminant validity of IQ tests is questionable due to positive manifold. Thus, it is not unusual for scores generated from the target IQ test to be significantly related to scores from instruments that measure concepts theoretically and empirically related to, but not the same as, the construct of interest (e.g., educational performance). Outcome validity, or what Messick (1989, 1995) termed *consequential aspects*, involves the assessment of the meaning of scores and the intended and unintended consequences of assessment use. Evidence of outcome validity related to IQ tests is particularly inadequate because of the widespread disagreement as to how IQ scores should be interpreted.

Generalizability data provide perhaps the weakest evidence of IQ score validity simply because intelligence is so inextricably embedded in culture. Indeed, Greenfield (1998) observed that (1) “cultures define intelligence by what is adaptive in their particular ecocultural niche” (p. 83) and (2) “definitions of intelligence are as much cultural ideals as scientific statements” (p. 83). Furthermore, as noted by Gould (1996), “Facts are not pure and unsullied bits of information; culture also influences what we see and how we see it. Theories, moreover, are not inexorable inductions from facts. The most creative theories are often imaginative visions imposed on facts; the source of the imagination is also strongly cultural” (p. 54). Even IQ tests designed expressly to be *culture fair*, such as Raven’s Progressive Matrices (Raven, Raven, & Court, 1995), necessitate conventional knowledge that is culture specific, such as the “ordinal relationship among the columns and among the rows as well as specific knowledge concerning what mental operations are relevant to perform on the test matrix” (Greenfield, 1998, p. 106).

Finally, there exists insufficient evidence of content-related validity with regard to IQ tests – specifically face validity, item validity, and sampling validity. Face validity is questionable because items on IQ tests are not relevant, important, or interesting for many test-takers. Indeed, negative attitudes can adversely affect the score validity of IQ tests (Steele & Aronson, 1995). Further, because IQ tests are so influenced by culture, the item content selected for IQ tests for one cultural group – even if psychometrically sound for that cultural group – likely is inappropriate for other cultural groups, thereby threatening both item validity and sampling validity.

Table 16.2 summarizes the quality of validity evidence pertaining to IQ tests extracted from the extant literature using Onwuegbuzie and colleagues’ (2009) meta-validation model. It can be seen from this table that inadequate validity evidence has been provided for IQ tests for the majority of validity types.

Table 16.2 *Interpretation of quality of validity evidence for IQ tests (using Onwuegbuzie et al.'s (2009) meta-validation model)*

Validity type	Evidence
Criterion-related:	
Concurrent validity	Strong
Predictive validity	Strong
Content-related:	
Face validity	Inadequate
Item validity	Weak
Sampling validity	Weak
Construct-Related:	
Substantive validity	Weak
Structural validity	Adequate
Convergent validity	Strong
Discriminant validity	Inadequate
Divergent validity	Adequate
Outcome validity	Weak
Generalizability	Weak

Socioeconomic Status and IQ

But let us for a moment suspend belief and assume that intelligence tests are psychometrically flawless. What of the relationship between socioeconomic status (SES) and IQ? Much of the criticism of Herrnstein and Murray (1994) centered around their quick dismissal of SES as a mitigating factor in the difference between Blacks and Whites on measures of IQ (e.g., Gardner, 1995; Lind, 1995; Nisbett, 1995). Yet, SES not only has been found to correlate with IQ (von Stumm & Plomin, 2015), but also it has been found to be associated with a number of IQ correlates, including achievement test scores (Brooks-Gunn, Guo, & Furstenberg, 1993), grade retentions, and functional literacy (Baydar, Brooks-Gunn, & Furstenberg, 1993). Noble, Norman, and Farah (2005) found that SES differences were associated with specific disparities in cognitive performance involving the brain's language and executive function systems. More recently, SES has been found to vary with language (Fernald, Marchman, & Weisleder, 2013; Hoff, 2013), executive function (Raver, Blair, & Willoughby, 2013), and memory ability (Markant et al., 2016).

Other factors that vary systematically with SES and likely play a role in creating the SES disparity in ability and achievement include physical health, home environment, neighborhood characteristics, and early education (Bornstein & Bradley, 2003). For example, SES is an important predictor of an array of health and illness outcomes (e.g., Adler & Ostrove, 1999; Anderson & Armstead, 1995; Cundiff, Matthews, & Karen, 2017), with researchers consistently documenting a strong SES gradient (i.e., lower SES corresponding to poorer health and vice versa) for cardiovascular disease, tuberculosis, chronic respiratory disease, gastrointestinal disease, arthritis, diabetes, metabolic syndrome, and adverse birth outcomes (Cantwell et al., 1998; Cundiff et al., 2017; Cunningham & Kelsey, 1984; Kaplan & Keil, 1993; Matthews et al., 1989; O'Campo et al., 1997; Pamuk et al., 1998; Robbins et al., 2001). SES also has been found to be positively related to perceptions of access and safety for physical activity, as well as to physical activity behaviors (D. K. Wilson et al., 2004), and, most recently, Jokela and colleagues (2009) found that SES largely explains the relationship between low IQ and early mortality in the United States. Furthermore, the relationships between SES and prenatal care (e.g., Elangovan et al., 2017; Lia-Hoagberg et al., 1990) and SES and nutrition (e.g., Brown & Pollitt, 1996; Kapp, Chan, & Mann, 2018) are well documented.

Home environment factors include number of siblings (Blake, 1989); the presence of two parents (Amato & Keith, 1991); home literacy or disciplinary style (Jackson et al., 2000); household resources such as books, computers, and a study room, as well as availability of after-school and summer educational services (Eccles, Lord, & Midgley, 1991; Entwisle & Astone, 1994; McLoyd, 1998); and cognitive stimulation and emotional stress levels (Noble et al., 2005).

In addition to home resources, SES, which is a primary determinant in the location of a child's neighborhood and school, also provides what Coleman (1988) referred to as *social capital*, the supportive relationships among individuals and institutions that promote the sharing of social norms and values necessary to school success. Furthermore, according to the National Research Council (1999), SES is the most important determinant of school financing in the United States, because nearly one half of all public school funding is based on local property taxes. Research comparing low-SES and higher-SES schools provided evidence of significant differences in instructional arrangements, materials, teacher experience, teacher retention, and teacher-student ratio (Wenglinsky, 1998), as well as poorer-quality relationships between school personnel and parents (Watkins, 1997). Children who live in poor school districts also have to contend with the stressors of limited social services, more violence, homelessness, and illicit drug activity (W. J. Wilson, 1987).

Although it has been argued that the benefits of early childhood education dissipate soon after termination of the program (e.g., Haskins, 1989; Herrnstein & Murray, 1994), Brooks-Gunn and colleagues (1994) demonstrated that the positive effects of intervention on verbal ability and reasoning skills were still evident two years after the end of their randomized control trial. Furthermore, a meta-analysis of the long-term benefits of early childhood education programs led to the conclusion that these interventions produce persistent, cost-effective effects on academic achievement (Barnett, 1998).

Nature versus Nurture

The relationship between IQ and SES (and its many correlates) is only one argument challenging hereditarian assumptions about the largely genetic nature of intelligence. Bouchard and colleagues (1990) found that the IQs of individuals correlated more highly with their monozygotic twins, siblings, and parents if they grew up together (0.86, 0.47, 0.42, respectively) than if they did not (0.72, 0.24, 0.22, respectively). This suggests that family environment (e.g., child-rearing practices) plays at least some role in the acquisition of intelligence. A number of other environmental factors have been identified in the literature as having either a favorable or unfavorable impact on IQ. These include exposure to toxins or hazards; diet; illness; schooling; prenatal variables such as mother's use of cigarettes, drugs, or alcohol; even duration of breastfeeding, not to mention the variety of random individual life experiences that are impossible to quantify or to control (Toga & Thompson, 2005). There also appears to be some evidence that environment can determine the relative impact of genetic variation. Turkheimer and colleagues (2003), in a study of 320 pairs of twins tested at age seven, found that environmental factors had a far more significant impact on childhood IQ in poor families (heritability = 0.10) than in wealthier families (heritability = 0.72). This suggests that *nature* might be more important at the higher end of the socioeconomic spectrum and *nurture* might be more important at the lower end (Toga & Thompson, 2005).

Still more evidence for the impact of environment on IQ is the observation of population-level increases in IQ scores over generations, a phenomenon known as the *Flynn effect*. This occurrence has been detected across tests and groups and in more than a dozen countries (Flynn, 1987). Noted increases have been attributed to improvements in education, nutrition, and healthcare; advancements in technology; and improved access to information via television and the Internet.

Other research has focused on gene-environment correlations. For example, it has been posited that more intelligent individuals tend to seek out more stimulating or challenging mental activities or might, in fact, create or evoke situations that further enhance their intellectual prowess (Plomin & Kosslyn, 2001; Ridley, 2003). Whereas there is ample documentation of the impact of heredity on intelligence (e.g., Jensen, 1998; Herrnstein & Murray, 1994), the evidence has been misconstrued to imply that IQ is static and intelligence immutable. As the foregoing arguments suggest, this is simply not the case. Furthermore, we must remember that heritability estimates are *population* statistics and cannot be applied to individuals or their IQ scores. Nor can we infer that the proportion of IQ variance explained by heredity *within* groups is equivalent to the proportion of IQ variance that it explains *between* groups. Indeed, this is one of the most grievous errors of generalization made when interpreting findings on heritability. By way of illustration, Lewontin (1982) and others (Rosenberg et al., 2002; Tishkoff & Kidd, 2004) have demonstrated that approximately 85 percent of genetic variation in a given trait occurs between any two individuals *within* a socially defined racial group and only 6 to 7 percent occurs *between* socially defined racial groups.

Summary and Conclusions

This recent and still oncoming era of post-truth has witnessed an unsettling of T/truth, amid continuously shifting and unstable intersections among policy, methodology, and evidence (cf. Wolgemuth et al., 2018a). However, this era has generated both challenges and opportunities for scholars to rethink the purpose, justification, and value of their work, as well as the validity/legitimation of their knowledge claims (Wolgemuth et al., 2018b). Foucault (2003) warned against the field of social sciences being subjected to abuse wherein certain experiences, knowledges, and wisdom traditions are marginalized or eliminated in order to produce partial elements of truth and to effect governmental power. Thus, in this post-truth era, it is essential that the politics of *all* research undergo close scrutiny. However, no social science research area warrants closer scrutiny than does intelligence research. As admonished by Onwuegbuzie and Daley (2001):

> No branch of research necessitates these qualities [of ethical research] so much as does that pertaining to intelligence because, historically, findings from this area often have led to far-reaching political interventions such as the 1922 Immigration Act, the 1924 Sterilization Law, the 1954 *Brown v. Board of Education* ruling, and, most recently, affirmative action laws. In short, thus far, intelligence research has advanced the politics of oppression, further stigmatizing and marginalizing minority individuals to justify policies and strategies that focus on elitism and exclusion instead of on meritocracy, egalitarianism, and equal opportunity. As such, research on racial differences is not just an academic exercise. (p. 218)

Therefore, in this post-truth era, with respect to intelligence research, what we have is a strong relationship between two weak phenomena (i.e., race and intelligence), one of which – intelligence – is reported to be measurable with IQ tests that happen to correlate with socioeconomic status and that represent a narrowly defined set of cognitive skills which, not surprisingly, predict similarly defined academic skills and, therefore, occupational success and wealth, which, in turn, predict intelligence as represented by an IQ score. Flawed constructs, flawed instruments, and flawed relationships yield flawed inferences and flawed educational and social policies.

What's to be done? Race appears to be a phenomenon of our human tendency to classify, perhaps driven by a need to impose order on nature (Sternberg et al., 2005). The fact is, we have been socialized to label ourselves, and we will probably continue to do so. The problem arises when those in the scientific community reify social conceptions such that they are presented as biological certainties, thereby perpetuating erroneous beliefs about between-group differences. When these beliefs are used in an attempt to advance dubious political agendas, scientists risk becoming instruments of those who would attempt to stifle the progress of minorities in the United States and elsewhere. These authors agree with the position taken by Hunt and Carlson (2007b) that studies with immediate social relevance, such as those investigating group differences, be held to higher technical and methodological standards than those examining purely scientific issues, and that risk-benefit trade-offs be considered in making decisions to publish.

We need to be clear that *IQ* is not synonymous with *intelligence* and to continue in our efforts to reach a consensus on the substance of this elusive construct. In this regard, the authors are impressed with the work of Fagan and Holland (2002, 2007, 2009) who argue that intelligence is information processing and that cultural differences in the provision of information appear to account for observed racial differences in IQ. Specifically, what Fagan and Holland's research demonstrates is that differences in knowledge between Blacks and Whites for intelligence test items can be erased when equal opportunity is provided for exposure to the information to be tested. Other studies have yielded similar findings. For example, Bridgeman and Buttram (1975) found that training in verbal strategies eliminated the differences between Black and White schoolchildren on nonverbal analogies tests; Sternberg and colleagues (2002) demonstrated that teaching cognitive skills and strategies to Tanzanian children increased their scores relative to non-trained peers on tests of syllogisms, sorting, and twenty questions; and Skuy and colleagues (2002) reported that Black South African college students benefited more from a mediated learning experience on matrices tasks than did their White counterparts. Fagan and Holland (2002) state:

> We believe that the failure to develop tests of intelligence that can be fairly applied across racial groups stems from a theoretical bias to equate the IQ score with intelligence rather than with knowledge. If we define intelligence as information processing and the IQ score as knowledge, the possibility of culture-fair tests of intelligence based on estimates of information processing arises. (p. 385)

There is little doubt that valid, unbiased measures of intellectual ability – optimally developed using *both* quantitative and qualitative research techniques (i.e., mixed methods research; cf. Onwuegbuzie, Bustamante, & Nelson, 2010) – would be useful for the processes of selection, recruitment, and promotion of individuals to positions in which they can function most effectively, both in the educational and occupational arenas. However, we must remember that intelligence is only one of many collinear variables that determine success or failure in society; that what is considered *intelligent* behavior in one context might not be relevant or valued in another; and that even conceptions of success vary from culture to culture. Furthermore, as Sternberg (2000) notes, by confusing intelligence with what society says is intelligent, we risk overlooking or giving up on individuals who have valuable skills and abilities to contribute.

In conclusion, continued research on race and intelligence is important, particularly with regard to the etiology of differences in IQ scores. In conducting studies of this nature, however, investigators must be comprehensive, transparent, and cautious, given the potential for divisiveness and far-reaching sociopolitical implications. For this reason, all such explorations should be subjected to rigorous peer review, regardless of the distinction of the authors involved. It is only by holding such research to the highest standards that we can hope to make constructive and meaningful contributions to the field in a post-truth era and beyond.

References

Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: What we know and what we don't. In N. E. Adler, M. Marmot, B. S. McEwen, & J. Stewart (Eds.), *Annals of the New York Academy of Sciences*, vol. 896, *Socioeconomic status and health in industrial nations: Social, psychological, and biological pathways*. (pp. 3–15). New York: New York Academy of Sciences.

Amato, P. R., & Keith, B. (1991). Parental divorce and adult well-being: A meta-analysis. *Journal of Marriage and the Family*, *53*, 43–58. https://doi.org/10.2307/353132

Anderson, N. B., & Armstead, C. A. (1995). Toward understanding the association of socioeconomic status and health: A new challenge for the biopsychosocial approach. *Psychosomatic Medicine*, *57*, 213–225. https://doi.org/10.1097/00006842-199505000-00003

Barnett, W. S. (1998). Long-term cognitive and academic effects of early childhood education on children in poverty. *Preventive Medicine*, *27*, 204–207. https://doi.org/10.1006/pmed.1998.0275

Baydar, N., Brooks-Gunn, J., & Furstenberg, F. (1993). Early warning signs of functional illiteracy: Predictors in childhood and adolescence. *Child Development*, *64*, 815–829. https://doi.org/10.2307/1131220

Binet, A., & Simon, T. (1916). *The development of intelligence in children* (trans. E. S. Kite). Baltimore, MD: Williams & Wilkins.

Blake, J. (1989). Number of siblings and educational attainment. *Science*, *245*, 32–37. https://doi.org/10.1126/science.2740913

Boring, E. G. (1923). Intelligence as the tests test it. *New Republic*, *36*, 35–37.

Bornstein, M. H., & Bradley, R. H. (2003). *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Lawrence Erlbaum.

Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota Study of Twins Reared Apart. *Science*, *250*, 223–228. https://doi.org/10.1126/science.2218526

Bridgeman, B., & Buttram, J. (1975). Race differences on nonverbal analogy test performance as a function of verbal strategy training. *Journal of Educational Psychology*, *67*, 586–590. https://doi.org/10.1037/h0077030

Brooks-Gunn, J., Guo, G., & Furstenberg, F. (1993). Who drops out of and who continues beyond high school? *Journal of Research on Adolescence*, *3*, 271–294. https://doi.org/10.1207/s15327795jra0303_4

Brooks-Gunn, J., McCarton, C., Casey, P., McCormick, M., Bauer, C., Bernbaum, J., & Tyson, J. (1994). Early intervention in low birthweight, premature infants. *Journal of the American Medical Association*, *272*, 1257–1262. https://doi.org/10.1001/jama.1994.03520160041040

Brown, J. L., & Pollitt, E. (1996). Malnutrition, poverty, and intellectual development. *Scientific American*, *274*(2), 38–43. https://doi.org/10.1038/scientificamerican0296-38

Cantwell, M. F., Mckenna, M. T., McCray, E., & Onorato, I. M. (1998). Tuberculosis and race/ethnicity in the United States: Impact of socioeconomic status. *American Journal of Respiratory and Critical Care Medicine*, *157*, 1016–1020. https://doi.org/10.1164/ajrccm.157.4.9704036

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, MA: Cambridge University Press.

Caruso, J. C., & Cliff, N. (1998). The factor structure of the WAIS-R: Replicability across age-groups. *Multivariate Behavioral Research*, *33*, 273–293. https://doi.org/10.1207/s15327906mbr3302_4

Ceci, S., & Williams, W. (2009). Darwin 200: Should scientists study race and IQ? YES: The scientific truth must be pursued. *Nature*, *457*, 788–789. https://doi.org/10.1038/457788a

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, *94*, S95–S120. https://doi.org/10.1086/228943

Collins, K. M. T., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed methods research in special education and beyond. *Learning Disabilities: A Contemporary Journal*, *4*, 67–100.

Cundiff, J., Matthews, M., & Karen, A. (2017). Is subjective social status a unique correlate of physical health? A meta-analysis. *Health Psychology*, *36*(12), 1109–1125. https://doi.org/10.1037/hea0000534

Cunningham, L. S., & Kelsey, J. L. (1984). Epidemiology of musculoskeletal impairments and associated disability. *Journal of Public Health*, *74*, 574–579. https://doi.org/10.2105/AJPH.74.6.574

Dorn, C. (2017). *For the common good: A new history of higher education in America*. Ithaca, NY: Cornell University Press.

Eccles, J. S., Lord, S., & Midgley, C. (1991). What are we doing to early adolescents? The impact of educational context on early adolescents. *American Journal of Education*, *99*, 521–542. https://doi.org/10.1086/443996

Elangovan, G. P., Muthu, J., Periyasamy, I. K., Balu, P., & Kumar, R. S. (2017). Self-reported prenatal oral health-care practices of preterm low birth weight-delivered women belonging to different socioeconomic status: A postnatal survey. *Journal of Indian Society of Periodontology*, *21*, 489–493. https://doi.org/10.4103/jisp.jisp_79_16

Entwisle, D. R., & Astone, N. M. (1994). Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Development*, *65*, 1521–1540. https://doi.org/10.2307/1131278

Fagan, J. F. (1992). Intelligence: A theoretical viewpoint. *Current Directions in Psychological Science*, *1*, 82–86. https://doi.org/10.1111/1467-8721.ep10768727

Fagan, J. F. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law*, *6*, 168–179. https://doi.org/10.1037/1076-8971.6.1.168

Fagan, J. F., & Holland, C. (2002). Equal opportunity and racial differences in IQ. *Intelligence*, *30*, 361–387. https://doi.org/10.1016/S0160-2896(02)00080-6

Fagan, J. F., & Holland, C. (2007). Racial equality in intelligence: Predictions from a theory of intelligence as processing. *Intelligence*, *35*, 319–334. https://doi.org/10.1016/j.intell.2006.08.009

Fagan, J. F., & Holland, C. (2009). Culture-fair prediction of academic achievement. *Intelligence*, *37*, 62–67. https://doi.org/10.1016/j.intell.2008.07.004

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Development Science*, *16*, 234–248. https://doi.org/10.1111/desc.12019

Fish, J. M. (2002). A scientific approach to understanding race and intelligence. In J. M. Fish (Ed.), *Race and intelligence: Separating science from myth* (pp. 1–28). Mahwah, NJ: Erlbaum.

Flynn, J. R. (1987). Massive IQ gains in 4 nations: What IQ tests really measure. *Psychological Bulletin*, *101*, 171–191. https://doi.org/10.1037/0033-2909.101.2.171

Foucault, M. (2003). *Society must be defended: Lectures at the Collège de France 1975–1976*. Eds. M. Bertani & A. Fontana, trans. D. Macey. New York: Picador.
Frank, G. (1983). *The Wechsler enterprise: An assessment of the development, structure, and use of the Wechsler test of intelligence*. New York: Pergamon.
Galton, F. (1892). *Hereditary genius*. London: Macmillan.
Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
Gardner, H. (1995). Cracking open the IQ box. In S. Fraser (Ed.), *The bell curve wars: Race, intelligence, and the future of America* (pp. 23–35). New York: Basic Books.
Gardner, H. (2006). *Multiple intelligences: New horizons*. New York: Basic Books.
Geary, D. C., & Whitworth, R. H. (1988). Dimensional structure of the WAIS-R: A simultaneous multi-sample analysis. *Educational and Psychological Measurement*, *48*, 945–956. https://doi.org/10.1177/0013164488484009
Geiger, R. L. (2015). *The history of American higher education: Learning and culture from the founding to World War II*. Princeton: Princeton University Press.
Gleason, N. W. (Ed.) (2018). *Higher education in the era of the fourth industrial revolution*. London: Palgrave Macmillan.
Glenday, C. (Ed.) (2013). *Guinness world records*. London: Jim Pattison Group.
Gottfried, A. W., Gottfried, A. E., Bathurst, K., Guerin, D. W., & Parramore, M. M. (2003). Socioeconomic status in children's development and family environment: Infancy through adolescence. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting and child development* (pp. 189–207). Mahwah, NJ: Erlbaum.
Gould, S. J. (1996). *The mismeasure of man*. New York: Norton.
Greenfield, P. M. (1998). The cultural evolution of IQ. In U. Neisser (Ed.), *The rising curve* (pp. 81–124). Washington: American Psychological Association.
Haskins, R. (1989). Beyond metaphor: The efficacy of early childhood education. *American Psychologist*, *44*, 274–282. https://doi.org/10.1037/0003-066X.44.2.274
Herrnstein, R. J., & Murray, C. (1994). *The bell curve*. New York: Simon & Schuster.
Hocutt, M., & Levin, M. (1999). The *Bell Curve* case for heredity. *Philosophy of the Social Sciences*, *29*, 389–415. https://doi.org/10.1177/004839319902900303
Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*, 4–14. https://doi.org/10.1037/a0027238
Hoffman, S. (2006). "Racially tailored" medicine unraveled. *American University Law Review*, *55*, 395–452.
Hunt, E., & Carlson, J. (2007a). Considerations relating to the study of group differences in intelligence. *Perspectives on Psychological Science*, *2*, 194–213. https://doi.org/10.1111/j.1745-6916.2007.00037.x
Hunt, E., & Carlson, J. (2007b). The standards for conducting research on topics of immediate social relevance. *Intelligence*, *35*, 393–399. https://doi.org/10.1016/j.intell.2006.10.002
Jackson, A. P., Brooks-Gunn, J., Huang, C., & Glassman, M. (2000). Single mothers in low-wage jobs: Financial strain, parenting and preschoolers' outcomes. *Child Development*, *71*, 1409–1423. https://doi.org/10.1111/1467-8624.00236
Jazayeri, A. R., & Poorshahbaz, A. (2003). Reliability and validity of Wechsler Intelligence Scale for Children – Third Edition (WISC-III) in Iran. *Journal of Medical Education*, *2*, 75–80.
Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*(1), 1–123. https://doi.org/10.17763/haer.39.1.l3u15956627424k7

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.1177/1558689806298224

Jokela, M., Elovainio, M., Singh-Manoux, A., & Kivimäki, M. (2009). IQ, socioeconomic status, and early death: The US National Longitudinal Survey of Youth. *Psychosomatic Medicine*, *71*, 322–328. https://doi.org/10.1097/PSY.0b013e31819b69f6

Kamin, L. J. (1997). Twin studies, heritability, and intelligence. *Science*, *278*, 1385.

Kamphaus, R. W., Benson, J., Hutchison, S., & Platt, I. O. (1994). Identification of factor models for the WISC-III. *Educational and Psychological Measurement*, *54*, 174–186. https://doi.org/10.1177/0013164494054001023

Kaplan, G. A., & Keil, J. E. (1993). Socioeconomic factors and cardiovascular disease: A review of the literature. *Circulation*, *88*, 1973–1998. https://doi.org/10.1161/01.CIR.88.4.1973

Kapp, D. S., Chan, J., & Mann, A. (2018). Socioeconomic disparities in inflammatory response on cancer mortality using national health and nutrition examination survey. *Journal of Clinical Ontology*, e13574. https://doi.org/10.1200/JCO.2018.36.15_suppl.e13574

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children – Second Edition*. San Antonio, TX: Pearson/PsychCorp.

Kolar, G. M. (2001). A literature review and critical analysis of the concurrent validity of the Differential Ability Scales and the Cognitive Assessment System. Unpublished master's thesis, University of Wisconsin-Stout, Menomonie, Wisconsin.

Layzer, D. (1995). Science or superstition? In R. Jacoby & N. Glauberman (Eds.), *The bell curve debate: History, documents, opinions* (pp. 653–681). New York: Times Books/Random House.

Lewontin, R. C. (1982). *Human diversity*. New York: Freeman.

Li, D., & Koedel, C. (2017). Representation and salary gaps by race-ethnicity and gender at selective public universities. *Educational Researcher*, *46*, 343–354. https://doi.org/10.3102/0013189X17726535

Lia-Hoagberg, B., Rode, P., Skovholt, C., Oberg, C., Berg, C., Mullett, S., & Choi, T. (1990). Barriers and motivators to prenatal care among low-income women. *Social Science and Medicine*, *30*, 487–495. https://doi.org/10.1016/0277-9536(90)90351-R

Liaw, F. R., & Brooks-Gunn, J. (1994). Cumulative familial risks and low birthweight children's cognitive and behavioral development. *Journal of Clinical Child Psychology*, *23*, 360–372. https://doi.org/10.1207/s15374424jccp2304_2

Lind, M. (1995). Brave new right. In S. Fraser (Ed.), *The bell curve wars: Race, intelligence, and the future of America* (pp. 172–178). New York: Basic Books.

Littlefield, A., Lieberman, L., & Reynolds, L. T. (1982). Redefining race: The potential demise of a concept in anthropology. *Current Anthropology*, *23*, 641–647. https://doi.org/10.1086/202915

Markant, J., Ackerman, L. K., Nussenbaum, K., & Amso, D. (2016). Selective attention neutralizes the adverse effects of low socioeconomic status on memory in 9-month-old infants. *Developmental Cognitive Neuroscience*, *18*, 26–33. https://doi.org/10.1016/j.dcn.2015.10.009

Massey, J. T. (1980). A comparison of interviewer observed race and respondent reported race in the National Health Interview Survey. In *Proceedings of the American Statistical*

Association, Social Statistics Section (pp. 425–428). Washington: American Statistical Association.

Matthews, K. A., Kelsey, S. F., Meilahn, E. N., Kuller, L. H., & Wing, R. R. (1989). Educational attainment and behavioral and biologic risk factors for coronary heart disease in middle-aged women. *American Journal of Epidemiology*, *129*, 1132–1144. https://doi.org/10.1093/oxfordjournals.aje.a115235

McCardle, J. J. (1998). Contemporary statistical models for examining test bias. In J. J. McCardle & R. W. Woodcock (Eds.), *Human cognitive abilities: Theory and practice* (pp. 157–196). Mahwah, NJ: Erlbaum.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist*, *53*, 185–204. https://doi.org/10.1037/0003-066X.53.2.185

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). Old Tappan, NJ: Macmillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Naglieri, J. A., & Das, J. P. (1997). *Das-Naglieri Cognitive Assessment System*. Rolling Meadows, IL: Riverside.

National Research Council (1999). *Equity and adequacy in education finance: Issues and perspectives*. Washington: National Research Council Committee on Education Finance.

Neisser, U. (1998). Rising test scores. In U. Neisser (Ed.), *The rising curve* (pp. 3–22). Washington: American Psychological Association.

Nisbett, R. (1995). Race, IQ, and scientism. In S. Fraser (Ed.), *The bell curve wars: Race, intelligence, and the future of America* (pp. 36–57). New York: Basic Books.

Noble, K., Norman, M., & Farah, M. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, *8*(1), 74–87. https://doi.org/10.1111/j.1467-7687.2005.00394.x

O'Campo, P., Xue, X., Wang, M. C., & Caughy, M. (1997). Neighborhood risk factors for low birthweight in Baltimore: A multilevel analysis. *American Journal of Public Health*, *87*, 1113–1118. https://doi.org/10.2105/AJPH.87.7.1113

O'Grady, K. (1989). Factor structure of the WISC-R. *Multivariate Behavioral Research*, *24*, 177–193. https://doi.org/10.1207/s15327906mbr2402_3

O'Grady, K. (1990). A confirmatory maximum factor analysis of the WPPSI. *Personality and Individual Differences*, *11*, 135–190. https://doi.org/10.1016/0191-8869(90)90005-C

Onwuegbuzie, A. J. (2003). Expanding the framework of internal and external validity in quantitative research. *Research in the Schools*, *10*(1), 71–90.

Onwuegbuzie, A. J., Bustamante, R. M., & Nelson, J. A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research*, *4*, 56–78. https://doi.org/10.1177/1558689809355805

Onwuegbuzie, A. J., & Daley, C. E. (1996). Myths surrounding racial differences in intelligence: A statistical, sociological, social psychological, and historical critique of The Bell Curve. Paper presented to students and faculty at the University of Cape Town, South Africa, May.

Onwuegbuzie, A. J., & Daley, C. E. (2001). Racial differences in IQ revisited: A synthesis of nearly a century of research. *Journal of Black Psychology*, *27*, 209–220. https://doi.org/10.1177/0095798401027002004

Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. T. (2009). A meta-validation model for assessing the score-validity of student teacher evaluations. *Quality and Quantity: International Journal of Methodology*, *43*, 197–209. https://doi.org/10.1007/s11135-007-9112-4

Pamuk, E., Makuc, D., Heck, K., Reuben, C., & Lochner, K. (1998). *Socioeconomic status and health chartbook. Health, United States, 1998*. Hyattsville, MD: National Center for Health Statistics.

Pearson, H. (1995). Developing the rage to win. In S. Fraser (Ed.), *The bell curve wars: Race, intelligence, and the future of America* (pp. 164–171). New York: Basic Books.

Pesta, B. J., & Poznanski, P. J. (2014). Only in America: Cold winters theory, race, IQ and well-being. *Intelligence*, *46*, 271–274. https://doi.org/10.1016/j.intell.2014.07.009

Plomin, R., & Kosslyn, S. M. (2001). Genes, brain and cognition. *Nature Neuroscience*, *4*, 1153–1154. https://doi.org/10.1038/nn1201-1153

Raven, J., Raven, J. C., & Court, J. H. (1995). *Manual for Raven's Progressive Matrices and Vocabulary Scales* (Section J, General Overview).Oxford: Oxford Psychologists Press.

Raver, C. C., Blair, C., & Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: New perspectives on models of differential susceptibility. *Developmental Psychology*, *49*, 292–304. https://doi.org/10.1037/a0028343

Ridley, M. (2003). *Nature via nurture: Genes, experience, and what makes us human*. New York: HarperCollins.

Rindermann, H., Becker, D., & Coyle, T. R. (2016). Survey of expert opinion on intelligence: Causes of international differences in cognitive ability tests. *Frontiers in Psychology*, 7. https://doi.org/10.3389/fpsyg.2016.00399

Robbins, J. M., Vaccarino, V., Zhang, H., & Kasl, S. V. (2001). Socioeconomic status and type 2 diabetes in African American and non-Hispanic white women and men: Evidence from the Third National Health and Nutrition Examination Survey. *American Journal of Public Health*, *91*, 76–83. https://doi.org/10.2105/AJPH.91.1.76

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales – Fifth Edition*. Rolling Meadows, IL: Riverside.

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Field, K. K., Zhivotovsky, L. A., et al. (2002). Genetic structure of human populations. *Science*, *298*, 2381–2385. https://doi.org/10.1126/science.1078311

Rushton, J. P. (2000). *Race, evolution, and behavior: A life-history perspective* (3rd ed.). Port Huron, MI: Charles Darwin Research Institute.

Rushton, J. P., Skuy, M., & Fridjohn, P. (2003). Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White engineering students in South Africa. *Intelligence*, *31*, 123–137. https://doi.org/10.1016/S0160-2896(02)00140-X

Schaefer, R. T. (1988). *Racial and ethnic groups* (3rd ed.). Glenview, IL: Scott Foresman.

Skuy, M., Gewer, A., Osrin, Y., Khunou, D., Fridjohn, P., & Rushton, J. P. (2002). Effects of mediated learning experiences on Raven's matrices scores of African and non-African university students in South Africa. *Intelligence*, *30*, 221–232. https://doi.org/10.1016/S0160-2896(01)00085-X

Smedley, A., & Smedley, B. (2005). Race as biology is fiction, racism as a social problem is real. *American Psychologist*, *60*, 16–26. https://doi.org/10.1037/0003-066X.60.1.16

Smith, J., Brooks-Gunn, J., & Klebanov, P. (1997). Consequences of living in poverty for young children's cognitive and verbal ability and early school achievement. In

G. Duncan & J. Brooks-Gunn (Eds.), *Consequences of growing up poor* (pp. 132–189). New York: Russell Sage.
Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811. https://doi.org/10.1037/0022-3514.69.5.797
Sternberg, R. J. (1997a). *Successful intelligence*. New York: Plume.
Sternberg, R. J. (1997b). The triarchic theory of intelligence. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 92–104). New York: Guilford Press.
Sternberg, R. J. (2000). Implicit theories of intelligence as exemplar stories of success: Why intelligence test validity is in the eye of the beholder. *Psychology, Public Policy, and Law*, *6*, 159–167. https://doi.org/10.1037/1076-8971.6.1.159
Sternberg, R. J., Grigorenko, E. L., & Kidd, K. K. (2005). Intelligence, race, and genetics. *American Psychologist*, *60*(1), 46–59. https://doi.org/10.1037/0003-066X.60.1.46
Sternberg, R. J., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C., et al. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence*, *30*, 141–162. https://doi.org/10.1016/S0160-2896(01)00091-5
Stoddard, G. D. (1943). *The meaning of intelligence*. New York: Macmillan.
Telzrow, C. F. (1990). Does *PASS* pass the test? A critique of the Das-Naglieri Cognitive Assessment System. *Journal of Psychoeducational Assessment*, *6*, 344–355. https://doi.org/10.1177/073428299000800310
Tesich, S. (1992). A government of lies. *The Nation*, *254*(1), 12–14.
Tishkoff, S. A., & Kidd, K. K. (2004). Implications of biogeography of human populations for “race” and medicine. *Nature Genetics*, *36*(11, Suppl.), S21–S27. https://doi.org/10.1038/ng1438
Toga, A. W., & Thompson, P. M. (2005). Genetics of brain structure and intelligence. *Annual Review of Neuroscience*, *28*, 1–23. https://doi.org/10.1146/annurev.neuro.28.061604.135655
Turkheimer, E., Haley, A., Waldron, M., D’Onofrio, B., Gottesman, I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, *14*, 623–628. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x
von Stumm, S., & Plomin, R. (2015). Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence*, *48*, 30–36. https://doi.org/10.1016/j.intell.2014.10.002
Watkins, T. J. (1997). Teacher communications, child achievement, and parent traits in parent involvement models. *Journal of Educational Research*, *91*, 3–14. https://doi.org/10.1080/00220679709597515
Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore, MD: Williams & Wilkins.
Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence – Third Edition*. San Antonio, TX: Pearson/PsychCorp.
Wechsler, D. (2003). *Wechsler Intelligence Scale for Children – Fourth Edition*. San Antonio, TX: Pearson/PsychCorp.
Wechsler, D. (2008). *Wechsler Adult Intelligence Scale – Fourth Edition*. San Antonio, TX: Pearson/PsychCorp.
Wenglinsky, H. (1998). Finance equalization and within-school equity: The relationship between education spending and the social distribution of achievement. *Educational Evaluation and Policy Analysis*, *20*, 269–283. https://doi.org/10.3102/01623737020004269

Wicherts, J. M., Borsboom, D., & Dolan, C. V. (2010a). Evolution, brain size, and the national IQ of peoples around 3,000 years BC. *Personality and Individual Differences*, *48*, 104–106. https://doi.org/10.1016/j.paid.2009.08.020

Wicherts, J. M., Borsboom, D., & Dolan, C. V. (2010b). Why national IQs do not support evolutionary theories of intelligence. *Personality and Individual Differences*, *48*, 91–96. https://doi.org/10.1016/j.paid.2009.05.028

Wilson, D. K., Kirtland, K. A., Ainsworth, B. E., & Addy, C. L. (2004). Socioeconomic status and perceptions of access and safety for physical activity. *Annals of Behavioral Medicine*, *28*, 20–28. https://doi.org/10.1207/s15324796abm2801_4

Wilson, L. C., & Williams, D. R. (1998). Issues in the quality of data on minority groups. In V. C. McLoyd & L. Steinberg (Eds.), *Studying minority adolescents: Conceptual, methodological, and theoretical issues* (pp. 237–250). Mahwah, NJ: Lawrence Erlbaum Associates.

Wilson, W. J. (1987). The hidden agenda. In W. J. Wilson (Ed.), *The truly disadvantaged: The inner city, the underclass and public policy* (pp. 140–164). Chicago: University of Chicago Press.

Wolgemuth, J. R., Koro-Ljungberg, M., Marn, T. M., Onwuegbuzie, A. J., & Dougherty, S. M. (Eds.) (2018a). Rethinking education policy and methodology in a post-truth era. Special issue, *Education Policy Analysis Archives*, *26*(145).

Wolgemuth, J. R., Koro-Ljungberg, M., Marn, T. M., Onwuegbuzie, A. J., & Dougherty, S. M. (2018b). Start here, or here, no here: Introductions to rethinking education policy and methodology in a post-truth era. *Education Policy Analysis Archives*, *26*(145), 1–8.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2007). *Woodcock-Johnson III NU Tests of Cognitive Abilities*. Rolling Meadows, IL: Riverside.

PART IV

Biology of Intelligence

17 Animal Intelligence

Thomas R. Zentall

The notion of an evolutionary scale with humans at the top is popularly held but also self-serving. We tend to undervalue the exceptional sensory skills of tracking and drug-detecting dogs as well as the navigational abilities of homing pigeons, whales, and monarch butterflies. Conversely, we tend to overvalue our problem-solving ability, capacity to modify our environment, and ability to communicate with each other. This bias notwithstanding, taken as a whole, clearly the sum of our intellectual capacity, measured in almost any way, exceeds that of other animals. The role of our intelligence in the domination of our species over others seems obvious, but in the broader perspective of evolutionary success, as measured by the number of surviving members of a species, intelligence, as a general characteristic, correlates only superficially (and perhaps even negatively) with most measures of evolutionary success. Consider the relatively small numbers of our closest relatives, the great apes, compared with the large numbers of considerably more "primitive" insects, bacteria, and viruses. And it is estimated that if a massive disaster were to occur (e.g., if Earth were hit by a large asteroid or suffered a self-inflicted nuclear disaster), many simpler organisms would likely survive much better than large intelligent animals like us.

From a purely biological perspective, the ideal survival machine is a simple, perhaps even one-celled, organism (e.g., the amoeba) that has survived in one of two ways. Either it has needed to undergo little change in morphology or behavior for millions of years because it exists in a remarkably stable (predictable) environment, in which case there has been little need for change, or if its environment does change, it relies on natural selection by means of very rapid reproduction and occasional mutation (e.g., bacteria and viruses). This ability to reproduce quickly and often ensures the survival of many of these organisms (albeit not necessarily in the same form) even in the event of a major catastrophe.

Many other organisms whose rate of reproduction has not been able to keep up with relatively rapid changes in the environment have relied on their ability to modify their behavior during their lifetime. Intelligence, in its simplest form, can be thought of as the genetic flexibility that allows organisms to adjust their behavior to relatively rapidly changing environments. For some animals, a stable supply of a highly specific food may be predictable (e.g., eucalyptus leaves for the koala or bamboo leaves for the giant panda) – at least until recently. For most animals, however, environments are much less predictable and their predisposed eating preferences have had to be much more flexible. For still other animals, the environment is sufficiently unpredictable that it is impossible to specify (by genetic means alone) what food will be available to an individual (consider the varied diet of the city-dwelling rat). For these animals to survive, more general (abstract) rules must be available. Rules about what to eat may not be based on the genetically predisposed sight or taste of what is ingested but on the association of those sensory attributes with their consequences. Instead of instructing the animal to eat eucalyptus leaves or to eat a certain class of seeds, these genes tell the animal that if it feels sick after eating a new food, it should avoid eating more of that food. Such general rules allow for the behavioral flexibility that we call learning.

But there is a price to pay for this added flexibility. The animal must sometimes suffer the consequences of eating something bad. If the novel food is poisonous, the animal may not survive to use its newfound knowledge. The creation and maintenance of a nervous system capable of such learning represents a cost as well. For many animals, the benefits of the capacity for simple associative learning outweigh the cost, and for some animals, the negative consequences are sufficiently costly that simple learning rules are not enough.

Some animals have found ways to reduce this cost. Rats, which live in highly unpredictable environments in which resources are variable, have evolved the ability to learn the consequences of eating a small amount of a novel food in a single experience, even when those consequences are experienced hours after the food was ingested (Garcia & Koelling, 1966). Rats have also developed the ability to transmit food preferences socially. If a rat experiences the smell of a novel food on the breath of another rat, it will prefer food with that smell over another, equally novel food (Galef, 1988) and it may also be able to assess the consequences to the other rat (e.g., sickness) of having eaten a novel food (Kuan & Colwill, 1997).

But what if this degree of flexibility in learning is still not enough to allow for survival? In the case of humans, for example, our poorly developed sense of smell, our relatively poorly developed gross motor response (e.g., slow running speed relative to other large predators), and our relative physical weakness may not have allowed us to hunt competitively with other large predators. The competition with other animals for food must have come about slowly enough for us to develop weapons and tools, complex forms of communication (language), and complex social structure (allowing for cooperation, teamwork, and reciprocation). According to this hypothesis, although our intellect appears to have given us a clear advantage over other animals, its evolution is likely to have emerged because of our relative weakness in other areas. Other animals have compensated for their weaknesses by developing strengths in nonintellectual areas (e.g., the snail compensates for its lack of rapid mobility by building a protective shell around itself). Discussions of animal intelligence often assume, inappropriately, that intelligence is inherently good. In our case, it has turned out to be generally true (at least to the present). For us, intelligence has had a runaway effect on our ability to adapt to change (an effect that Dawkins, 1976, calls hypergamy), which has allowed us to produce radical changes in our environment. However, from a biological perspective, in general, intelligence can be viewed as making the best out of a bad situation, or producing a complex solution to problems that other species have often solved in simpler ways. As we evaluate the various intellectual capacities of nonhuman animals, we should try to keep in mind that they have survived quite well (until recently when we have dominated the planet) without the need for our complex intellectual skills.

The Comparative Approach: Two Caveats

First, most people have a vague idea of the relative intelligence of animals. As a general rule, those species that are more like us physically are judged to be more intelligent. But we should be careful in making such judgments because we humans are the ones deciding what intelligent behavior is. We make up the rules and the testing procedures and those tests are often biased in favor of our particular capacities. Isn't it interesting that animals that have similar sensory, motor, and motivational systems just happen to be judged as more intelligent?

Bitterman (1975) has suggested that a relational view of animal learning should be used to correct for peripheral differences in sensory capacity and motor coordination. He suggests that rather than looking for differences in the rate at which different species can learn, we might look at differences, for example, in an animal's ability to learn from the experience of learning. In other words, to what extent can learning facilitate new learning (learning to learn)? Then, using the rate of original learning as a baseline, one can determine the degree to which later learning, presumably involving the same processes, is facilitated. For example, in serial reversal learning, one assesses how much a particular species improves in learning successive reversals of a simple simultaneous discrimination, relative to the rate of original learning. This approach is not always possible, however, and we must be aware that our assessment may be biased by the use of testing procedures not well suited for the species we are studying.

Second, we must guard against the opposite bias – the tendency to interpret behavior as intelligent because of its similarity to intelligent human behavior. In evaluating research addressing the cognitive capacity of animals I will adopt C. Lloyd Morgan's (1894) position that it is not necessary to interpret behavior as complex (more cognitive) if a simpler (less cognitive) account will suffice. Thus, higher-level cognitive interpretations will always be contrasted with simpler, contiguity- and contingency-based, associative-learning accounts. I will start with several classical issues concerned with the nature of learning and intelligence in animals, transition to more complex behavior thought to be uniquely human, and end with examples of presumably complex behavior that are likely to be based on simpler predisposed processes in animals, including possibly humans.

Absolute Versus Relational Learning

One of the most basic cognitive functions is not being bound to the absolute properties of a stimulus. Although Hull (1943) suggested that learning may involve solely the absolute properties of a stimulus, he proposed that animals will appear to respond relationally because they will respond similarly to similar stimuli, a process known as stimulus generalization. Spence (1937) elaborated on this theory by proposing that discrimination learning establishes predictable gradients of excitation and inhibition that surround the training stimuli and that summate algebraically. And this theory of generalization gradient summation can account for a number of phenomena that were formerly explained as relational learning (see Riley, 1968). The fact that one sees little discussion of this issue in the modern literature suggests that animals are capable of using either the absolute or relative properties of a stimulus in making discriminations.

Learning to Learn

Can an animal use prior learning to facilitate new learning? That is, can animals learn to learn? If an animal learns a simple discrimination between two stimuli (an S+, to which responses are reinforced and an S− to which responses are extinguished) and then, following acquisition, the discrimination is reversed (the S+ becomes S− and the S− becomes S+), and then reversed again, repeatedly, are successive reversals learned faster than earlier reversals? Animals trained on such a serial-reversal task often show improvement within a few reversals and the rate of improvement can be used as a measure of learning to learn. For example, rats show more improvement than pigeons, and pigeons show more improvement than goldfish (Bitterman & Mackintosh, 1969). Mackintosh (1969) attributes these differences in serial-reversal learning to the differential ability of these species to maintain attention to the relevant dimension.

A different approach to learning to learn is to look for improvement in the rate at which discriminations involving new stimuli are learned. This phenomenon, known as learning set (Harlow, 1949), has been studied primarily using visual discriminations with monkeys but good evidence for such effects has also been found with olfactory discriminations with rats (Slotnick & Katz, 1974). In the limit, learning of a new discrimination, or of a reversal, can occur in a single trial. When it does, it is referred to as a win-stay-lose-shift strategy because stimulus choice is completely controlled by the consequences of choice on the preceding trial. One means of developing such a strategy is to learn to forget the consequences of trials prior to the immediately preceding trial. In fact, research has shown that memory for the specific characteristics of the stimuli from prior discriminations does decline as the number of discriminations learned increases (Meyer, 1971). Thus, animals approach optimal learning by learning to ignore the effects of all but the most recent experience.

Stimulus Class Formation

Perceptual Classes

Pigeons are remarkably adept at responding selectively to photographs of natural scenes depending on whether the scene involves a human form (Herrnstein & Loveland, 1964) or trees or water (Herrnstein, Loveland, &

Cable, 1976), and those objects need not be anything that they might have actually encountered in their past (e.g., underwater pictures of fish; Herrnstein & deVilliers, 1980). To demonstrate that the pigeons do not simply memorize a list of pictures and their appropriate responses, Herrnstein and colleagues (1976) showed that the pigeons could respond appropriately to new examples of the positive and negative stimulus sets.

What is interesting about perceptual classes is that it is difficult to specify what features humans or pigeons use to discriminate members from nonmembers of the perceptual class. However, examination of the kinds of errors made can tell us about the attributes that were used to categorize the exemplars and the similarities in the underlying processes. For example, pigeons make errors similar to those of young children (e.g., they often erroneously assign a picture of a bunch of celery to the category "tree").

Equivalence Relations

The emergent relations that may arise when arbitrary, initially unrelated stimuli are associated with the same response are often referred to as functional equivalence (see Zentall & Smeets, 1996) because the two stimuli can be thought of as "having the same meaning." The most common procedure for demonstrating functional equivalence involves training on two conditional discriminations. In the first, for example, a red hue (sample) signals that a response to a circle will be reinforced (but not a response to a dot) and a green hue signals that a response to a dot will be reinforced (but not a response to a circle); see Figure 17.1. In the second conditional discrimination, a vertical line signals that a response to the circle will be reinforced (but not a response to the dot) and a horizontal line signals that a response to the dot will be reinforced (but not a response to the circle). Thus, red and vertical line mean choose the circle and green and horizontal line mean choose the dot. This procedure has been referred to as many-to-one matching because training involves the association of two samples with the same comparison stimulus. To show that an emergent relation has developed between the red hue and the vertical line and between the green hue and the horizontal line, one can train new associations between one pair of the original samples (e.g., the red and green hues) and a new pair of comparison stimuli (e.g., blue and white hues, respectively). Then on test trials, one can show that emergent relations have developed when, without further training, an animal chooses blue when the sample is a vertical line and chooses white when the sample is a horizontal line (Urcuioli et al., 1989; Wasserman, DeVolder, & Coppage, 1992; Zentall, 1998).

The ability of animals to develop emergent stimulus classes involving arbitrary stimuli has important implications for human language learning because stimulus class formation plays an integral role in the acquisition of that aspect of human language known as *semantics* – the use of symbols (words) to stand for objects, actions, and attributes. The ability of small-brained organisms like pigeons to develop stimulus classes made up of arbitrary stimuli suggests that this capacity is much more pervasive than was once thought.

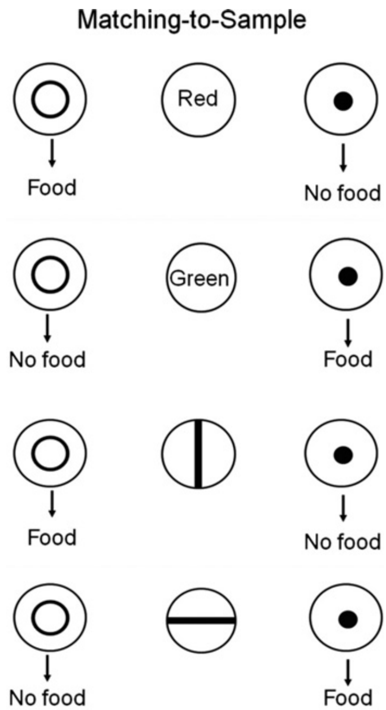

Figure 17.1 *Many-to-one matching training used to show that pigeons will learn that red and vertical (as well as green and horizontal) "mean the same thing." If red and green samples are then associated with new comparison stimuli, blue and white, respectively, there is evidence that the vertical and horizontal lines are also associated with the blue and white stimuli, respectively.*

Memory Strategies

The task most often used to study working memory in animals is delayed matching-to-sample, in which following acquisition of matching-to-sample, a delay is inserted between the offset of the sample and the onset of the comparison stimuli (Roberts & Grant, 1976). However, the retention functions typically found with this procedure generally greatly underestimate the animal's working memory capacity for two reasons. First, in many studies, the novel delay interval is quite similar to the intertrial interval, the end of the trial event. When the delay interval and the intertrial interval are made distinctive, the retention functions obtained may provide a very different picture of the animal's memory (Sherburne, Zentall, & Kaiser, 1998). Second, the novelty of the delays may result in a generalization decrement that is confounded with memory loss. When delays are not expected, they may not be "understood." When pigeons are trained with delays so they are not novel, considerably better memory has been found (Dorrance, Kaiser, & Zentall, 2000). Of more interest in the assessment of animal intelligence are strategies that animals may use to enhance memory.

Prospective Processes

Traditionally, animal memory has been viewed as a rather passive process. According to this view, sensory events can leave a trace that may control responding even when the event is no longer present (Roberts & Grant, 1976). However, it has been suggested that animals can also actively translate or code the representation of a presented stimulus into an expectation of a yet-to-be-presented event (Honig & Thompson, 1982). The use of expectations, or prospective coding processes, has important implications for the cognitive capacities of animals. If the expectation of a stimulus, response, or outcome can serve as an effective cue for responding, it suggests that animals may be capable of exerting active control over memory, and in particular, it may suggest they have the capacity for active planning.

The notion of expectancy as an active purposive process can be attributed in part to Tolman (1932). Although one can say that a dog salivates when it hears a bell because it *expects* food to be placed in its mouth, the demonstration that an expectation can serve as a discriminative stimulus (i.e., as the basis for making a choice) suggests that the expectancy has additional cognitive properties.

The differential outcomes effect. If a conditional discrimination is designed such that a correct response following one sample results in one kind of outcome (e.g., food) and a correct response following the other sample results in a different kind of outcome (e.g., water), one can show that acquisition of the conditional discrimination is facilitated (Trapold, 1970) and retention is better when a delay is inserted between the conditional and choice stimuli (Peterson, Wheeler, & Trapold, 1980). Furthermore, there is evidence from transfer-of-training experiments that in the absence of other cues, outcome anticipations can serve as sufficient cues for comparison choice. That is, if the original samples are replaced by other stimuli associated with the same differential outcomes, positive transfer has been found (Edwards et al., 1982; Peterson, 1984). This line of research indicates that presentation of a sample creates an expectation of a particular kind of outcome, and that expectation can then serve as the basis for comparison choice. In most cases, the differential outcomes have differential hedonic value (e.g., a high probability of food versus a low probability of food) and it is possible that outcome anticipation can elicit differential emotional states in the animal. But there is evidence that nondifferentially hedonic events such as the anticipation of a particular colored stimulus also can affect response accuracy (Kelly & Grant, 2001; Miller, Friedrich, Narkavic, & Zentall, 2009; Williams, Butler, & Overmier, 1990).

Planning ahead. One of the hallmarks of human cognitive behavior is our ability to consciously plan for the future. Although animals sometimes appear to plan for the future (birds build nests, rats hoard food), these behaviors are likely to be under genetic control. One also needs to distinguish between planning for the future and learning with a long delay of reinforcement. Suddendorf and Corballis (1997) have suggested that for behavior to be considered future planning, the behavior must occur at a time when the relevant motivation is not present. Thus, Roberts (2002) reported the absence of planning by monkeys which, when given their daily portion of food, after eating, threw out of their cage whatever food remained but requested more food later in the day. Further laboratory research suggested, however, that monkeys can learn to plan for the future and, under the right conditions, would choose a smaller amount of food over a larger amount (1) if more food would be provided later after they selected the smaller amount but not the larger amount or (2) if choosing the larger amount resulted in the removal of much of what was selected (Naqshbandi & Roberts, 2006).

More convincing evidence for planning was reported by Raby and colleagues (2007). Western scrub jays, which cache food for future use, learned that they might spend the night in a compartment in which, in the morning, they would find one kind of food (peanuts) or in a different compartment in which they would find a different kind of food (kibble). On test trials, they were allowed to eat and cache food in either compartment the night before. When they were given peanuts in the evening, after eating they tended to cache them in the kibble compartment and when they were given kibble in the evening, after eating they tended to cache them in the peanut compartment (i.e., the compartment in which they would not find the particular cached food in the morning).

Directed (Intentional) Forgetting

The notion of directed or intentional forgetting is borrowed from human memory research. It implies that memory is an active rather than an automatic process. Presumably, following presentation, items that participants are instructed to forget may not be well stored or maintained in memory and thus should not be well retained. In a directed forgetting task with animals, for example, pigeons are trained on a matching task and then a delay of a fixed duration is introduced between the sample and the comparisons. On "forget" trials, during the delay, the pigeons are given a cue that indicates that there will be no test of sample memory. On probe trials, however, the forget cue is presented but there is also a test of sample memory. Matching accuracy on these probe trials is generally well below that on "remember" trials on which there was an expected test of sample memory (Grant, 1981). But this design confounds differential motivation on remember and forget trials with sample memory effects (Roper & Zentall, 1993). In a more complex design that controls for motivational effects and that better approximates the human-directed forgetting procedure by allowing the animal to *reallocate* its memory from the sample to an alternative memory, better evidence for directed forgetting in pigeons has been demonstrated (Roper, Kaiser, & Zentall, 1995). Thus, under certain conditions it appears that animals do have at least some active control over their memory processes.

Episodic Memory

Human memory can be identified by the kinds of processes presumed to be involved. Procedural memory involves memory for actions (e.g., riding a bicycle) and it has been assumed that much learned behavior by animals involves this kind of memory. Human declarative memory is assumed to be more cognitive because it involves memory for facts (semantic memory) and memory of personal experiences (episodic memory). Although animals cannot typically describe factual information, their conditional rule-based learning can be thought of as a kind of semantic memory (e.g., if the sample is red choose the vertical line, if the sample is green choose the horizontal line). More controversial is whether animals have episodic memory.

Tulving (1972) proposed that an episodic memory should include the *what, where*, and *when* of a specific experience. Clayton and Dickinson (1999) showed that western scrub jays that cached peanuts and wax worms (*what*) in different compartments of an ice cube tray (*where*) learned that their preferred wax worms would be edible after one day, but after four days only the peanut would be edible (*when*; see also Babb & Crystal, 2006, for a similar finding with rats). But it can be argued that it is insufficient to retrieve the *what, where*, and *when* of an episode because those rules have been explicitly trained (i.e., they are semantic or rule-based memories). Instead, better evidence for episodic memory would come from the finding that animals can retrieve information about a past episode when there is no expectation that they will be requested to retrieve that memory in the future (Zentall et al., 2001). That is, imagine that pigeons are first trained to report the location where they recently pecked (e.g., if left, choose red, if right, choose green). Then they are trained on an unrelated conditional discrimination in which choice of a vertical line was correct when the sample is blue and choice of the horizontal line is correct when the sample is yellow. Singer and Zentall (2007) found that on probe trials on which following a vertical- or horizontal-line comparison response the pigeons were asked *unexpectedly* to report the location of the stimulus that they had pecked (they were presented with a choice between red and green stimuli), they did so reliably. Thus, by either criterion (what-where-when or responding to an unexpected question), pigeons show some evidence of episodic-like memory.

Navigation

Compared to many animals, humans have relatively poor navigational skills. Consider how dependent we are on external supports such as compasses, maps, and more recently global positioning devices. Many animals (e.g., migrating whales, birds, monarch butterflies) can navigate over many hundreds of miles using magnetic fields, chemical gradients, and star patterns. And homing pigeons use a number of these navigational systems, including landmarks consisting of natural and man-made geographic features (Lipp et al., 2004).

However, many humans have the ability to imagine a route that they will take and even to imagine how to get to a familiar destination by a novel path. This ability, known as cognitive mapping, consists of knitting together landmarks one has experienced, such that the relation among them can be used to determine a novel path to arrive at a goal. To qualify as cognitive mapping, landmarks should be needed to *form* a cognitive map but they should not be necessary to *use* it. Can animals form a cognitive map?

Some animals have the remarkable ability to navigate in the absence of landmarks or other external cues. This ability, known as path integration (or dead reckoning), involves the representation of direction and distance one has traveled from a starting point. Desert ants, which live in an environment with few stable landmarks, are particularly adept at path integration, as can be shown not only by the direct path that they take to return to their nest after a foraging trip but also by the systematic error incurred if they are displaced just before they attempt to return home (Collette & Graham, 2004). The distinction between path integration and cognitive mapping has been a point of controversy. However, under conditions that cannot be accounted for with either landmark use or path integration, there is evidence for the development of a simple cognitive map in dogs (Chapuis & Varlet, 1987), as well as in rats (Singer, Abroms, & Zentall, 2007).

Counting

The term *numerical competence* is often used in animal research because the more common term, *counting*, carries with it the surplus meaning that accompanies the human verbal labels given to numbers. That this distinction is an arbitrary one, based on limitations of response (output) capacity of animals rather than conceptual ability, is suggested by Pepperberg's (1987) work with generalized number use (with verbal English) in an African gray parrot.

An excellent review of the animal-counting literature is provided by Davis and Memmott (1982), who concluded that counting does not come easily to animals. "Although counting is obtainable in infrahumans, its occurrence requires considerable environmental support" (p. 566). In contrast, however, Capaldi (1993) concluded that under the right conditions, animals "count" routinely. In simple but elegant experiments, Capaldi and Miller (1988) demonstrated that following training, rats can anticipate whether they will get fed or not for running down an alley, depending solely on the exact number of successive times they have run down that alley and found food on successive earlier trials.

The difference in the conclusions reached by Davis and Memmott (1982) and by Capaldi and Miller (1988) has general implications for the study of intelligence in animals (including humans). The context in which one looks for a particular capacity may determine whether one will find evidence for it. Because we, as human experimenters, devise the tasks that serve as the basis for the assessment of intelligence, we must be sensitive to the possibility that these tasks may not be optimal for eliciting the behavior we are assessing. As noted earlier, much of our view of the evolutionary scale of intelligence may be biased in this way by species differences in sensory, response, and motivational factors.

Perhaps the most impressive demonstration of numerical competence in an animal was reported by Boysen and Berntson (1989). A chimpanzee, Sheba, was first trained on the correspondence between Arabic numerals and the number of objects present. When Sheba was then shown a number of objects seen at two different locations (e.g., three objects at one site and one object at another), she pointed to the numeral "4," the sum of the objects. Finally, she was shown Arabic numerals at two different sites and she spontaneously pointed to the numeral that represented the sum of the two numerals she had seen.

Reasoning

Reasoning can be thought of as a class of cognitive behavior for which correct responding on test trials requires an inference based on incomplete experience. Although, for obvious reasons, most research on reasoning in animals has been done with higher primates (e.g., chimpanzees), there is evidence that some reasoning-like behavior can be demonstrated in a variety of species.

Object Permanence

The ability to understand that an object still exists when it is no longer visible is called object permanence. According to Piaget (1954), if a child sees an object disappear behind a screen or into a container, and by reaching behind the screen or into the container the child indicates an understanding that the object is still there, it is referred to as succeeding at a *visual displacement* of the object. If the container into which an object has been placed is moved and the child indicates an understanding that the object is still in the container, it is referred to as succeeding at an *invisible displacement* of the object.

Casual observation of pet dogs and cats suggests that they too can search appropriately for an object that they can no longer see (e.g., their behavior when a ball rolls under a couch; controlled experiments have supported this observation; Triana & Pasnak, 1981). In one version of the invisible displacement task, after the object is placed in a container and the container is moved behind a screen, the container reappears but the object is no longer in it. Children are considered successful if they look for the object behind the screen because it suggests that they infer the location of the missing object. Although several species of animal, including gorillas (Natale et al., 1986), chimpanzees (Call, 2001), Eurasian jays (Zucca, Milos, & Vallortigara, 2007), and dogs (Gagnon & Doré, 1992), have been shown to search accurately for visibly displaced objects, using this procedure, the evidence for accurate searches for invisibly displaced objects has been less conclusive.

A simpler version of the invisible displacement task involves a rotating beam with a container at either end. Evidence for high accuracy on the invisible displacement task by dogs has been found when, after placing an object in one container, the beam is rotated 90°, but when the beam is rotated 180° the dogs have generally failed to choose correctly (Miller et al., 2009). The 180° rotation appears to be more difficult because after the rotation the apparatus appears exactly as it was before the rotation. There is also some evidence that pigeons are able to track the invisible displacement of an object (Zentall & Raley, 2019).

Tool Use

One of the characteristics of modern humans is the use of tools to solve problems. A tool can be defined broadly as a device or implement (especially one held in the hand) used to carry out a particular function. Thus, it is relatively easy to demonstrate tool use in an animal that has hands. Chimpanzees clearly use tools when they strip leaves from a stem and probe large termite mounds, attracting termites to the stem so that they can eat them when the stem is withdrawn from the mound (McGrew, 1992). And tool use is demonstrated when they break open a nut by collecting a flat rock on which they place the nut and hit it with a second handheld rock to obtain the edible part inside (Boesch & Boesch, 1990). Other uses of tools are not so direct and may be controversial. Certainly, when a sea otter swims to the surface holding a clam and a rock and opens the clam by hitting it against the rock placed on its chest (Hall & Schaller, 1964) it would qualify as tool use. However, when a seagull picks up a clam and drops the clam over rocks, does that qualify (Maron, 1982)?

A secondary question is how did the behavior come about? In the case of the clam-dropping gull, did that behavior come about by trial and error? Was the drop accidental and then it became part of the bird's repertoire because it was reinforced?

The spontaneous use of a tool has been found in an Asian elephant (Foerder et al., 2011). When offered food on a branch that was not within reach, without prior experience of climbing on boxes to obtain food, the elephant moved a box to a location where it could stand on the box and obtain the food with its trunk.

There is evidence that bottlenose dolphins use marine sponges as foraging tools (Krützen et al., 2014), and that behavior appears to be culturally transmitted because Mann and Sargeant (2003) found that sponge foraging appears to be passed down from mother to calf.

Tool use as a problem-solving strategy has also been studied in New Caledonian crows (Weir, Chappell, & Kacelnik, 2002), birds that naturally break off twigs and probe them into dead logs to force out insects and grubs. When brought into the laboratory, these birds quickly learn to use a bent piece of wire to hook the handle on a small bucket with food that has been placed at the bottom of a tall cylinder. More surprising, however, when the bent hook was not available, one bird spontaneously bent a straight wire into a hook to retrieve the food bucket. Thus, in this case, the bird manufactured the tool from material not generally available in nature, so the behavior could not easily be attributable to prior experience or genetically prepared behavior.

Transitive Inference

In its simplest form, the transitive inference task can be described as follows: If A is greater than B (A > B), and B is greater than C (B > C), then it can be inferred that A > C (where the letters A, B, and C represent arbitrary stimuli such as colors or shapes). It is assumed that a correct response on this relational learning task requires that an inference be made about the relation between A and C, a relation that can only be derived from the two original propositions involving B. To avoid potential problems with "endpoint effects" that could produce a spurious nonrelational solution (i.e., C is never said to be greater and A is always said to be greater), in experimental research, tasks generally involve four propositions – A > B, B > C, C > D, and D > E – and the test involves the choice between B and D, each of which during training had been sometimes greater and sometimes lesser.

When humans are tested for transitive inference, the use of language allows for the propositions to be completely relational. Relative size may be assigned to individuals identified only by name (e.g., given that Anne is taller than Betty, and Betty is taller than Carol, who is taller, Anne or Carol?). With animals, however, there is no way to present such relational propositions without also presenting the actual stimuli. And if the stimuli differ in observable value (e.g., size), then a correct response can be made without the need to make an inference.

McGonigle and Chalmers (1977) suggested that a nonverbal relational form of the task could be represented by simple simultaneous discriminations in which one stimulus is associated with reinforcement (+) and the other is not (–). A > B can be represented as A+ B–, B > C as B+ C–, and so on. With four propositions an animal would be exposed to A+ B–, B+ C–, C+ D–, and D+ E–. A is always positive and E is always negative but B and D, stimuli that were never paired during training, would share similar reinforcement histories. If animals order the stimuli from A is best to E is worst, then B should be preferred over D.

Findings consistent with transitive inference have been reported in research with species as diverse as chimpanzees (Gillan, 1981), rats (Davis, 1992), pigeons (Fersen et al., 1991) and even fish (Grosenick, Clement, & Fernald, 2007). Although some have argued that these results can be accounted for without postulating that an inference has been made (Couvillon & Bitterman, 1992; Fersen et al., 1991; Steirn, Weaver, & Zentall, 1995), transitive inference effects have been found when these presumably simpler mechanisms have been controlled (Lazareva & Wasserman, 2006; Weaver, Steirn, & Zentall, 1997). Recent evidence suggests, however, that the nonverbal transitive inference effect found with simple simultaneous discriminations may result from differential tendencies to reject the test stimuli acquired during training (see Galizio et al., 2017; Zentall, Peng, & Miles, in press).

Conservation

The conservation of liquid volume task, made popular as a test of cognitive development by Piaget (1952), was developed to test for the inference that if two liquid volumes are initially the same and one of the volumes is transformed by pouring it into a container of a different shape (following transformation, the heights of the liquids in the two containers are quite different), the volumes are still the same. Woodruff, Premack, and Kennel (1978) developed a nonverbal version of this task that they used to test for conservation in Sarah, a chimpanzee. Not only did Sarah indicate (by means of previously acquired use of tokens representing "same" and "different") that transformation of shape did not cause two like volumes to be different, but she also indicated that two dissimilar volumes continued to be different following a transformation that resulted in liquid levels of similar height. Furthermore, importantly, Sarah was unable to correctly judge the relative volume of the liquids if the transformation was made out of sight. Thus, correct responding required observation of the original state of the containers and the transformation. This series of experiments is particularly noteworthy for its careful control of possible extraneous variables.

Analogical Reasoning

Another example of reasoning by a chimpanzee, analogical reasoning, has been reported by Gillan, Premack, and Woodruff (1981). Sarah, the chimpanzee, was shown pictures of objects she had previously encountered in the relation: A is to B as A′ is to X, with a choice of B′ and C′ as a replacement for X. Sarah's reliance on the analogical relationship was tested by varying only the initial stimulus pair. Thus, on one trial she was presented with, for example, "lock" is to "key" as "paint can" is to "?" with a choice of "can opener" and "paint brush," and on another trial with "paper" is to "pencil" as "paint can" is to "?" with a choice of the same "can opener" and "paint brush." In the first case, Sarah selected the "can opener" (indicating something with which to open the paint can), in the second, the "paint brush" (indicating something with which to paint). Thus, at least one chimpanzee appears to understand and be able to use analogical reasoning.

Language

We are the only species to develop, on our own, the flexible form of communication based on arbitrary symbols that we call language. With training, however, other species may be able to acquire a rudimentary form of symbolic communication. One of the most widely reported and least understood lines of research in animal intelligence involves projects concerned with the acquisition of language by chimpanzees. The three best known of these projects are Gardner and Gardner's (1969) sign learning project (see also Patterson's [1978] work with a gorilla and Herman, Pack, and Morrel-Samuels' [1993] work with dolphins), Premack's (1976) token learning project, and the Rumbaughs' (see Rumbaugh, 1977) keyboard learning project.

Although these projects are identified by the nature of the responses required of their animals, they are better distinguished by differences in their conceptual approaches. The Gardners chose sign language because it is an accepted form of human language, and acquisition and mastery skills by a chimpanzee could be compared directly with those of humans by objective sign-trained observers unfamiliar with the animals. The use of tokens in Premack's research allowed for more careful control over the set of possible responses. Premack's research focused more on the *conceptual* nature of language, including such characteristics as same/different learning, negation, property of, and causality. The Rumbaughs' work with Austin and Sherman focused on the functional use of language in communication between conspecifics (Savage-Rumbaugh, 1984). For example, they established conditions in which solution of a problem by one chimpanzee required the production and reception by another chimpanzee of a list of symbols representing a request for a tool.

Whether the communication skills acquired by these chimpanzees qualify as language depends in part on how language is defined. Unfortunately, there is little agreement on the necessary and sufficient characteristics of language. Such a definition must be sufficiently liberal to include not only hearing-impaired humans who use sign language but also young children and many developmentally delayed adults who have restricted but functionally adequate language skills.

Taking the Perspective of Others

An organism can take the perspective of another when it demonstrates an understanding of what the other may know. For example, when Susan sees a hidden object moved to a second hidden location after Billy has left the room and Susan understands that Billy will probably look for the object in the first location rather than the second, we would say that Susan can take the perspective of Billy, or that she has a theory of mind, because she understands that Billy doesn't know that the object has been moved (see Frye, 1993). To demonstrate perspective-taking in an animal is a bit more complex because, in the absence of language, theory of mind must be inferred from other behavior.

Self-Recognition

Recognition of the similarity between ourselves and other humans would seem to be a prerequisite for perspective-taking. If we can recognize ourselves in a mirror, we can see that we are similar to others of our species. Gallup (1970) has shown that not only will chimpanzees exposed to a mirror use it for grooming, but if their face is marked while they are anesthetized, they will use the mirror to explore the mark visually and tactually (i.e., they pass the mark test). Furthermore, both prior experience with the mirror and the presence of the mirror following marking appear to be necessary for mark exploration to occur. Mirror-directed mark exploration appears to occur in other higher apes (orangutans and perhaps also in gorillas), however, evidence of self-recognition in monkeys has been mixed. Gallup and Suarez (1991) failed to find evidence for it, even with extensive mirror experience, but recent evidence indicates that they too show it under the right conditions (Rajala et al., 2010). There is also some evidence of self-recognition in dolphins (Reiss & Marino, 2001) elephants (Plotnik, de Waal, & Reiss, 2006) and magpies (Prior, Schwatz, & Güntürkün, 2008). Thus, self-recognition appears to occur in several species thought to have other cognitive skills. There is even suggestive evidence that ants have some degree of self-recognition, as they react quite differently to a mirror image of themselves than to the image of another member of the same species (Cammaerts & Cammaerts, 2015).

Imitation

A more direct form of perspective-taking involves the capacity to imitate another (Piaget, 1951), especially opaque imitation for which the observer cannot see itself perform the response (e.g., clasping one's hands behind one's back). But evidence for true imitative learning requires that one rule out (or control for) other sources of facilitated learning following observation (see Whiten & Ham, 1992; Zentall, 1996). A design that appears to control for artifactual sources of facilitated learning following observation is the two-action procedure based on a method developed by Dawson and Foss (1965). For example, imitation is said to occur if observers, exposed to a demonstrator performing a response in one of two topographically different ways, perform the response with the same topography as their demonstrator (Heyes & Dawson, 1990). Akins and Zentall (1996) trained Japanese quail demonstrators to either step on a treadle or peck the treadle for food reinforcement. When observer quail were exposed to one or the other demonstrator, they matched the behavior of their demonstrator with a high probability (see also Zentall, Sutton, & Sherburne, 1996, for similar evidence with pigeons). Furthermore, there is some evidence that pigeons can imitate a sequence of two responses: operating a treadle, by stepping or pecking, and pushing a screen, to the left or to the right (Nguyen, Klein, & Zentall, 2005).

Perhaps the most impressive example of animal imitation comes from a test of generalized imitative learning reported by Hayes and Hayes (1952) with a home-raised chimpanzee named Viki. Using a set of seventy gestures, Viki was trained to replicate each gesture when the experimenter said, “Do this.” More important, Viki also accurately performed ten novel arbitrary gestures when directed to with the “Do this” command (see also Custance, Whiten, & Bard, 1995). Recent evidence suggests that other species can also follow the do-as-I-do command, in particular dogs (Topal et al., 2006), parrots (Moore, 1992), and dolphins (Herman, 2002).

If Piaget’s reasoning is correct, the ability to imitate requires the ability to take the perspective of another. But children do not develop the ability to take the perspective of another until they are 5–7 years old, yet they are able to imitate others at a much earlier age. Furthermore, if pigeons and Japanese quail can imitate, it is unlikely that they do so by taking the perspective of the demonstrator, at least not in the sense that Piaget implied. Thus, although cognitively interesting, imitation may not provide evidence for the kind of cognitive behavior implied by perspective-taking.

Another mechanism by which imitation may occur is cross-modal matching (see, e.g., Mitchell, 1997). The idea is that the brain of some animals is wired such that seeing a particular response being made by a conspecific stimulates the motor neurons associated with the same response in the observer. There is, in fact, some evidence of mirror neurons that fire both when monkeys make a response as well as when they see that response made by a human or another monkey (Rizzolatti et al., 1996). It is very likely that brain mechanisms involving mirror neurons play a role in imitation, but they cannot readily account for opaque imitation in which there is no visual match between what is seen and what is felt. Such a neural connection has not yet been found.

An interesting finding with children not shown by chimpanzees is “over-imitation” (Horner & Whiten, 2005). When chimpanzees are shown a method for obtaining a treat from a box, they tend to omit aspects of the demonstration that are irrelevant. Children, however, tend to imitate the irrelevant aspects of the demonstration as well. That is, they imitate more than is necessary. This finding has led to the interesting speculation that humans and other animals are motivated by different aspects of the task. Whereas chimpanzees are motivated by the pragmatics of obtaining the treat, children learn from others with an added focus on the social consequences of social learning (Nielsen & Haun, 2016; Over & Carpenter, 2012).

Animal Culture

When a particular behavior is imitated by all members of a group but not by other groups, some researchers have taken it as evidence that the species has a form of culture (see Laland & Galef, 2009), but this question depends in part on how one defines culture. If one defines culture as an anthropologist might, characterized by a group having socially learned laws, ethics, rituals, religion, and morality, then no group of nonhuman animals has culture. If, however, one defines culture as the transmission of innovations among members of a group (some have argued that tradition may be a less controversial term; Laland & Galef, 2009), then animals may have some characteristics of culture. Much of the evidence for culture in animals comes from animals living in natural settings in which the members of one group exhibit a particular behavior, whereas those of other nearby groups do not (e.g., grooming posture in chimpanzees; McGrew & Tutin, 1978). The problem is, if group differences in behavior are to be attributed to culture, it must be clearly shown that they do not result either from genetic differences between the groups or from environmental differences that could have encouraged one group to develop the novel behavior by individual learning (Schaik, 2012).

Better controlled studies can be carried out in the laboratory, where one can control for the environmental conditions and also for genetic differences by randomly assigning animals to groups (Dally, Clayton, & Emery, 2008). And perhaps the best example of a simulation of animal culture in the laboratory is the serial transmission of food preference among rats (Galef & Whiskin, 1998).

Theory of Mind

A version of the child's game with a hidden object described in the section Taking the Perspective of Others (Frye, 1993) was attempted by Povinelli, Nelson, and Boysen (1990). Chimpanzees were trained to select a box toward which a trainer was pointing to receive a reward. When they were tested with two trainers (who were pointing at different boxes) – one who had been present when the box was baited (the "knower") and the other who had been absent (the "guesser") – the chimpanzees chose the box indicated by the "knower" over that indicated by the "guesser." But as Heyes (1998) has noted, in this and other similar procedures (involving, for example, a "guesser" with a bag over his head), the preference for the "knower" did not show up on early trials and the number of test trials was sufficient for the chimpanzees to have learned to use the "knower's" behavior as a cue.

A different approach to theory of mind focused on the natural competitiveness and dominance hierarchy of chimpanzees (e.g., Hare, Call, & Tomasello, 2001). They found that when a subordinate chimpanzee could observe that a dominant chimpanzee could see where food was hidden, it tended to avoid that location. But the subordinate was less likely to avoid a location when the dominant could not have seen where the food was hidden. Although these and related experiments provide the best evidence to date for theory of mind in animals, it may be that cues provided by the dominant chimpanzee played a role in the results. That is, if the dominant chimpanzee was staring at the location where it saw food hidden, it may have inhibited the subordinate from approaching that location.

Recently, animal researchers have used techniques borrowed from developmental psychologists to study false belief, a version of theory of mind, in young children without the need for a verbal response. Southgate, Senju, and Csibra (2007) found that children often looked at a location where they expected an actor to search for an object that the child, but not the actor, had seen moved to a different location. Krupenye and colleagues (2016) used this procedure to show that apes appeared to understand that an actor would look in a place where the hidden object had been but no longer was. Thus, there is some evidence that apes understand that others are autonomous agents that may not have the same information as themselves.

Deception

If an animal can purposefully deceive another, one could argue that it must be able to take the perspective of the other. The broken wing act of the killdeer can be considered a functional act of deception (Ristau, 1991) but it is very likely that this act of luring a potential predator away from its nest is genetically predisposed and may not represent purposeful deception.

Certainly, functional deception can be trained. Woodruff and Premack (1979) trained chimpanzees to point to the container that held food in order to receive the food. The chimpanzees then learned that one trainer would give them the food for pointing to the correct container, whereas the other would allow them to have the food only if they pointed to an incorrect container. Although the chimpanzees learned to respond appropriately, there was no indication that they *intended* to deceive the trainer (Dennett, 1983).

One can find anecdotes in the literature suggestive of intentional deception (e.g., Heyes, 1998), but the problem with the attribution of deception is that intentionality must be inferred from behavior, and intentionality is particularly difficult to assess in a nonverbal organism. Perhaps the best example of deception by animals that appears to have aspects of intentionality comes from pilfer-avoidance behavior of certain food-caching birds. Western scrub jays will go to great lengths to avoid caching in the presence of other scrub jays, and when there is another scrub jay present that might have seen them caching, they will often wait until the other bird has left and then move their cached food to a different location (Dally, Emery, & Clayton, 2005). Whether this deception is purposeful is difficult to determine but the various forms that it takes (Dally, Emery, & Clayton, 2004) suggests that it may be.

Cooperation and Altruism

Cooperation and altruism are special cases of intelligent behavior because they represent a form of social behavior for which the actions of the organism have implications for the well-being of another. Although true cooperation and altruism are closely related to theory of mind, many forms of these behaviors (e.g., the cooperation among dogs hunting in a pack, and maternal behavior) are strongly biologically predisposed and so cognitive accounts are unnecessary. Other cases of cooperation can more parsimoniously be interpreted as the use of another animal as a discriminative stimulus. Skinner (1962), for example, trained pigeons to “hunt” on each trial for the response location (randomly designated) to which a response would be reinforced. He then placed two pigeons side by side and added the contingency that on a given trial, the same two response locations were correct and must be pecked simultaneously to obtain a reward. The pigeons readily adjusted to the new contingency and often got fed, but their functional cooperation can be explained as the use of the movement of one pigeon toward a response location as a discriminative stimulus for the other pigeon to peck the corresponding location (see Tan & Hackenberg, 2016, for similar results with rats).

A better example of altruistic behavior was described by Bartal, Decety, and Mason (2011), who found that rats would work to release a cage-mate (but not an inanimate object) from a restraint. But in studies of altruism it is critical to ensure that the presumed altruist does not receive some tangible reinforcement for its behavior, and Hackenberg (2017) has found that the opportunity to interact with the cage-mate may provide sufficient social reinforcement to account for the releasing behavior.

Recently it has been proposed that rats actually show reciprocal altruism, choosing to feed a rat that has recently provided them with food (Rutte & Taborsky, 2008), but before accepting such results one needs to be sure that simpler conditioning effects are not responsible for the observed behavior (see Zentall, 2016).

Examples of altruistic behavior based on variants of parental behavior (e.g., adoption of an unrelated offspring) can be explained more parsimoniously in terms of “errors” in biologically predisposed behavior. Even altruistic acts such as those that occur between unrelated humans in wartime may be based on biological predispositions that evolved in hunter-gatherer times as a form of kin selection (the tendency that genes predispose the bodies that they are in to look out for themselves and copies of themselves in others – i.e., kin). In the case of wartime bravery, the closeness of the military unit may mimic the relatedness of a hunter-gatherer hunting party. Furthermore, one could argue that although a certain level of intelligence may be required to produce true, cognitively based cooperation and altruism in humans, considering the range of individual differences in altruism and cooperation among humans, intelligence is certainly not predictive of either. Theory of mind in animals is a relatively new area of research that is fraught with problems of interpretation; however, clever techniques for assessing what animals know (e.g., Gallup, 1970; Hare et al., 2001) promise to get us closer to the goal of understanding the relation between the cognitive abilities of humans and those of other animals.

What Animals Can Tell Us about Human Reasoning

Cognitive Dissonance

I have saved for last the discussion of four programs of research directed to similarities between the behavior of humans and that of other animals because they have important implications for how we interpret human behavior. The first has to do with a phenomenon extensively studied in humans called cognitive dissonance (Festinger, 1957). Cognitive dissonance is the discomfort that comes when there is a discrepancy between one's beliefs and one's behavior. For example, if one believes that one should tell the truth, one is likely to feel dissonance on occasions when one fails to do so. That dissonance may be resolved by deciding that there are some conditions under which lying is appropriate or that the person lied to may have deserved it. Cognitive dissonance presumably comes about because of a need to be consistent or to avoid being labeled a hypocrite. Does this represent a kind of social intelligence? And if so, would nonhuman animals show a similar effect? But how would one go about asking this question of animals?

One approach involves a version of cognitive dissonance called justification of effort (Aronson & Mills, 1959). In their study, undergraduates who underwent an unpleasant initiation to become part of a group reported that they were more inclined to want to join the group than those who underwent a less unpleasant initiation. It is assumed that those individuals gave more value to membership in the group to justify undergoing the unpleasant initiation.

The justification of effort design allows for a direct test of cognitive dissonance in animals. For example, if on some trials, a pigeon has to work hard to receive Signal A that says food is coming but on other trials, the pigeon does not have to work hard to receive Signal B that says the same food is coming, will the pigeon show a preference for Signal A over Signal B? Several studies have shown that they will (e.g., Clement et al., 2000; Kacelnik & Marsh, 2002). But is this cognitive dissonance? Do animals need to justify to themselves why they worked harder for one signal than the other?

Alternatively, we have suggested that this choice behavior results from the contrast between the relatively negative emotional state of the organism at the end of the effort and on presentation of the signal for reinforcement (Zentall & Singer, 2007). That difference would be greater when more effort is involved. Thus, the subjective value of the signal for reinforcement might be judged to be greater. Under these conditions, contrast provides a more parsimonious account of the pigeons' choice behavior. Could contrast also be involved when similar behavior is shown by humans? It would be worth studying this possibility.

Suboptimal Choice: Gambling Behavior

Humans often gamble (e.g., they play the lottery) even when the odds against winning are very high. This form of gambling behavior may be attributable to an inaccurate assessment of the probability of winning, perhaps resulting in part from public announcements of the winners but not the losers (an availability heuristic). Would animals show a similar kind of suboptimal gambling behavior? According to optimal foraging theory, they should not because such inappropriate behavior should have been selected against by evolution. Furthermore, if the choice is to have any meaning for the animal, it would have to have experienced the probability associated with winning (reinforcement) and that should increase the likelihood that the animal would be able to assess the probability of winning and losing. However, we have recently found conditions under which pigeons will prefer 20 percent reinforcement over 50 percent reinforcement (Stagner & Zentall, 2010) in the context of a gamble. The procedure was as follows: If the pigeon chose the left alternative, 20 percent of the time a red stimulus appeared and was followed by food ten seconds later. The remainder of the time it chose the left alternative, a green stimulus appeared which was never followed by food. Thus, food appeared 20 percent of the time for the choice of left alternative. If the pigeon chose the right alternative, whatever stimulus appeared, it was followed by food 50 percent of the time. Curiously, the pigeons strongly preferred the left alternative and they did so in spite of the fact that they would have gotten two and a half times more food for choosing the right alternative. The results suggest that the probability of getting the cue for reinforcement was not important, only the predictive value of that stimulus. This prediction was recently confirmed when pigeons were given a choice between a 50 percent chance and a 100 percent chance of receiving a perfect predictor of food (Smith & Zentall, 2016). As predicted, although one of those stimuli occurred twice as often as the other, because both stimuli were perfect predictors of food, the pigeons were indifferent between the alternatives. Recent research suggests that the positive contrast between the expected probability of reinforcement (e.g., 50% in the example just given) and the actual probability of reinforcement when the signal for reinforcement appears may provide an added incentive to choose the sub-optimal alternative (Case & Zentall, 2018).

Such results suggest that gambling behavior is likely to have a simple biological basis and although social and cognitive factors may contribute to human gambling behavior, the underlying mechanism is likely to be present in other animals. A more nearly complete analysis of the mechanisms responsible for this irrational behavior and its relation to human gambling will have to wait for further research, but at this point it is clear that pigeons are no more appropriate in their choice behavior than are humans.

Sunk Cost

The sunk cost effect occurs when one allows an amount of money, time, or other resource already invested in an activity to affect one's decision to invest more resources. For example, one may sit through a film that one does not like because to leave would be to waste one's investment of the price of the ticket. But there is no way to recoup that already expended cost of the ticket and by continuing to sit through the film, one is spending additional resources – one's time that could be better spent in other activities. Similarly, although economists generally warn against it, one may choose to continue with a failing business because of the past investment one has made in it, rather than in the likelihood of future success.

The sunk cost effect comes under the general rubric of prospect theory (Kahneman & Twersky, 1979) which suggests that humans will take greater risks to avoid a loss than to obtain a gain. The question is to what extent this behavior stems from the cultural beliefs that one should avoid wasting resources and that one should complete a task that one has started. If one could show that animals, such as pigeons, show similar behavior, it would suggest that sunk cost is a general phenomenon that has basic behavioral origins.

In fact, evidence for sunk cost has been found in pigeons (Navarro & Fantino, 2005; Pattison, Watanabe, & Zentall, 2012). For example, pigeons learn that pecking a green light requires thirty pecks, whereas pecking a red light requires only ten pecks. They then learn that after pecking the green light a number of times (which varies from trial to trial), they are able to choose to continue with the green light (to complete the thirty pecks) or switch to the red light for which ten pecks are always required. Results indicate that pigeons often choose to continue with the green light, even when twenty more pecks are required (see also, Magalhães & White, 2013). Thus, pigeons show a sunk cost effect that is very similar to that shown by humans. Why pigeons show the sunk cost effect is not clear. One can speculate that it arises from the fact that in nature switching to a different foraging patch often involves uncertainty, some travel time, and possibly increased danger. In any event, one can conclude that culture and language are not necessary components of the effect.

When Less Is More

When humans are asked to judge the value of a set of objects of excellent quality, they often give it higher value than those same objects with the addition of objects of lesser quality (Hsee, 1998). This effect has been found when humans are asked to judge the value of sets of dishes or sets of baseball cards (Hsee, 1998), and also when academics are asked to judge the quality of a curriculum vita (CV; Hayes, 1983). For example, a CV with three publications in excellent journals is judged better than a CV with the same three publications in excellent journals, plus six more in lesser-quality journals. This bias is an example of the affect heuristic in which humans appear to judge the average quality of a set of objects to determine the value of the set, rather than judge the value of the total number of objects in the set. The phenomenon has become known as a *less is more* or *less is better* effect (Hsee, 1998).

Recent evidence suggests that even pigeons are susceptible to this bias. For example, pigeons will work for dried peas and dried milo seeds but when given a choice between the two, they prefer the peas. However, when they are given a choice between a pea and a pea together with a milo seed, they prefer the pea alone (Zentall et al., 2014). Apparently, the pigeons too are averaging the high-quality pea with the lower-quality milo seed and value the pair less than the pea by itself. Similar results have been found with monkeys (Kralik et al., 2012) and dogs (Pattison & Zentall, 2014). The basis of this bias may originate in the need to make rapid decisions, presumably because of intense competition from conspecifics and the possibility of predation, and they use this heuristic in the laboratory even when speed is not a factor, although level of motivation may also play a role (Zentall et al., 2014). Once again, the fact that other animals show this suboptimal choice indicates that the bias is probably not dependent on human cultural influence.

Conclusions

The broad range of positive research findings that have come from investigating the cognitive abilities of animals suggests that many of the "special capacities" attributed to humans may be more quantitative than qualitative. In the case of many cognitive learning tasks, once we learn how to ask the question appropriately (i.e., in a way that is accommodating to the animal), we may be surprised with the capacity of animals to use complex relations.

In evaluating the animal (and human) intelligence literature, we should be sensitive to both overestimation of capacity (what appears to be higher-level functioning in animals that can be accounted for more parsimoniously at a lower level; see Zentall, 1993) and underestimation of capacity (our bias to present animals with tasks convenient to our human sensory, response, and motivational systems). Underestimation can also come from the difficulty in providing animals with task instructions as can easily be done with humans (see Zentall, 1997). The accurate assessment of animal intelligence will require vigilance, on the one hand, to evaluate cognitive functioning against simpler accounts and, on the other hand, to determine the conditions that will maximally elicit the animal's cognitive capacity.

References

Akins, C., & Zentall, T. R. (1996). Evidence for true imitative learning in Japanese quail. *Journal of Comparative Psychology*, *110*, 316–320.

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, *59*, 177–181.

Babb, S. J., & Crystal, J. D. (2006). Episodic-like memory in the rat. *Current Biology*, *16*, 1317–1321.

Bartal, I. B.-A., Decety, J., & Mason, P. (2011). Helping a cagemate in need: Empathy and pro-social behavior in rats. *Science*, *334*, 1427–1430.

Bitterman, M. E. (1975). The comparative analysis of learning. *Science*, *188*, 699–709.

Bitterman, M. E., & Mackintosh, N. J. (1969). Habit reversal and probability learning: Rats, birds, and fish. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 163–185). New York: Academic Press.

Boesch, C., & Boesch, H. (1990). Tools use and tool making in wild chimpanzees. *Folia Primatologica*, *54*, 86–99.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, *103*, 23–31.

Call, J. (2001). Object permanence in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and children (*Homo sapiens*). *Journal of Comparative Psychology*, *115*, 159–171.

Cammaerts, M. C., & Cammaerts, R. (2015). Are ants (*Hymenoptera, Formicidae*) capable of self recognition? *Journal of Sciences*, *5*, 521–532.

Capaldi, E. J. (1993). Animal number abilities: Implications for a hierarchical approach to instrumental learning. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence* (pp. 191–209). Hillsdale, NJ: Erlbaum.

Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 3–17.

Case, J. P., & Zentall, T. R. (2018). Suboptimal choice in pigeons: Does the predictive value of the conditioned reinforcer alone determine choice? *Animal Cognition*, *22*, 81–87.

Chapuis, N., & Varlet, C. (1987). Short cuts by dogs in natural surroundings. *Quarterly Journal of Experimental Psychology*, *39*, 49–64.

Clayton, N. S., & Dickinson, A. (1999). Scrub jays (*Aphelocoma coerulescens*) remember the relative time of caching as well as the location and content of their caches. *Journal of Comparative Psychology*, *113*, 403–416.

Clement, T. S., Feltus, J., Kaiser, D. H., & Zentall, T. R. (2000). "Work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin and Review*, *7*, 100–106.

Collette, T. S., & Graham, P. (2004). Animal navigation: Path integration, visual landmarks and cognitive maps. *Current Biology*, *14*, 475–477.

Couvillon, P. A., & Bitterman, M. E. (1992). A conventional conditioning analysis of "transitive inference" in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*, 308–310.

Custance, D. M., Whiten, A., & Bard, K. A. (1995). Can young chimpanzees imitate arbitrary actions? Hayes and Hayes (1952) revisited. *Behaviour*, *132*, 837–859.

Dally, J. M., Clayton, N. S., & Emery, N. J. (2008). Social influences on foraging by rooks (*Corvus frugilegus*). *Behaviour*, *145*, 1101–1124.

Dally, J. M., Emery, N. J., & Clayton, N. S. (2004). Cache protection strategies by western scrub-jays (*Aphelocoma californica*): Hiding food in the shade. *Proceedings of the Royal Society B: Biological Sciences*, *271*, S387–S390.

Dally, J. M., Emery, N. J., & Clayton, N. S. (2005). Cache protection strategies by western scrub-jays (*Aphelocoma californica*): Implications for social cognition. *Animal Behaviour*, *70*, 1251–1263.

Davis, H. (1992). Transitive inference in rats (*Rattus norvegicus*). *Journal of Comparative Psychology*, *106*, 342–349.

Davis, H., & Memmott, J. (1982). Counting behavior in animals: A critical evaluation. *Psychological Bulletin*, *92*, 547–571.

Dawkins, R. (1976). *The selfish gene*. New York: Oxford University Press.

Dawson, B. V., & Foss, B. M. (1965). Observational learning in budgerigars. *Animal Behaviour*, *13*, 470–474.

Dennett, D. C. (1983). Intentional systems in cognitive ecology: The "panglossian paradigm" defended. *Behavioral and Brain Sciences*, *6*, 343–355.

Dorrance, B. R., Kaiser, D. H., & Zentall, T. R. (2000). Event duration discrimination by pigeons: The choose-short effect may result from retention-test novelty. *Animal Learning and Behavior*, *28*, 344–353.

Edwards, C. A., Jagielo, J. A., Zentall, T. R., & Hogan, D. E. (1982). Acquired equivalence and distinctiveness in matching-to-sample by pigeons: Mediation by reinforcer-specific expectancies. *Journal of Experimental Psychology: Animal Behavior Processes*, *8*, 244–259.

Fersen, L. V., Wynne, C. D. L., Delius, J. D., & Staddon, J. E. R. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 334–341.

Festinger, L (1957). *A theory of cognitive dissonance*. Stanford: Stanford University Press.

Foerder, P., Galloway, M., Barthel, T., Moore, D. E. III, & Reiss, D. (2011). Insightful problem solving in an Asian elephant. *PLoS One*, *6*(8), e23251. https://doi.org/10.1371/journal.pone.0023251

Frye, D. (1993). Causes and precursors of children's theory of mind. In D. F. Hay & A. Angold (Eds.), *Precursors and causes of development and psychopathology*. Chichester, UK: Wiley.

Gagnon, S., & Doré, F. Y. (1992). Search behavior in various breeds of adult dogs (*Canis familiaris*): Object permanence and olfactory cues. *Journal of Comparative Psychology*, *106*, 58–68.

Galef, B. G., Jr. (1988). Imitation in animals: History, definition, and interpretation of data from the psychological laboratory. In T. R. Zentall & B. G. Galef, Jr. (Eds.), *Social learning: Psychological and biological perspectives* (pp. 3–28). Hillsdale, NJ: Erlbaum.

Galef, B. G., Jr., & Whiskin, E. E. (1998). Determinants of the longevity of socially learned food preferences of Norway rats. *Animal Behaviour*, *55*, 967–975.

Galizio, A., Doughty, A. H., Williams, D. C., & Saunders, K. J. (2017). Understanding behavior under nonverbal transitive-inference procedures: Stimulus-control-topography analyses. *Behavioural Processes*, *140*, 202–215.

Gallup, G. G. (1970). Chimpanzees self-recognition, *Science*, *167*, 86–87.

Gallup, G. G., & Suarez, S. D. (1991). Social responding to mirrors in rhesus monkeys: Effects of temporary mirror removal. *Journal of Comparative Psychology*, *105*, 376–379.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science*, *4*, 123–124.

Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, *165*, 664–672.

Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*, 150–164.

Gillan, D. J., Premack, D., & Woodruff, G. (1981). Reasoning in the chimpanzee: I. Analogical reasoning. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*, 1–17.

Grant, D. S. (1981). Stimulus control of information processing in pigeon short-term memory. *Learning and Motivation*, *12*, 19–39.

Grosenick, L., Clement, T. S., & Fernald, R. D. (2007). Fish can infer social rank by observation alone. *Nature*, *445*, 429–432.

Hackenberg, T. D. (2017). To free or not to free: Determinants of social release in rats. Paper presented at the meeting of the Society for the Quantitative Analysis of Behavior, Denver, CO, May 25.

Hall, K. R. L., & Schaller, G. B. (1964). Tool-using behavior of the California sea otter. *Journal of Mammalogy*, *45*, 287–298.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*, 139–151.

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*, 51–65.

Hayes, S. C. (1983). When more is less: Quantity versus quality of publications in the evaluation of vitae. *American Psychologist*, *38*, 1398–1400.

Hayes, K. J., & Hayes, C. (1952). Imitation in a home-raised chimpanzee. *Journal of Comparative and Physiological Psychology*, *45*, 450–459.

Herman, L. M. (2002). Vocal, social, and self-imitation by bottlenosed dolphins. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 63–108). Cambridge, MA: MIT Press.

Herman, L. M., Pack, A. A., & Morrel-Samuels, P. (1993). Representational and conceptual skills of dolphins. In H. L. Roitblat, L. M. Herman, & P. E. Nachtigall (Eds.), *Language and communication: Comparative perspectives* (pp. 403–442). Hillsdale, NJ: Erlbaum.

Herrnstein, R. J., & deVilliers, P. A. (1980). Fish as a natural category for people and pigeons. *Psychology of Learning and Motivation*, *14*, 59–95.

Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science*, *146*, 549–551.

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*, 285–301.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, *21*, 101–134.

Heyes, C. M., & Dawson, G. R. (1990). A demonstration of observational learning in rats using a bidirectional control. *Quarterly Journal of Experimental Psychology*, *42B*, 59–71.

Honig, W. K., & Thompson, R. K. R. (1982). Retrospective and prospective processing in animal working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (vol. 16, pp. 239–283). Orlando, FL: Academic Press.

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, *8*, 164–181.

Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, *11*, 107–121.

Hull, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.

Kacelnik, A., & Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behaviour*, *63*, 245–250.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Kelly, R., & Grant, D. S. (2001). A differential outcomes effect using biologically neutral outcomes in delayed matching-to-sample with pigeons. *Quarterly Journal of Experimental Psychology*, *54B*, 69–79.

Kralik, J. D., Xu, E. R., Knight, E. J., Khan, S. A., & Levine, J. W. (2012). When less is more: Evolutionary origins of the affect heuristic. *PLoS One*, *7*, e46240. https://doi.org/10.1371/ journal.pone.0046240

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*, 110–114.

Krützen, M., Kreicker, S., MacLeod, C. D., Learmonth, J., Kopps, A. M., Walsham, P., et al. (2014). Cultural transmission of tool use by Indo-Pacific bottlenose dolphins (*Tursiops sp.*) provides access to a novel foraging niche. *Proceedings of the Royal Society B: Biological Sciences 281*, 20140374.

Kuan, L.-A., & Colwill, R. (1997). Demonstration of a socially transmitted taste aversion in the rat. *Psychonomic Bulletin and Review*, *4*, 374–377.

Laland, K. N., & Galef, B. G., Jr. (Eds.) (2009). *The question of animal culture*. London: Harvard University Press.

Lazareva, O. F., & Wasserman, E. A. (2006). Effect of stimulus orderability and reinforcement history on transitive responding in pigeons. *Behavioural Processes*, *72*, 161–172.

Lipp, H.-P., Vyssotski, A. L., Wolfer, D. P., Renaudineau, S., Savini, M., Tröster, G., et al. (2004). Pigeon homing along highways and exits. *Current Biology*, *14*, 1239–1249.

Mackintosh, N. J. (1969). Comparative studies of reversal and probability learning: Rats, birds, and fish. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning* (pp. 137–162). New York: Academic Press.

Magalhães, P., & White, K. G. (2013). Sunk cost and work ethic effects reflect suboptimal choice between different work requirements. *Behavioural Processes*, *94*, 55–59.

Mann, J., & Sargeant, B. (2003). Like mother, like calf: the ontogeny of foraging traditions in wild Indian ocean bottlenose dolphins (*Tursiops* sp.). In D. Fragaszy & S. Perry (Eds.), *The biology of traditions* (pp. 236–266). Cambridge, UK: Cambridge University Press.

Maron, J. L. (1982). Shell-dropping behavior of western gulls (*Larus occidentalis*). *The Auk*, *99*, 565–569.

McGonigle, B. O., & Chalmers, M. (1977). Are monkeys logical? *Nature*, *267*, 694–696.

McGrew, W. C. (1992). *Chimpanzee material culture: Implications for human evolution*. Cambridge, UK: Cambridge University Press.

McGrew, W. C., & Tutin, C. E. G. (1978). Evidence for a social custom in wild chimpanzees? *Man*, *13*, 234–251.

Meyer, D. R. (1971). Habits and concepts of monkeys. In L. E. Jarrard (Ed.), *Cognitive processes of nonhuman primates* (pp. 83–102). New York: Academic Press.

Miller, H. C., Friedrich, A. M., Narkavic, R. J., & Zentall, T. R. (2009). A differential outcomes effect using hedonically-nondifferential outcomes with delayed matching-to-sample by pigeons. *Learning and Behavior*, *37*, 161–166.

Miller, H. C., Gipson, C. D., Vaughan, A., Rayburn-Reeves, R., & Zentall, T. R. (2009). Object permanence in dogs: Invisible displacement in a rotation task. *Psychonomic Bulletin and Review*, *16*, 150–155.

Mitchell, R. W. (1997). A comparison of the self-awareness and kinesthetic-visual matching theories of self-recognition: Autistic children and others. *New York Academy of Sciences*, *818*, 39–62.

Moore, B. R. (1992). Avian movement imitation and a new form of mimicry: Tracing the evolution of a complex form of learning. *Behaviour*, *122*, 231–263.

Morgan, C. L. (1894). *An introduction to comparative psychology*. London: Scott.

Naqshbandi, M., & Roberts, W. A. (2006). Anticipation of future events in squirrel monkeys (*Saimiri sciureus*) and rats (*Rattus norvegicus*): Tests of the Bischof-Kohler hypothesis. *Journal of Comparative Psychology*, *120*, 345–357.

Natale, F., Antinucci, F., Spinozzi, F., & Poti', P. (1986). Stage 6 object permanence in nonhuman primate cognition: A comparison between gorilla (*Gorilla gorilla*) and Japanese macaque (*Macaca fuscata*). *Journal of Comparative Psychology*, *100*, 335–339.

Navarro, A. D., & Fantino, E. (2005). The sunk cost effect in pigeons and humans. *Journal of the Experimental Analysis of Behavior*, *83*, 1–13.

Nguyen, N. H., Klein, E. D., & Zentall, T. R. (2005). Imitation of two-action sequences by pigeons. *Psychonomic Bulletin and Review*, *12*, 514–518.

Nielsen, M., & Haun, D. (2016). Why developmental psychology is incomplete without comparative and crosscultural perspectives. *Philosophical Transactions of the Royal Society: B*, *371*, 20150071.

Over, H., & Carpenter, M. (2012). Putting the social into social learning: Explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*, *126*, 182–192.

Patterson, F. G. (1978). The gestures of a gorilla: Language acquisition in another pongid. *Brain and Language*, *5*, 72–97.

Pattison, K. F., & Zentall, T. R. (2014). Suboptimal choice by dogs: When less is better than more. *Animal Cognition*, *17*, 1019–1022.
Pattison, K. F., Zentall, T. R., & Watanabe, S. (2012). Sunk cost: Pigeons (*Columba livia*) too show bias to complete a task rather than shift to another. *Journal of Comparative Psychology*, *126*, 1–9.
Pepperberg, I. M. (1987). Interspecies communication: A tool for assessing conceptual abilities in an African grey parrot. In G. Green-berg & E. Tobach (Eds.), *Language, cognition, and consciousness: Integrative levels* (pp. 31–56). Hillsdale, NJ: Erlbaum.
Peterson, G. B. (1984). How expectancies guide behavior. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 135–148). Hillsdale, NJ: Erlbaum.
Peterson, G. B., Wheeler, R. L., & Trapold, M. A. (1980). Enhancement of pigeons' conditional discrimination performance by expectancies of reinforcement and nonreinforcement. *Animal Learning and Behavior*, *8*, 22–30.
Piaget, J. (1951). *Play, dreams, and imitation in childhood*. New York: W. W. Norton.
Piaget, J. (1952). *The child's concept of number*. New York: W. W. Norton.
Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
Plotnik, J. M., de Waal, F. B. M., & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences*, *103*, 17053–17057.
Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1990). Inferences about guessing and knowing by chimpanzees. *Journal of Comparative Psychology*, *104*, 203–210.
Premack, D. (1976). *Intelligence in ape and man*. Hillsdale, NJ: Erlbaum.
Prior, H., Schwatz, A., & Güntürkün, O. (2008). Mirror-induced behavior in the magpie (*Pica pica*): Evidence of self-recognition. *PLoS Biology*, *6*(8), e202. https://doi.org/10.1371/journal. pbio.0060202
Raby, C. R., Alexis, D. M., Dickinson, A., & Clayton, N. S. (2007). Empirical evaluation of mental time travel. *Behavioral Brain Sciences*, *30*, 330–331.
Rajala, A. Z., Reininger, K. R., Lancaster, K. M., & Populin, L. C. (2010). Rhesus monkeys (*Macaca mulatta*) do recognize themselves in the mirror: Implications for the evolution of self-recognition. *PLoS One*, *5*(9), e12865. https://doi.org/10.1371/journal.pone.0012865
Reiss, D., & Marino, L. (2001). Self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences*, *98*, 5937–5942.
Riley, D. A. (1968). *Discrimination learning*. Boston: Allyn & Bacon.
Ristau, C. A. (1991). Aspects of the cognitive ethology of an injury-feigning bird, the piping plover. In C. Ristau (Ed.), *Comparative ethology: The minds of other animals* (pp. 79–89). Hillsdale, NJ: Lawrence Erlbaum.
Rizzolatti, G., Fadiga, L, Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.
Roberts, W. A. (2002). Are animals stuck in time? *Psychological Bulletin*, *128*, 473–489.
Roberts, W. A., & Grant, D. S. (1976). Studies of short-term memory in the pigeon using the delayed matching-to-sample procedure. In D. L. Medin, W. A. Roberts, & R. T. Davis (Eds.), *Processes of animal memory* (pp. 79–112). Hillsdale, NJ: Erlbaum.
Roper, K. L., Kaiser, D. H., & Zentall, T. R. (1995). Directed forgetting in pigeons: The role of alternative memories in the effectiveness of forget cues. *Animal Learning and Behavior*, *23*, 280–285.
Roper, K. L., & Zentall, T. R. (1993). Directed forgetting in animals. *Psychological Bulletin*, *113*, 513–532.

Rumbaugh, D. M. (Ed.) (1977). *Language learning by a chimpanzee: The Lana project*. New York: Academic Press.

Rutte, C., & Taborsky, M. (2008). The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): Direct vs generalised reciprocity. *Behavioral Ecology and Sociobiology*, *62*(4), 499–505.

Savage-Rumbaugh, E. S. (1984). Acquisition of functional symbol use in apes and children. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 291–310). Hillsdale, NJ: Erlbaum.

Schaik, C. P. van (2012). Animal culture: Chimpanzee conformity? *Current Biology*, *22*, R402–R404.

Sherburne, L. M., Zentall, T. R., & Kaiser, D. H. (1998). Timing in pigeons: The choose-short effect may result from "confusion" between delay and intertrial intervals. *Psychonomic Bulletin and Review*, *5*, 516–522.

Singer, R. A., Abroms, B. D., & Zentall, T. R. (2007). Formation of a simple cognitive map by rats. *International Journal of Comparative Psychology*, *19*, 417–425.

Singer, R. A., & Zentall, T. R. (2007). Pigeons learn to answer the question "Where did you just peck?" and can report peck location when unexpectedly asked. *Learning and Behavior*, *35*, 184–189.

Skinner, B. F. (1962). Two "synthetic social relations." *Journal of the Experimental Analysis of Behavior*, *5*, 531–533.

Slotnick, B. M., & Katz, H. M. (1974). Olfactory learning-set formation in rats. *Science*, *185*, 796–798.

Smith, A. P., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Behavior Processes*, *42*, 212–220.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*, 587–592.

Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, *44*, 430–444.

Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin and Review*, *17*, 412–416.

Steirn, J. N., Weaver, J. E., & Zentall, T. R. (1995). Transitive inference in pigeons: Simplified procedures and a test of value transfer theory. *Animal Learning and Behavior*, *23*, 76–82.

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs 123*, 133–167.

Tan, L., & Hackenberg, T. D. (2016). Functional analysis of mutual behavior in laboratory rats (*Rattus norvegicus*). *Journal of Comparative Psychology*, *130*, 12–23.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Century.

Topal, J., Byrne, R. W., Miklosi, A., & Csanyi, V. (2006). Reproducing human actions and action sequences: "Do as I do!" in a dog. *Animal Cognition*, *9*, 355–367.

Trapold, M. A. (1970). Are expectancies based on different reinforcing events discriminably different? *Learning and Motivation*, *1*, 129–140.

Triana, E., & Pasnak, R. (1981). Object permanence in cats and dogs. *Animal Learning and Behavior*, *9*, 135–139.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 382–403). New York: Academic Press.

Urcuioli, P. J., Zentall, T. R., Jackson-Smith, P., & Steirn, J. N. (1989). Evidence for common coding in many-to-one matching: Retention, intertrial interference, and transfer. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*, 264–273.

Wasserman, E. A., DeVolder, C. L., & Coppage, D. J. (1992). Non-similarity based conceptualization in pigeons via secondary or mediated generalization. *Psychological Science*, *6*, 374–379.

Weaver, J. E., Steirn, J. N., & Zentall, T. R. (1997). Transitive inference in pigeons: Control for differential value transfer. *Psychonomic Bulletin and Review*, *4*, 113–117.

Weir, A. A. S., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science*, *297*, 981.

Whiten, A., & Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. *Advances in the Study of Behavior*, *21*, 239–283.

Williams, D. A., Butler, M. M., & Overmier, J. B. (1990). Expectancies of reinforcer location and quality as cues for a conditional discrimination in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *16*, 3–13.

Woodruff, G., & Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, *7*, 333–362.

Woodruff, G., Premack, D., & Kennel, K. (1978). Conservation of liquid and solid quantity by the chimpanzee. *Science*, *202*, 991–994.

Zentall, T. R. (1993). Animal cognition: An approach to the study of animal behavior. In T. R. Zentall (Ed.), *Animal cognition: A tribute to Donald A. Riley* (pp. 3–15). Hillsdale, NJ: Erlbaum.

Zentall, T. R. (1996). An analysis of imitative learning in animals. In C. M. Heyes & B. G. Galef, Jr. (Eds.), *Social learning and tradition in animals* (pp. 221–243). New York: Academic Press.

Zentall, T. R. (1997). Animal memory: The role of instructions. *Learning and Motivation*, *28*, 248–267.

Zentall, T. R. (1998). Symbolic representation in pigeons: Emergent stimulus relations in conditional discrimination learning. *Animal Learning and Behavior*, *26*, 363–377.

Zentall, T. R. (2016). Reciprocal altruism in rats: Why does it occur? *Learning and Behavior*, *44*, 7–8.

Zentall, T. R., Clement, T. S., Bhatt, R. S., & Allen, J. (2001). Episodic-like memory in pigeons. *Psychonomic Bulletin and Review*, *8*, 685–690.

Zentall, T. R., Laude, J. R., Case, J. P., & Daniels, C. W. (2014). Less means more for pigeons but not always. *Psychonomic Bulletin and Review*, *21*, 1623–1628.

Zentall, T. R., Peng, D., & Miles, L. (in press). Transitive inference in pigeons may result from differential tendencies to reject the test stimuli acquired during training. *Animal Cognition*.

Zentall, T. R., & Raley, O. L. (2019). Object permanence in the pigeon: Insertion of a delay prior to choice facilitates visible- and invisible-displacement accuracy. *Journal of Comparative Psychology*, *133*, 132–139.

Zentall, T. R., & Singer, R. A. (2007). Within-trial contrast: Pigeons prefer conditioned reinforcers that follow a relatively more rather than less aversive event. *Journal of the Experimental Analysis of Behavior*, *88*, 131–149.

Zentall, T. R., & Smeets, P. M. (Eds.) (1996). *Stimulus class formation in humans and animals*. Amsterdam: North Holland.

Zentall, T. R., Sutton, J. E., & Sherburne, L. M. (1996). True imitative learning in pigeons. *Psychological Science*, *7*, 343–346.

Zucca, P., Milos, N., & Vallortigara, G. (2007). Piagetian object permanence and its development in Eurasian jays (*Garrulus glandarius*). *Animal Cognition*, *10*, 243–258.

18 The Evolution of Intelligence

Reconstructing the Pathway to the Human Mind

Lucy A. Bates and Richard W. Byrne

By almost any measure, humans as a species are extremely "brainy" – but why? Where did our intelligence come from? What selective advantages did evolving intelligence give our human ancestors, and are we the only intelligent species? Using a comparative approach – that is, by considering the brains and cognitive skills of other animal species, including those closely and more distantly related to humans – we can attempt to discover how, when, and even why our intellectual skills evolved.

Our everyday willingness to attribute intelligence to other species varies greatly. Many people are happy to call certain species "clever" (perhaps apes, dogs, elephants, or dolphins), while many other animals are dismissed as automatons (insects, even reptiles for some people). These popular characterizations often, and understandably, rest upon seeing features of our own behavior reflected in animals. For example, we know that elephants take great care of their youngsters, touching, caressing, and showing signs of empathy (Bates et al., 2008; Plotnik & de Waal, 2014) – it is easy to assume they do so thoughtfully; it is much harder to see anything of ourselves in creepy spiders or scuttling crabs. However, such human-centric assumptions about what is "clever" are anathema to biologists and comparative psychologists (Shettleworth, 2009). All species alive today have evolved to fit their environmental niche: Natural selection ensures that species which survive are optimized for their environment. Humanlike intelligence is evidently one way to succeed as a species, given our current dominance over the planet (Bar-On, Phillips, & Milo, 2018), but as any "Evolutionary Biology 101" course would tell us, we should never assume it is the only way nor that we represent some "pinnacle of evolution" that all other species struggle to attain. Evolution does not equate to "progressive" improvements toward an eventual human ideal (Dawkins, 1976; Dawkins & Wong, 2016).

Nevertheless, when we consider specifically how *human* intelligence evolved, it is both necessary and proper to look for correlates and antecedents in animals of those cognitive capacities that we consider inherent to our own intelligence. To chart the evolutionary path of human intelligence, we must first chart the differences between animal species in their cognitive aptitudes. We can then convert this into knowledge about the likely abilities of evolving *Homo sapiens* and our predecessors, utilizing the method of *evolutionary reconstruction* (Byrne, 1995). This requires having a taxonomy of a group of living species that accurately reflects their evolutionary history of relatedness (phylogeny). Such phylogenies are most reliably derived from

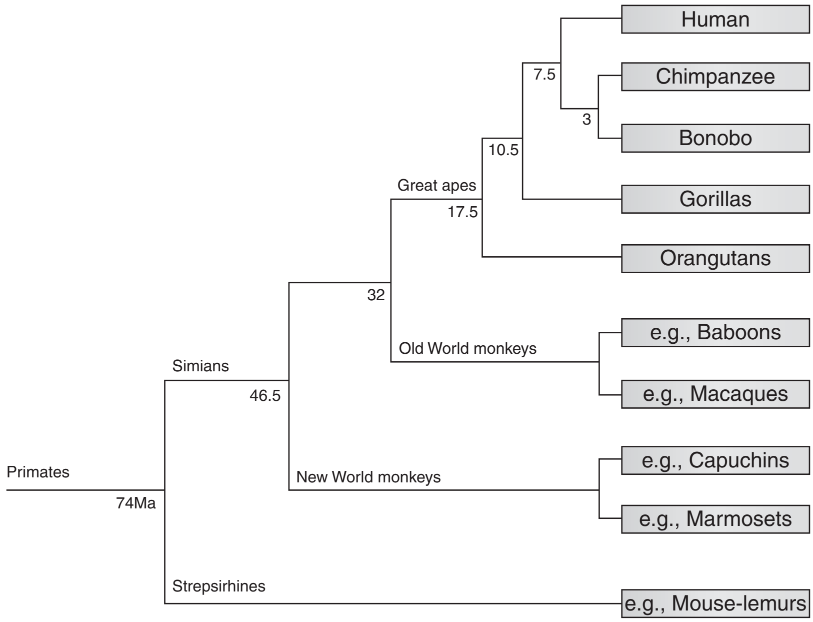

Figure 18.1 *Simplified primate phylogeny.*

Each branch point in the taxonomy implies the definite existence of an extinct ancestor of a group of living species; groups of this kind are called clades*. Each clade has a distinctive complex of features, as a direct result of their shared ancestry. Some of these shared features are* primitive *ones, retained from much earlier stages in the branching process; but some are novel, specific to the most recent common ancestor: the so-called* derived *features. Numbers show the estimated divergence date (Ma; million years ago) – or the last common ancestor – between two clades, with estimates taken from Pozzi and colleagues (2014).*

DNA sequence data (see Figure 18.1). A phylogeny allows us to know the sequence – and with appropriate calibration against fossil evidence, estimate the times – of divergences between the human lineage and that of each of our living relatives. We can then apply what we know about the behavior and cognition of each living species to ascribe characteristics to each common ancestor implied by our phylogeny, based on traits shared by all or most of their descendants. What follows is based on using this method to reconstruct the evolution of human intellectual capacities, as they varied at different points in the human lineage.

Our Shared Primate Ancestors

The two main lineages of modern primates (the prosimians, such as lemurs and bush babies, and the simians, which includes the monkeys and apes) are unlike each other in many respects. The most reasonable assumption is that their common ancestor showed the more primitive trait, in this case, the typical mammalian version. (In taxonomy, "primitive" is a technical term, not an insult: A trait inherited from a long line of earlier ancestors is presumably a highly successful one.) On this basis, our earliest primate ancestors – living around 74 Ma (million years ago; Pozzi et al., 2014) – were relatively small and not very social animals, with brains as expected in a mammal of this body size. They were nocturnal but with considerable binocular vision, and quadrupedal but with a five-fingered hand whose fingernails allowed dexterous gripping. If we wanted to think of a modern animal that was something like this, a mouse lemur *Microcebus* from Madagascar is perhaps the closest.

With such manual dexterity, these early primates would have been investigative animals, and like all modern mammals they must have possessed the ability to respond to correlations in the environment (classical, Pavlovian conditioning) and to modify their behavior flexibly on the basis of experience (instrumental, Skinnerian conditioning; Mackintosh, 1983). As with modern mammals, the efficiency of their learning would be greatly increased by biological predispositions to learn connections appropriate for survival (Morand-Ferron, 2017; Roper, 1983). To the limited extent that they were social, their exploration and learning would have benefited from a focus narrowed to the essentials by attending to the actions of others ("stimulus enhancement"). We may not consider them particularly intelligent by human standards, but we can be sure that like modern-day mouse lemurs, these animals were adept at coping with their environments, displaying learning, dexterous manipulation, exploration, and curiosity.

Our Simian Ancestors

Our simian ancestors, diverging from other primates around 46 Ma (Pozzi et al., 2014), were very different animals. Living by day rather than night and forming long-lasting social groups, these animals were larger – perhaps the size of a domestic cat – yet with a brain twice as large as would be expected for a typical mammal of that size, resulting chiefly from enlargement of the neocortex and cerebellum. We know this because these traits pervade their descendants, the living monkeys and apes. Early attempts to compare learning abilities of extant simian primates with those of other mammals in the psychologist's laboratory found no qualitative differences, but considerably greater *speed* of learning in simians (Passingham, 1981). Subsequent field studies have found very much more dramatic differences from typical mammals, and most of these center around social skill.

Monkey and ape species treat other group members as individuals: They know who is close kin to whom and who outranks whom, both in relation to themselves and third parties (Cheney & Seyfarth, 1990). Many species acquire rank by recruiting kin support; they invest grooming time, building up a network of reliable allies, and make an effort to reconcile with those allies if they should come into conflict; they recall and repay favors by their allies; they maintain these alliances over long timescales and predict the distribution of mutual help; and they notice whether interactions among third parties follow the "usual" rules of who is dominant or submissive to whom (Bergman et al., 2003; Byrne & Bates, 2010; Cheney & Seyfarth, 1990; Cheney, Seyfarth, & Silk, 1995; Crockford et al., 2007; Dunbar, 1991; Wittig et al., 2007). Moreover, all simians show occasional use of social tactics, such as deception, to navigate and manipulate these relationships and interactions, and attain personal goals at the expense of other group members (Byrne & Whiten, 1990; Whiten & Byrne, 1988). Tactics recorded include: suppressing copulation calls when mating with disallowed individuals (and so achieving secrecy); suddenly staring into the distance when pursued (and so distracting the pursuers); a juvenile screaming as if hurt when near an adult with food (and so gaining the food, when the mother attacked the other adult); and withholding food-advertisement calls when discovering small amounts of food (and so monopolizing the food).

This intricate system of social interaction must depend on sophisticated perception of other individuals and their actions, coupled with good memory for the history of each (Byrne, 1997; Byrne & Bates, 2007, 2010). Social perception must include selective-attention mechanisms that allow a sustained focus on key areas to pick up relevant information and avoid irrelevance. For simian primates, then, perceptual and memory load for social information becomes very significant as group size rises (Byrne, 2016). In addition, the individually tailored social manipulations shown by monkeys and apes imply very efficient learning powers in social contexts. Although these tactics work by manipulating the mental states of the target individuals, in most cases no understanding of the mental-state mechanism is necessary; rather, very rapid, perhaps one-trial, learning must instead be envisaged (Byrne, 1997). All the evidence thus points to a simian specialization of *rapid learning ability* in social contexts, supported by *increased perceptual and memory efficiency.*

Explaining Simian Intelligence

Given the complexity in social manipulation shared across extant simian primates, it is not surprising that sociality has repeatedly been posited as the key driver of the large brains and intelligence evident in monkeys and apes. Indeed, much research effort has been dedicated to illustrating the connections between sociality and large brains in monkeys and apes, and more recently also in a range of other animal species. Versions of this idea come in subtly different guises– *Machiavellian intelligence* (Byrne & Whiten, 1988; de Waal, 1982); *the social brain* (Brothers, 1990; Dunbar, 1998); *cultural intelligence* (van Schaik, Isler, & Burkart, 2012; van Schaik & Burkart, 2011) – but they share a core of assumptions. These are (1) some selective force encouraged our diurnal and rather small simian ancestors to live in increasingly large and tight-knit social groups; most often the driver of this is suggested to be predation pressure (van Schaik, 1983); (2) the resulting increase in sociality caused greater *within*-group competition, over access to resources such as food and/or mates; (3) living in close-knit social groups therefore selected for increased ability to manage the complex social interactions that resulted.

In short, intelligence can be considered an adaptation to group living (Humphrey, 1976). Critically, other individuals present a "moving target" of continually changing behavior, able to respond to others' strategies with their own, so creating an "arms race" of increasing abilities to manipulate others through a balance of both cooperation and competition. The net result was an increase in perception and memorization of the increasing number of relationships and personal idiosyncrasies in these enlarged groups: Rapid learning of social tactics ensured maximization of individual gains while maintaining group cohesion. On this account, the earliest origins of human memory abilities and mental swiftness arose from the social demands confronting our ancestors from 46 Ma.

There is considerable evidence supporting the notion of "social intelligence." Social species tend to do better on cognitive tests compared to closely related but more-solitary species – for example, more-social Sumatran orangutans outperformed more solitary Bornean orangutans on a battery of tests (Forss et al., 2016). Specific adaptations for social cognition are known in the primate brain (Platt, Seyfarth, & Cheney, 2016; Rushworth, Mars, & Sallet, 2013), and much of the additional support for the theory of "social intelligence" has used brain size – either relative to body size or the size of one brain area relative to others – as an index of specialization for intelligence. For example, in both macaque monkeys and humans, prefrontal cortex volume correlates with social network size (Bickart et al., 2011; Sallet et al., 2011).

However, the validity of using of brain size to index intelligence cannot simply be assumed (Healy & Rowe, 2007). For a start, do larger brain sizes really equate to greater intelligence? Until recently, this was often just assumed, but there is now considerable empirical evidence to support this idea. Among primates, for instance, species with larger brains (or brain parts) show generally enhanced cognitive performance and/or adaptability (Deaner, van Schaik, & Johnson, 2006); specifically, neocortex volume predicts the rate of tactical deception (Byrne & Corp, 2004) and the rate of social learning (Reader & Laland, 2002). The principle also holds for other animal taxa. Benson-Amram and colleagues showed that problem-solving ability was predicted by brain size in mammalian carnivores, with larger-brained species (relative to body size) better able to solve puzzle boxes baited with food than those with smaller relative brain sizes (Benson-Amram et al., 2016). MacLean and colleagues showed that absolute brain volume was the best predictor of performance on tasks measuring self-control across thirty-six mammalian and bird species, including all nonhuman great apes, various Old and New World monkey species, lemurs, rodents, carnivores, Asian elephants, domestic pigeons, jays, and sparrows (MacLean et al., 2014). There is even evidence that within certain species, larger brains equate to better performance. For example, female guppies (a small, live young–bearing fish) with larger brains performed better on a range of cognitive tasks than smaller-brained females (Kotrschal et al., 2013).

More controversially, if larger brains do indicate greater intelligence, do they really derive from the needs of sociality? Discriminating between alternative hypotheses for the evolution of intelligence is hugely difficult. Approaches rely on finding which proxy measure of evolutionary challenge best predicts brain sizes: Group size or mating system has typically been employed to index social challenge; diet type or range area for environmental challenge. Many studies reported stronger correlations with social measures than direct environmental challenges, supporting the idea of social living as the main driver of primate brain expansion (Dunbar, 1992, 1995, 2012; Dunbar & Shultz, 2007, 2017; Shultz & Dunbar, 2007). The relationships between sociality, brain size, and behavioral complexity/flexibility also apparently hold outside the primate lineage: For example, K. C. R. Fox, Muthukrishna, and Shultz (2017) showed that large brain size in cetaceans (whales and dolphins) correlates with social structure and group size, and with the breadth of social and cultural behaviors observed. In contrast, both historic and recent analyses of large data sets found that overall, primate brain size was best predicted by ecological variables: either home range size (Clutton-Brock & Harvey, 1980), diet type (frugivory or folivory) (DeCasien, Williams, & Higham, 2017); or both combined (Powell, Isler, & Barton, 2017). MacLean and colleagues (2014) even showed that dietary breadth, but not social group size, was the best predictor of self-control across the primate, other mammal, and bird species they tested. These conflicting results have yet to be resolved, but certainly, even if the evolution of simian intelligence was driven by sociality, it seems unlikely that no other environmental factors played a role beyond encouraging sociality in the first place.

Primates generally must build up adequate nutrition from dispersed, largely vegetable foods in complex tropical environments. Remembering where the key resources are and when they are likely to bear suitable food items poses a significant intellectual challenge (Milton, 1988), and there is evidence that primates use their considerable intelligence to solve such foraging problems. For example, gray-cheeked mangabey monkeys take weather variables into account when searching for figs, being more likely to revisit a fruiting tree after several days of warm and sunny weather than after cooler, cloudier periods (Janmaat, Byrne, & Zuberbühler, 2006). Moreover, Janmaat and colleagues further showed that chimpanzees use long-term spatial memory to monitor the fruit states of large feeding trees, and plan their travel routes to take in the most valuable trees (Ban et al., 2016; Janmaat, Ban, & Boesch, 2013; Janmaat et al., 2014).

Impressive abilities of a particular kind don’t unequivocally favor one evolutionary hypothesis over another: They could represent a transfer of skills. Under natural conditions, social skills, tool use, extractive foraging, and innovation are all correlated (Reader & Laland, 2002; Reader, Hager, & Laland, 2011), and a meta-analysis of an extensive range of laboratory tests of primate ability found a single factor underlying primate intelligence (Deaner et al., 2006). Street and colleagues argue that sociality, extended life history, and the propensity for social learning (cultural intelligence) all coevolved with increasing brain size in primates (Street et al., 2017). The evolution of large brains and sociality potentially promoted reliance on social transmission of information, which itself allowed for further increases in brain size and cognitive abilities. At the same time, the evolving improvements in (social) intellect could have allowed for enhancements in foraging efficiency (perhaps through social learning mechanisms – see, for example, Hobaiter et al., 2014; van de Waal, Borgeaud, & Whiten, 2013; van de Waal, Bshary, & Whiten, 2014 – as well as through other perceptual, learning and memory abilities). The improved nutrition would, in turn, have allowed further increases in brain size.

Moreover, when looking outside of mammalian taxa, it is evident that living in large social groups is not the only potential stimulus to cognitive enhancement. Among corvid birds, for example, remarkable social, tool-use, and problem-solving abilities have been found in a range of species that are not necessarily considered "social" (Emery & Clayton, 2004; Taylor, 2014). Clark's nutcrackers make several thousand caches of pine seeds each autumn, and retrieve them over the winter and spring. In the laboratory, they have consistently outperformed other, non-storing birds in spatial memory tasks. Scrub jays take account of the decay rates of cached food of different types when deciding which of their caches to excavate and consume (Clayton & Dickinson, 1998); they take account of the viewing opportunities of competitors when deciding where to cache (Dally, Emery, & Clayton, 2004); they re-cache food in new places, if competitors might have seen their original cache sites (Emery & Clayton, 2001); and they take account of which individual competitors have seen them make caches, when deciding which to consume first or re-cache in private (Dally, Emery, & Clayton, 2006). It has been argued that these abilities are based on "episodic-like" memory (Clayton & Dickinson, 1998; Correia, Dickinson, & Clayton, 2007). Ravens use a competitor's gaze direction when hiding food, distinguish between knowledgeable and ignorant competitors, and use several tactics to deceive competitors about location of foods (Bugnyar, 2002; Bugnyar & Heinrich, 2005). Eurasian jays attribute desires to conspecifics (Ostojić et al., 2013). Rooks readily learn to retrieve food from a tube with a gravity trap in it, whereas monkeys have proved inept at similar tasks (Seed et al., 2006). Both rooks and New Caledonian crows are adept at using and making tools in the laboratory, and New Caledonian crows regularly employ tools in the wild (Bird & Emery, 2009; Hunt & Gray, 2003; Taylor, 2014).

There is no clear relationship between social group size and cognitive abilities in these corvid birds (Emery et al., 2007). Rooks are colonial and feed in temporary feeding flocks; scrub jays may live in extended families; ravens are monogamous but form flocks as juveniles; nutcrackers are monogamous and a rather solitary species. Even where birds show social manipulation, it is often directed at outsiders rather than group members. So perhaps in these cases, it is the need to deal competitively with *other minds* (of observant and perhaps individually known competitors, rather than group members per se) in a crowded foraging environment that is the key driver (Byrne & Bates, 2007). Interestingly, a clear relationship is evident *within* a population of cooperatively breeding Australian magpies. Ashton and colleagues found that in these (non-corvid) birds, individuals from larger groups show increased cognitive performance across a battery of four tasks designed to measure general cognitive ability; moreover, this correlation emerged early in life, and there were clear selective benefits in the birds that displayed better cognitive performance (Ashton et al., 2018). For Australian magpies, then, sociality does shape cognition. While debates about the explanatory power of correlations between large brains and group size or ecological variables are likely to rumble on, the breadth of evidence supporting the link between sociality and intelligence cannot be disregarded.

Great Ape Ancestors

Few species of apes exist today, although a much wider range of fossil species are known. Good evidence of mental or psychological capacities is lacking for the many closely related species of gibbon or lesser ape, so our interest focuses on the great apes (*Homo, Pan, Gorilla*, and *Pongo*), which are thought to have last shared a common ancestor at 17.5 Ma (Pozzi et al., 2014). Of these, the (common) chimpanzee has been extensively studied in the field and in captivity, but our knowledge of the bonobo (or pygmy chimpanzee), the gorilla, and the orangutan is now catching up. Several times a remarkable aptitude has been first discovered in the chimpanzee and only later detected in other great ape species, so that even now it would be foolhardy to identify clear mental differences among the nonhuman great apes.

All great apes differ from monkeys in their use of objects as tools. Two monkey populations have been discovered to use rocks as hammers to gain access to foods protected by hard shells (capuchin monkeys in Brazil, cracking nuts [Fragaszy et al., 2004; Moura & Lee, 2004]; long-tailed macaque monkeys in Thailand, opening shellfish [Gumert, Kluck, & Malaivijitnond, 2009]). Field experiments have shown that the monkeys deploy considerable knowledge of everyday physics in choosing suitable rocks. Capuchin monkeys take account of properties of both the nut and anvil and adjust their tools and techniques accordingly (Luncz et al., 2016; Mangalam & Fragaszy, 2015; Ottoni & Izar, 2008). However, the skills of these monkeys do not appear to rival those of the chimpanzee.

Chimpanzees routinely use a wide range of tools in the wild (Boesch & Boesch, 1990; Goodall, 1986; McGrew, 1992; Sanz & Morgan, 2013), and for a wide range of purposes. These include pestle-pounding to gain palm-tree pulp, chiseling open bees-nests, hammering hard nuts on a stone anvil, drinking from a leaf-sponge, sexual solicitation by leaf-tearing, wiping off blood or feces with a leaf, clubbing large predators, and aimed throwing of stones. Moreover, chimpanzees *make* many of their tools, a propensity not known in any monkey. For instance, the flexible stems used to fish for insects inside a termite mound or an ant nest in a tree bole are made by picking a suitable stem, stripping off the leaves, and biting the end to remove irregularities. Different tools are made for different purposes: nine-inch flexible stems for insect-fishing, slender but rigid two-foot wands for ant-dipping, thick rods for chiseling. Since these habits appear to be local traditions, differing between chimpanzee populations, they have been described as cultures (McGrew, 1992; Whiten et al., 1999; but see Koops et al., 2014, for a review of the role of ecology in shaping tool use).

Once thought the exclusive province of chimpanzees, traditions of elaborate toolmaking and tool use are now also known from one population of orangutans (Gruber, Singleton, & van Schaik, 2012; van Schaik et al., 2003). In the wild, gorillas have seldom been recorded using tools (Breuer, Ndoundou-Hockemba, & Fishlock, 2005), and bonobos only rarely and only in a social context (Koops, Furuichi, & Hashimoto, 2015), but in captivity all great apes show ready ability to make and use tools (McGrew, 1989), with bonobos even making and using simple stone-cutting tools (Roffman et al., 2015; Roffman et al., 2012). The striking disparity in ape tool use between field and captivity seems at first sight puzzling. A clue that may resolve the puzzle comes from the plant-feeding methods of mountain gorillas, which feature many of the same traits that make chimpanzee tool traditions so remarkable (Byrne, 2016), yet without involving tools. These gorillas rely on abundant, nutritious herbs, which are difficult or painful to eat: nettles, thistles, hook-covered clambering plants, and soft piths encased in woody exteriors. The techniques used to deal with these difficulties are elaborate, in that they require correct sequencing of several different stages; each stage needs accurate coordination of the two hands performing different roles; they are flexible, with stages omitted if they are unnecessary in any given case; and they show hierarchical embedding, with subprocesses that can be iterated to a criterion. The methods for different plants are not alike, but several of them share these general features.

Although anthropological interest has always focused on the *tools* chimpanzees make and use, the manner in which they construct and employ tools is very similar to gorilla plant feeding in its structural complexity (Byrne, 2001, 2016). Chimpanzees, when eating certain plant foods, show similar characteristics to mountain gorillas (Byrne, Corp, & Byrne, 2001; Corp & Byrne, 2002; Stokes & Byrne, 2001), but it is in gaining access to insects that they more commonly employ elaborate, hierarchically embedded programs of action – and it is those tasks that involve tools. Both species show very strong behavioral laterality for these skillful tasks, but not in other, less complex aspects of their manual behavior. With only one wild population of gorillas known to show these techniques, no distinctive traditions have been noted, but the pattern of variance within the skills themselves strongly suggests that the general approach, or layout, is socially acquired in each case. Details of grip, movement, and hand preference vary idiosyncratically, but the organization of the skills is remarkably standardized despite their complexity and the range of apparently viable alternative methods. Moreover, a captive population of western lowland gorillas has been found to develop an efficient method of processing European nettles, structurally similar to the Rwandan species, and while this technique was structurally complex, it was distinctively different to that used by mountain gorillas (Byrne, Hobaiter, & Klailova, 2011). On current evidence, it is most likely that chimpanzee tool traditions and gorilla plant preparation skills are learnt observationally by *imitation of the overall organization or program*, just as humans learn motor skills from others (Byrne, 2003, 2016). Nothing in the behavior of monkeys suggests a similar ability. The stone-hammer methods used by a few monkey populations are almost certainly learnt socially, by dexterous individual exploration of the objects seen used by others; but the organization of the task is simple, and closely resembles the way in which monkeys break nuts percussively by hand.

To be able to build up novel motor skills that are elaborate in organization, and to copy the layout of these skills from other, more skillful practitioners, implies an understanding of how behavior causes changes to be effected in the physical world. Such *causal understanding* may also be seen in experiments where cooperation is required. For example, an apparatus was designed so that two parties, sitting either side, must collaborate for both to be delivered rewards: One individual can see which of two handles needs to be pulled, the other can reach and pull the handle (Mason & Hollis, 1962). Over many trials, both monkeys and apes readily learn either role. However, when rhesus monkeys are switched to the other's role, they have to learn it again, as if from scratch (Povinelli, Parks, & Novak, 1992). But when chimpanzees were tested under identical circumstances, they knew immediately how to take the other person's role (Povinelli, Nelson, & Boysen, 1992). Whether the apes' superior understanding is a matter of understanding physical cause and effect, or mental intentions and knowledge, is here a moot point; but there is also evidence that great apes *can* appreciate causes that lie purely in the mental realm, to which we turn next.

In using language, we routinely take account of the perspective of our audience. Great apes share something of this ability. In using their natural gestures for communication, all species of great ape have been shown to adjust their signaling according to the ability of the target audience to perceive their gestures. For an audience who is facing and paying attention to them, they readily use silent gestures, whereas for an audience facing away or not paying attention, they use tactile gestures rather than silent visual ones (Byrne et al., 2017; Genty et al., 2009; Hobaiter, Byrne, & Zuberbühler, 2017; Hobaiter & Byrne, 2011; Tomasello & Call, 1997). Moreover, apes can take account of the audience's understanding, as well as their perceptual abilities. Taking advantage of zoo-housed orangutans' tendency to beg treats from keepers, Cartmill set up a situation in which the keeper apparently misunderstood requests for a favorite treat, in two ways (Cartmill & Byrne, 2007). Two foods were placed equidistant from keeper and ape: The ape, of course, gestured enthusiastically at the favored one. When the keeper "misunderstood entirely," giving the orangutan instead a nutritious but unpopular food, it switched to new gesture types; when the keeper "slightly misinterpreted" the ape, giving the orangutan only part of the treat, it continued with the same gestures but increased its signaling rate.

Taking account of another's perspective and knowledge was also shown in an unpublished experiment using warning calls given to a "predator": a veterinarian with a dart gun stands in for a predator for a zoo animal. Researchers arranged that one chimpanzee could clearly see the predatory vet's approach, whereas the other – his close ally – might not be able to. The observer distinguished between cases when their ally was quite unable to see the predator, when they call zealously, and when both chimpanzees were well able to see the predator, when the observer often remained silent (see Byrne, 2016, p. 20 for a more detailed description of this study). The distinction is based on understanding that an individual that can *see* a risk must therefore *know* about it, and is evidently fundamental to real communication, in which audience knowledge is taken into account. However, monkeys in the same experimental situation make no such distinction, calling as avidly in either case (Cheney & Seyfarth, 1990). Recently, chimpanzees in the wild have been shown similarly to take into account the knowledge of others to whom they signal (Crockford et al., 2012; Crockford, Wittig, & Zuberbühler, 2017; Schel et al., 2013). Experimenters arranged for a realistic model snake to be revealed to a foraging party of chimpanzees; naturally, they reacted with both alarm and curiosity at first, then calmed as the "snake" showed no signs of attack. The interesting question was what would happen when other chimpanzees approached:

Would those who had already seen the snake, in no need of warning themselves, realize that newcomers might be vulnerable? Evidently they did, since they gave alarm calls specifically when their relatives or allies who had not been present when the snake was first revealed arrived at the scene.

That great apes, but not monkeys, can evaluate the perspective and likely knowledge of others is consistent with older findings on observational records of deception. As noted above, most records of tactical deception can be explained as a product of very rapid associative learning by individuals who may not be able to assess other's knowledge: as functional but not necessarily intentional deception. However, a subset of records could not plausibly be understood this way, and instead implied the ability to understand other-person knowledge (Byrne & Whiten, 1991). These records were not distributed at random, or in rough proportion to the number of cases from each species, but were tightly clumped in one taxonomic group: the great apes (Byrne & Whiten, 1992). Records came from gorillas, bonobos, and orangutans as well as chimpanzees, so this seems to be an aptitude general in great apes. These data were for many years controversial, since experimental tests failed to reveal any ability to understand false belief in nonhuman great apes (Call & Tomasello, 2008; Hare, Call, & Tomasello, 2001; Kaminski, Call, & Tomasello, 2008; Tomasello & Call, 1997; Tomasello, Call, & Hare, 2003), but recently *false-belief understanding* has been demonstrated in great apes (Krupenye et al., 2016). These experimenters used a method devised to explore knowledge in preverbal children, anticipatory looking, to find out whether chimpanzees, bonobos, and orangutans could understand the consequences of another individual's false belief. In all cases, their anticipatory looking was predicted by the other individual's belief rather than their own knowledge.

The implication is that an individual of the common ancestor species at 17.5 Ma already had some appreciation of (1) how behavior has its effects on the physical world, (2) how familiar actions can be organized into more novel and sometimes complex plans, and (3) the knowledge state and intentions of other individuals, insofar as they differ from its own. Great apes have a *qualitative increase in causal understanding and knowledge of "other minds"* compared to monkeys.

Explaining Great Ape Intelligence

The causal understanding of great apes is not associated with any greater social demands: Indeed, in gorillas and orangutans, social group sizes are considerably smaller than those of most monkey species. So to what environmental problem were these cognitive skills an adaptive solution? One possibility is that the problems of food acquisition favored individuals with the most cost-effective techniques (Byrne, 1997). This is plausible, because great apes are heavy animals, yet, since they are adapted to hanging in tree canopies to allow small-branch feeding, their locomotion on the ground is energetically less efficient than that of monkeys. Furthermore, throughout their range, apes compete for ripe fruit and tender leaves directly with Old World monkeys, which are known to be able to digest coarser material than apes and so are able to exploit ripening plant food before their ape competitors. Since monkeys appear to possess all the aces, it becomes a problem to explain why all great apes did not become extinct (although many species did, according to the geological record). Clearly, the surviving great apes must have some compensatory advantage, and the suggestion is that they were able to compete by developing skills to reach foods that monkeys could not reach, as shown today in their expertise at extracting insects and dealing with plant defenses. That expertise is itself enabled by apes' ability to build up complex structures of behavior (Byrne, 1997), giving them advantages in the domain of technical rather than social cognition.

This conjecture is supported by the fact that all genera of living great apes show special skills in manual food processing: *Pongo*, in circumventing spiny rattans and palm spines in accessing fruits (Russon, 1998) and for extracting honey and seeds with tools (E. Fox, Sitompul, & van Schaik, 1999); *Gorilla*, for processing physically defended herb resources (Byrne, 2001); *Pan*, in collecting insect foods with tools, often ones made themselves and sometimes sets of two tools for a more complex task (McGrew, 1992; Sanz & Morgan, 2007). Moreover, as noted already, in captivity all great apes demonstrate remarkably similar tool-using and tool-making abilities, though many populations show no tool use in the wild (McGrew, 1989), suggesting that the underlying cognitive skills – all ones to do with feeding – have an ancient origin in the common ancestry of all the modern lines. Acquiring these skills through observing others also fits with the increasing evidence that suggests the repertoire of cognitive skills in any individual great ape depends on opportunities for social acquisition of knowledge (van Schaik & Burkart, 2011; van Schaik, Deaner, & Merrill, 1999). Moreover, Barton and Venditti (2014) have shown that the cerebellum underwent rapid expansion throughout great ape evolution, giving great apes (including humans) significantly larger than expected cerebellums. The cerebellum has a role in sensory-motor control and in learning complex action sequences, so cerebellar specialization likely underpins the advanced technical capabilities of great apes. The particular kind of information that apes seem adapted to extract from watching others – the hierarchical organization of planned behavior, rather than the exact blow-by-blow series of movements – has the potential to give them insight of a causal/intentional sort (Byrne, 2003). The regularly occurring result of a specific, organized sequence of actions can be seen as the intentional goal of performing the actions (especially if the individual concerned appears satisfied); and the regularly occurring sequel to an action or series of actions can be seen as caused by the action(s). These attributions are statistical, lacking the deep understanding of causality and intention that (some) humans possess, but they may be quite sufficient to give the edge over animals that lack this sort of understanding, and – more to the point – for passing nonverbal tests of theory of mind and physical causality.

The Last Common Ancestors

Our closest relatives are the two *Pan* species, the chimpanzee and the bonobo; based on the best current estimates from molecular taxonomy, their ancestors split from ours only 7.5 Ma (Pozzi et al., 2014). Unfortunately, reconstruction of the psychology of this "last common ancestor" species is particularly difficult. The two *Pan* species themselves shared ancestry until about 3.5 Ma, yet now they differ in various ways (Gruber & Clay, 2016). Common chimpanzees are routine toolmakers with elaborate social customs of tool use, living in loose communities somewhat dominated by coalitions of males, who are effective group hunters of relatively large prey, and who exhibit occasional infanticide and intercommunity lethal aggression (Goodall, 1986; Wilson et al., 2014). Bonobos do not regularly show any of these traits, but instead live in more cohesive groups in which non-reproductive sex (which is found in a variety of different patterns in all great apes) seems to dominate social bonding, and females play a much more equal role in group activities, including leading hunts (Clay, Furuichi, & de Waal, 2016; Furuichi et al., 2015; Tokuyama & Furuichi, 2016). For many characteristics, the state of the *Pan* common ancestor at 3.5 Ma remains indeterminate; this makes it very difficult to use evolutionary reconstruction to decide on the nature of our common ancestor at 7.5 Ma.

Furthermore, one captive bonobo (named Kanzi) has proved able, given an enriched but essentially normal human upbringing and the use of a board of word icons with which to express himself (Savage-Rumbaugh & Lewin, 1994), to achieve remarkable linguistic success. Kanzi can understand some spoken English delivered at normal speech rates, reliably identifying words that differ on a single phoneme, and using complex syntax such as embedded relative clauses to decipher meaning (Savage-Rumbaugh et al., 1993; Savage-Rumbaugh, Shanker, & Taylor, 1998; Savage-Rumbaugh et al., 1996). This performance has been demonstrated under controlled testing, with novel (or even absurd) sentences. Bonobos and chimpanzees lack the breath control, lip, and tongue control, and to some extent larynx configuration to speak: Kanzi's production therefore lags far behind that of humans, but nevertheless serves to convey a rich range of meanings in an organized way. The fact that such a degree of language comprehension can be developed in a chimpanzee raises the possibility that language itself is not a genetically programmed adaptation, but a program, which is learnt by each child under normal human rearing (Byrne, 1995; Lock, 1980; Savage-Rumbaugh et al., 1998). Whether Kanzi's remarkable skills will prove special to bonobos, or to the naturalistic regime of rearing he has experienced, is not yet known; nor are any comparable skills established for bonobos in the wild. However, these are the least well-studied and the rarest great apes, and the future may overturn many of our current beliefs about them. At present, we cannot reliably attribute any greater cognitive skills to our ancestors of 7.5 Ma than those of 17.5 Ma, but the chimpanzee/human clade may well prove to show unique cognitive adaptations of some sort.

Finally, we can only speculate on what occurred in the uniquely human lineage after our split with the *Pan* clade. Evidently, differences exist between us and the other great apes, in our brain size and architecture (Raghanti et al., 2018), in our cognitive skills, and in our social and cultural complexity. Tomasello has proposed that humans are not just adept at learning useful things from others, but are also driven to conform to others in order to affiliate with them – our identification with and maintenance of cultural groups is what makes us uniquely different (Tomasello, 2014, 2016; Tomasello et al., 2005; Tomasello, Kruger, & Ratner, 1993). The role of culture in "human-unique" intelligence is gaining increasing research attention (Heyes, 2018; Laland, 2017), and receives support from observations that humans are particularly specialized in the social realm (Herrmann et al., 2007), that children rely more on social information than do nonhuman great apes (van Leeuwen, Call, & Haun, 2014), and – particularly – that we seem to possess an inherent motivation to associate with and therefore obtain information from knowledgeable "in-group" members (for example, see Begus, Gliga, & Southgate, 2014, 2016).

Building the Composite Picture

In this chapter we have traced the possible evolution of human intelligence through our ancestral primate lineage, and in so doing we have outlined and evaluated several of the theories that seek to explain the evolution of our (arguably) most remarkable trait: our intelligence. Evolutionary reconstruction is inherently more ambiguous about *why* changes occurred than *what* did occur. We have used correlational evidence to evaluate plausible speculations, but as with any history, we may never know for sure what caused a change. Below is our current best guess at the picture.

Although our stock before 70 Ma was typically mammal-like, these earliest primate ancestors should not be underestimated. Although rather unspecialized mammals, our ancestral population was even at this date adapted to survive by flexibility; not surprisingly, perhaps, when one considers that they competed alongside the dinosaurs for millions of years, and then survived the Cretaceous/Tertiary extinction event at 65 Ma.

By 46 Ma, the monkey and ape lineage had developed traits that today are closely associated with living in long-lasting groups. The nocturnal existence of the earliest primates had by this time given way to diurnal and arboreal living: an arboreal primate in daytime is conspicuous, inevitably increasing predation risk and the benefit of group living. The psychological skills of these simian ancestors would have been clearest in the social domain, although fast learning would have been evident in all contexts: very much as in modern monkeys. Characteristics would have included differentiated reactions among individuals, the importance of third parties in conflicts, collaboration and coalitions built up and repaired by targeted affiliative interactions, social manipulation, and deception. Impressive social skills were underpinned by good skills of memory and perception, allowing rapid learning of nuanced distinctions among social companions. In contrast, these animals would not have possessed any deep understanding of the causes of other individuals' behavior, even in the social domain. The implication for human origins is that in human evolution, complex sociality predated the ability to understand other people as causal agents with independent minds.

In the human lineage, only the great apes have developed these latter skills, whose origin therefore appears to be from 17.5 Ma. This ancestor population was again of forest-living individuals, adapted to forage delicately by hanging, despite their considerable bulk – but correspondingly less able to travel efficiently on the ground. They would have possessed a much more general ability to understand causality, both physical (how actions achieve changes to objects) and mental (how varying intentions and knowledge affect individuals' behavior). There is no sign that this species was any more social than its ancestors of 46 Ma onwards, so the enhanced understanding of mechanism seems unlikely to have a social origin. A possible alternative is that, in the face of the severe challenge presented by the cooling Miocene climate, the surviving great apes were those able to bring to bear enhanced feeding skills, by acquiring complex and efficacious techniques for manually obtaining and processing food (Byrne, 2016). This package of abilities included organizing structures of actions hierarchically into novel programs, understanding and hence copying aspects of the skilled behavior of others, seeing how actions achieve their effects, and – where useful – incorporating tools into programs of actions, tools which themselves sometimes must be fashioned with skill and dexterity. On this thesis, the understanding of other minds was secondarily derived from an understanding of (rather more visible) behavioral causality.

In any case, the last common ancestor at 7.5 Ma was equipped with several cognitive aptitudes that form the bedrock of later, distinctively human adaptations. These apes were already large-brained and able to keep track of extensive and complex social relationships; quick to learn and able to exploit other individuals in complex ways; able to understand how actions change objects in the world, and how their own and others' actions can be organized into hierarchical programs to achieve novel goals; and that other individuals sometimes have knowledge and goals different from their own. Animals like this would be able to see the purpose of true communication, in which speaker and hearer try to model the other's knowledge, and would be able to build up hierarchically embedded structures of action, so crucial to human language.

References

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*, *554*(7692), 364–367.

Ban, S. D., Boesch, C., Guessan, A. N., Kouakou, E., Goran, N., Tako, A., et al. (2016). Taï chimpanzees change their travel direction for rare feeding trees providing fatty fruits. *Animal Behaviour*, *118*, 135–147.

Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. *Proceedings of the National Academy of Sciences*, pnas1711842115.

Barton, R. A., & Venditti, C. (2014). Rapid evolution of the cerebellum in humans and other great apes. *Current Biology*, *24*(20), 2440–2444.

Bates, L. A., Lee, P. C., Njiraini, N., Poole, J., Sayialel, K., Sayialel, S., et al. (2008). Do elephants show empathy? *Journal of Consciousness Studies*, *15*(10–11), 204–225.

Begus, K., Gliga, T., & Southgate, V. (2014). Infants learn what they want to learn: Responding to infant pointing leads to superior learning. *PloS One*, *9*(10), e108817.

Begus, K., Gliga, T., & Southgate, V. (2016). Infants' preferences for native speakers are associated with an expectation of information. *Proceedings of the National Academy of Sciences*, *113*(44), 12397–12402.

Benson-Amram, S., Dantzer, B., Stricker, G., Swanson, E. M., & Holekamp, K. E. (2016). Brain size predicts problem-solving ability in mammalian carnivores. *Proceedings of the National Academy of Sciences*, *113*(9), 2532–2537.

Bergman, T. J., Beehner, J. C., Cheney, D. L., & Seyfarth, R. M. (2003). Hierarchical classification by rank and kinship in baboons. *Science*, *302*, 1234–1236.

Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C., & Barrett, L. F. (2011). Amygdala volume and social network size in humans. *Nature Neuroscience*, *14* (2), 163–164.

Bird, C. D., & Emery, N. J. (2009). Insightful problem solving and creative tool modification by captive nontool-using rooks. *Proceedings of the National Academy of Sciences*, *106*(25), 10370–10375.

Boesch, C., & Boesch, H. (1990). Tool use and tool making in wild chimpanzees. *Folia Primatol*, *54*, 86–99.

Breuer, T., Ndoundou-Hockemba, M., & Fishlock, V. (2005). First observation of tool use in wild gorillas. *PLoS Biology*, *3*(11), 2041–2043.

Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, *1*, 27–51.

Bugnyar, T. (2002). Observational learning and the raiding of food caches in ravens, *Corvus corax*: Is it "tactical" deception? *Animal Behaviour*, *64*, 185–195.

Bugnyar, T., & Heinrich, B. (2005). Ravens, *Corvus corax*, differentiate between knowledgeable and ignorant competitors. *Proceedings of the Royal Society B: Biological Sciences*, *272*, 1641–1646.

Byrne, R. W. (1995). *The thinking ape: Evolutionary origins of intelligence*. Oxford: Oxford University Press.

Byrne, R. W. (1997). The technical intelligence hypothesis: An additional evolutionary stimulus to intelligence? In A. Whiten & R. W. Byrne (Eds.), *Machiavellian intelligence II: extensions and evaluations* (pp. 289–311). Cambridge, UK: Cambridge University Press.

Byrne, R. W. (2001). Clever hands: The food processing skills of mountain gorillas. In M. Robbins, P. Sicotte, & K. Stewart (Eds.), *Mountain gorillas: Three decades of research at Karisoke* (pp. 294–313). Cambridge, UK: Cambridge University Press.

Byrne, R. W. (2003). Imitation as behaviour parsing. *Philosophical Transactions of the Royal Society of London Series B*, *358*(1431), 529–536.

Byrne, R. W. (2016). *Evolving insight*. Oxford: Oxford University Press.

Byrne, R. W., & Bates, L. A. (2007). Sociality, evolution and cognition. *Current Biology*, *17* (16), R714–R723.

Byrne, R. W., & Bates, L. A. (2010). Primate social cognition: uniquely primate, uniquely social, or just unique? *Neuron*, *65*(6), 815–830.

Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: Intentional communication with a rich set of innate signals. *Animal Cognition*, *20*(4), 755–769.

Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1549), 1693–1699.

Byrne, R. W., Corp, N., & Byrne, J. M. (2001). Manual dexterity in the gorilla: Bimanual and digit role differentiation in a natural task. *Animal Cognition*, *4*, 347–361.

Byrne, R. W., Hobaiter, C., & Klailova, M. (2011). Local traditions in gorilla manual skill: Evidence for observational learning of behavioral organization. *Animal Cognition*, *14*(5), 683–693.

Byrne, R. W., & Whiten, A. (Eds.) (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*. Oxford: Clarendon Press.

Byrne, R. W., & Whiten, A. (1990). Tactical deception in primates: the 1990 database. *Primate Report*, *27*, 1–101.

Byrne, R. W., & Whiten, A. (1991). Computation and mindreading in primate tactical deception. In A. Whiten (Ed.), *Natural theories of mind* (pp. 127–141). Oxford: Basil Blackwell.

Byrne, R. W., & Whiten, A. (1992). Cognitive evolution in primates: Evidence from tactical deception. *Man*, *27*, 609–627.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*(5), 187–192.

Cartmill, E. A., & Byrne, R. W. (2007). Orangutans modify their gestural signaling according to their audience's comprehension. *Current Biology*, *17*(15), 1345–1348.

Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.

Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions : Evidence for causal reasoning ? *Journal of Comparative Psychology*, 109(2), 134–141.

Clay, Z., Furuichi, T., & de Waal, F. B. M. M. (2016). Obstacles and catalysts to peaceful coexistence in chimpanzees and bonobos. *Behaviour*, *153*(9–11), 1–38.

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, *395*, 272–278.

Clutton-Brock, T. H., & Harvey, P. H. (1980). Primates, brains and ecology. *Journal of Zoology*, *190*(3), 309–323.

Corp, N., & Byrne, R. W. (2002). Leaf processing by wild chimpanzees: Physically defended leaves reveal complex manual skills. *Ethology*, *108*, 673–696.

Correia, S., Dickinson, A., & Clayton, N. (2007). Western scrub-jays anticipate future needs independently of their current motivational state. *Current Biology*, *17*, 856–861.

Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, *22*(2), 142–146.

Crockford, C., Wittig, R. M., Seyfarth, R. M., & Cheney, D. L. (2007). Baboons eavesdrop to deduce mating opportunities. *Animal Behaviour*, *73*, 885–890.

Crockford, C., Wittig, R. M., & Zuberbühler, K. (2017). Vocalizing in chimpanzees is influenced by social-cognitive processes. *Science Advances*, *3*(11), e1701742.

Dally, J. M., Emery, N. J., & Clayton, N. S. (2004). Cache protection strategies by western scrub-jays (*Aphelocoma californica*): Hiding food in the shade. *Proceedings of the Royal Society B: Biological Sciences*, *271*, S387–S390.

Dally, J. M., Emery, N. J., & Clayton, N. S. (2006). Food-caching western scrub-jays keep track of who was watching when. *Science*, *312*, 1662–1665.

Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.

Dawkins, R., & Wong, Y. (2016). *The ancestor's tale: A pilgrimage to the dawn of life*. London: Weidenfeld & Nicholson.

de Waal, F. B. M. (1982). *Chimpanzee politics*. London: Jonathan Cape.

Deaner, R. O., van Schaik, C. P., & Johnson, V. (2006). Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. *Evolutionary Psychology*, *4*, 149–196.

DeCasien, A. R., Williams, S. A., & Higham, J. P. (2017). Primate brain size is predicted by diet but not sociality. *Nature Ecology and Evolution*, *1*(5), 1–7.

Dunbar, R. I. M. (1991). Functional significance of social grooming in primates. *Folia Primatologica*, *57*, 121–131.

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *20*, 469–493.

Dunbar, R. I. M. (1995). Neocortex size and group size in primates: A test of the hypothesis. *Journal of Human Evolution*, *28*, 287–296.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, *6*, 178–190.

Dunbar, R. I. M. (2012). Bridging the bonding gap: The transition from primates to humans. *Philosophical Transactions of the Royal Society B – Biological Sciences*, *367* (1597), 1837–1846.

Dunbar, R. I. M., & Shultz, S. (2007). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society Series B – Biological Sciences*, *362*, 649–658.

Dunbar, R. I. M., & Shultz, S. (2017). Why are there so many explanations for primate brain evolution? *Philosophical Transactions of the Royal Society Series B – Biological Sciences*, *372*, 20160244.

Emery, N. J., & Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, *414*, 443–446.

Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, *306*(5703), 1903–1907.

Emery, N. J., Seed, A. M., von Bayern, A. M. P., & Clayton, N. S. (2007). Cognitive adaptations of social bonding in birds. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1480), 489–505.

Forss, S. I. F., Willems, E., Call, J., & van Schaik, C. P. (2016). Cognitive differences between orang-utan species: A test of the cultural intelligence hypothesis. *Scientific Reports*, *6*, 1–12.

Fox, E., Sitompul, A., & van Schaik, C. P. (1999). Intelligent tool use in wild Sumatran orangutans. In S. T. Parker, H. L. Miles, & R. W. Mitchell (Eds.), *The mentality of gorillas and orangutans* (pp. 99–116). Cambridge, UK: Cambridge University Press.

Fox, K. C. R., Muthukrishna, M., & Shultz, S. (2017). The social and cultural roots of whale and dolphin brains. *Nature Ecology and Evolution*. https://COMP: Linkdoi.org/10.1038/s41559-017-0335-y

Fragaszy, D., Izar, P., Visalberghi, E., Ottoni, E. B., & De Oliveira, M. G. (2004). Wild capuchin monkeys (*Cebus libidinosus*) use anvils and stone pounding tools. *American Journal of Primatology*, *64*, 359–366.

Furuichi, T., Sanz, C., Koops, K., Sakamaki, T., Ryua, H., Tokuyamaa, N., & Morgan, D. (2015). Why do wild bonobos not use tools like chimpanzees do? *Behaviour*, *52* (3–4), 425–460.

Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (*Gorilla gorilla*): Repertoire, intentionality and possible origins. *Animal Cognition*, *12*(3), 527–546.

Goodall, J. (1986). *The chimpanzees of Gombe*. Cambridge, MA: Belknap Press.

Gruber, T., & Clay, Z. (2016). A comparison between bonobos and chimpanzees: A review and update. *Evolutionary Anthropology*, *25*(5), 239–252.

Gruber, T., Singleton, I., & van Schaik, C. (2012). Sumatran orangutans differ in their cultural knowledge but not in their cognitive abilities. *Current Biology*, *22*(23), 2231–2235.

Gumert, M. D., Kluck, M., & Malaivijitnond, S. (2009). The physical characteristics and usage patterns of stone axe and pounding hammers used by long-tailed macaques in the Andaman Sea region of Thailand. *American Journal of Primatology*, *71*(7), 594–608.

Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, *61*(1), 139–151.

Healy, S. D., & Rowe, C. (2007). A critique of comparative studies of brain size. *Proceedings of the Royal Society B: Biological Sciences*, *274*, 453–464.

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., Tomasello, M., Hernandez-Lloreda, M. V., et al. (2007). Humans have evolved specialised skills of social cognition: The cultural intelligence hypothesis. *Science*, *317*(5843), 1360–1366.

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Cambridge, MA: Harvard University Press.

Hobaiter, C., & Byrne, R. W. (2011). Serial gesturing by wild chimpanzees: Its nature and function for communication. *Animal Cognition*, *14*(6), 827–838.

Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, *71*(96). https://doi.org/10.1007/s00265-017-2325-1

Hobaiter, C., Poisot, T., Zuberbühler, K., Hoppitt, W., & Gruber, T. (2014). Social network analysis shows direct evidence for social transmission of tool use in wild chimpanzees. *PLoS Biology*, *12*(9), e1001960.

Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge, UK: Cambridge University Press.

Hunt, G., & Gray, R. (2003). Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1571), 867–874.

Janmaat, K. R. L., Ban, S. D., & Boesch, C. (2013). Chimpanzees use long-term spatial memory to monitor large fruit trees and remember feeding experiences across seasons. *Animal Behaviour*, *86*(6), 1183–1205.

Janmaat, K. R. L., Byrne, R. W., & Zuberbühler, K. (2006). Primates take weather into account when searching for fruits. *Current Biology*, *16*, 1232–1237.

Janmaat, K. R. L., Polansky, L., Ban, S. D., & Boesch, C. (2014). Wild chimpanzees plan their breakfast time, type, and location. *Proceedings of the National Academy of Sciences*, *111*(46), 16343–16348.

Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, *109*(2), 224–234.

Koops, K., Furuichi, T., & Hashimoto, C. (2015). Chimpanzees and bonobos differ in intrinsic motivation for tool use. *Scientific Reports*, *16*(5), 11356.

Koops, K., Visalberghi, E., & van Schaik, C. P. (2014). The ecology of primate material culture. *Biology Letters*, *10*(11), 20140508.

Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I., et al. (2013). Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. *Current Biology*, *23*(2), 168–171.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*(6308), 110–114.

Laland, K. N. (2017). *Darwin's unfinished symphony: How culture made the human mind*. Princeton: Princeton University Press.

Lock, A. (1980). *The guided reinvention of language*. London: Academic Press.

Luncz, L. V., Falótico, T., Pascual-Garrido, A., Corat, C., Mosley, H., & Haslam, M. (2016). Wild capuchin monkeys adjust stone tools according to changing nut properties. *Scientific Reports*, *6*, 1–9.

Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press.

MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., et al. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences*, *111* (20), E2140–E2148.

Mangalam, M., & Fragaszy, D. M. (2015). Wild bearded capuchin monkeys crack nuts dexterously. *Current Biology*, *25*(10), 1334–1339.

Mason, W. A., & Hollis, J. H. (1962). Communication between young rhesus monkeys. *Animal Behaviour*, *10*(3–4), 211–221.

McGrew, W. C. (1989). Why is ape tool use so confusing? In V. Standen & R. A. Foley (Eds.), *Comparative socioecology: The behavioural ecology of humans and other mammals* (pp. 457–472). Oxford: Blackwell Scientific.

McGrew, W. C. (1992). *Chimpanzee material culture : Implications for human evolution*. Cambridge, UK: Cambridge University Press.

Milton, K. (1988). Foraging behaviour and the evolution of primate intelligence. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans* (pp. 285–305). Oxford: Clarendon Press.

Morand-Ferron, J. (2017). Why learn? The adaptive value of associative learning in wild populations. *Current Opinion in Behavioral Sciences*, *16*, 73–79.

Moura, A. C. de A., & Lee, P. C. (2004).Capuchin stone tool use in Caatinga dry forest. *Science*, *306*(5703), 1909.

Ostojić, L., Shaw, R., Cheke, L., & Clayton, N. (2013). Evidence suggesting that desire-state attribution may govern food sharing in Eurasian jays. *Proceedings of the National Academy of Sciences*, *110*, 4123–4128.

Ottoni, E. B., & Izar, P. (2008). Capuchin monkey tool use: Overview and implications. *Evolutionary Anthropology*, *17*, 171–178.

Passingham, R. E. (1981). Primate specializations in brain and intelligence. *Symposia of the Zoological Society of London*, *46*, 361–388.

Platt, M. L., Seyfarth, R. M., & Cheney, D. L. (2016). Adaptations for social cognition in the primate brain. *Philosophical Transactions of the Royal Society Series B*, *371*, 20150096.

Plotnik, J. M., & de Waal, F. B. M. (2014). Asian elephants (*Elephas maximus*) reassure others in distress. *PeerJ*, *2*, e278.

Povinelli, D. J., Nelson, K. E., & Boysen, S. T. (1992). Comprehension of role reversal in chimpanzees: Evidence of empathy? *Animal Behaviour*, *43*(4), 633–640.

Povinelli, D. J., Parks, K. A., & Novak, M. A. (1992). Role reversal by rhesus monkeys, but no evidence of empathy. *Animal Behaviour*, *44*, 269–281.

Powell, L. E., Isler, K., & Barton, R. A. (2017). Re-evaluating the link between brain size and behavioural ecology in primates. *Proceedings of the Royal Society B: Biological Sciences*, *284*, 20171765.

Pozzi, L., Hodgson, J. A., Burrell, A. S., Sterner, K. N., Raaum, R. L., & Disotell, T. R. (2014). Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. *Molecular Phylogenetics and Evolution*, *75*(1), 165–183.

Raghanti, M. A., Edler, M. K., Stephenson, A. R., Munger, E. L., Jacobs, B., Hof, P. R., et al. (2018). A neurochemical hypothesis for the origin of hominids. *Proceedings of the National Academy of Sciences*, 1719666115.

Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society Series B – Biological Sciences*, *366*(1567), 1017–1027.

Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences*, *99*(7), 4436–4441.

Roffman, I., Savage-Rumbaugh, S., Rubert-Pugh, E., Ronen, A., & Nevo, E. (2012). Stone tool production and utilization by bonobo-chimpanzees (*Pan paniscus*). *Proceedings of the National Academy of Sciences*, *109*(36), 14500–14503.

Roffman, I., Savage-Rumbaugh, S., Rubert-Pugh, E., Stadler, A., Ronen, A., & Nevo, E. (2015). Preparation and use of varied natural tools for extractive foraging by bonobos (*Pan paniscus*). *American Journal of Physical Anthropology*, *158*(1), 78–91.

Roper, T. J. (1983). Learning as a biological phenomenon. In T. R. Halliday & P. J. B. Slater (Eds.), *Animal behaviour*, vol. 3, *Genes, development and learning* (pp. 178–212). Oxford: Blackwell Scientific.

Rushworth, M. F. S., Mars, R. B., & Sallet, J. (2013). Are there specialized circuits for social cognition and are they unique to humans? *Current Opinion in Neurobiology*, *23*(3), 436–442.

Russon, A. E. (1998). The nature and evolution of intelligence in orangutans (*Pongo pygmaeus*). *Primates*, *39*, 485–503.

Sallet, J., Mars, R. B., Noonan, M. P., Andersson, J. L., O'Reilly, J. X., Jbabdi, S., et al. (2011). Social network size affects neural circuits in macaques. *Science*, *334*, 697–700.

Sanz, C. M., & Morgan, D. B. (2007). Chimpanzee tool technology in the Goualougo Triangle, Republic of Congo. *Journal of Human Evolution*, *52*, 420–433.

Sanz, C., & Morgan, D. (2013). Ecological and social correlates of chimpanzee tool use. *Philosophical Transactions of the Royal Society Series B*, *368*, 20120416.

Savage-Rumbaugh, E. S., & Lewin, R. (1994). *Kanzi. The ape at the brink of the human mind*. New York: John Wiley.

Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, *58*, 1–252.

Savage-Rumbaugh, E. S., Shanker, S. G., & Taylor, T. J. (1998). *Apes, language and the human mind*. New York: Oxford University Press.

Savage-Rumbaugh, E. S., Williams, S. L., Furuichi, T., & Kano, T. (1996). Language perceived: *Paniscus* branches out. In W. C. McGrew, L. F. Marchant, & T. Nishida (Eds.), *Great ape societies* (pp. 173–184). Cambridge, UK: Cambridge University Press.

Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call production meets key criteria for intentionality. *PLoS One*, *8* (10), e76674.

Seed, A. M., Tebbich, S., Emery, N., & Clayton, N. S. (2006). Investigating physical cognition in rooks, *Corvus frugilegus*. *Current Biology*, *16*, 697–701.

Shettleworth, S. J. (2009). The evolution of comparative cognition : Is the snark still a boojum? *Behavioural Processes*, *80*, 210–217.

Shultz, S., & Dunbar, R. I. M. (2007). The evolution of the social brain: Anthropoid primates contrast with other vertebrates. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1624), 2429–2436.

Stokes, E. J., & Byrne, R. W. (2001). Cognitive capacities for behavioural flexibility in wild chimpanzees (*Pan troglodytes*): The effect of snare injury on complex manual food processing. *Animal Cognition*, *4*, 11–28.

Street, S. E., Navarrete, A. F., Reader, S. M., & Laland, K. N. (2017). Coevolution of cultural intelligence, extended life history, sociality, and brain size in primates. *Proceedings of the National Academy of Sciences*, *114*(30), 7908–7914.

Taylor, A. H. (2014). Corvid cognition. *WIREs Cognitive Science*, https://doi.org/10.1002/wcs.1286

Tokuyama, N., & Furuichi, T. (2016). Do friends help each other ? Patterns of female coalition formation in wild bonobos at Wamba. *Animal Behaviour*, *119*, 27–35.

Tomasello, M. (2014). The ultra-social animal. *European Journal of Social Psychology*, *44* (3), 187–194.

Tomasello, M. (2016). The ontogeny of cultural learning. *Current Opinion in Psychology*, *8*, 1–4.

Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford: Oxford University Press.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states – The question is which ones and to what extent. *Trends in Cognitive Sciences*, *7*(4), 153–156.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–91; discussion 691–735.

Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*, 495–552.

van de Waal, E., Borgeaud, C., & Whiten, A. (2013). Potent social learning and conformity shape a wild primate's foraging decisions. *Science*, *340*, 483–485.

van de Waal, E., Bshary, R., & Whiten, A. (2014). Wild vervet monkey infants acquire the food-processing variants of their mothers. *Animal Behaviour*, *90*, 41–45.

van Leeuwen, E. J. C., Call, J., & Haun, D. B. M. (2014). Human children rely more on social information than chimpanzees do. *Biology Letters*, *10*, 20140487.

van Schaik, C. P. (1983). Why are diurnal primates living in groups? *Behaviour*, *87*, 120–147.

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., et al. (2003). Orangutan cultures and the evolution of material culture. *Science*, *299*, 102–105.

van Schaik, C. P., & Burkart, J. M. (2011). Social learning and evolution: The cultural intelligence hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1567), 1008–1016.

van Schaik, C. P., Deaner, R. O., & Merrill, M. Y. (1999). The conditions for tool use in primates: Implications for the evolution of material culture. *Journal of Human Evolution*, *36*(6), 719–741.

van Schaik, C. P., Isler, K., & Burkart, J. M. (2012). Explaining brain size variation: From social to cultural brain. *Trends in Cognitive Sciences*, *16*(5), 277–284.

Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences*, *11*, 233–273.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., et al. (1999). Cultures in chimpanzees. *Nature*, *399*(6737), 682–685.

Wilson, M. L., Boesch, C., Fruth, B., Furuichi, T., Gilby, I. C., Hashimoto, C., et al. (2014). Lethal aggression in *Pan* is better explained by adaptive strategies than human impacts. *Nature*, *513*(7518), 414–417.

Wittig, R. M., Crockford, C., Wikberg, E., Seyfarth, R. M., & Cheney, D. L. (2007). Kin-mediated reconciliation substitutes for direct reconciliation in female baboons. *Proceedings of the Royal Society B: Biological Sciences*, *274*, 1109–1115.

19 The Biological Basis of Intelligence

Richard J. Haier

Our knowledge about the biological basis of intelligence is growing rapidly, based on studies using genetics and neuroimaging methods to augment classic psychometrics. I have detailed major advances and new issues in *The Neuroscience of Intelligence* (Haier, 2017), but since that publication there is even more that is new and exciting to describe. This chapter is not a complete review of the literature but it will cover some key findings since that book manuscript was finished in late 2015.

There is overwhelming empirical evidence that intelligence is best described by a general factor that is common among all tests of mental ability and by specific factors like verbal, numerical, and spatial ability (Hunt, 2011). The general factor (*g*) alone accounts for about half the variance in IQ and other intelligence test scores. That is why the *g*-factor is the primary focus of most recent genetic and neuroimaging research. The *g*-factor cannot be measured in the same way as weight or distance (Haier, 2014). It can only be estimated, usually as a latent variable extracted from a battery of tests, and best interpreted for a person as a percentile compared to other individuals. Some individual tests are more *g*-loaded than others and single high *g*–loaded tests are often used in research for time/administration convenience. However, the best estimates of *g* come from batteries of tests like standard IQ tests (Johnson, te Nijenhuis, & Bouchard, 2008). Another useful distinction is between fluid and crystallized intelligence (Carroll, 1993); fluid intelligence is closely related to *g* and often the two are used interchangeably. Crystallized intelligence is more knowledge-based rather than novel reasoning.

Many studies are now using a proxy for *g* that is readily available for large representative samples being collected as part of multinational consortia, some of which are collecting DNA and neuroimaging data, on thousands and even over a million individual volunteers. Since it is not logistically feasible to collect cognitive test batteries or even a single test taking forty-five minutes on such large numbers of people, proxies of the *g*-factor are useful. The most widely used proxy used on large multinational samples recently is educational attainment, assessed as the number of years of schooling completed. It has an estimated correlation with

This chapter is an update of my original *Cambridge Handbook of Intelligence* chapter (Haier, 2011), my book *The Neuroscience of Intelligence* (Haier, 2017), and of two more recent updates written for undergraduates (Haier, 2018, 2019 in press). This chapter is written for more advanced students with an emphasis on important findings from the most recent studies through spring, 2018. These newest studies provide a road map for more in-depth investigations for those who may wish to do research on intelligence. For a historical perspective, explanations of neuroimaging methods, and a review of genetic and neuroimaging/intelligence research through 2015, the reader is referred to my book.

intelligence of about 0.6 (Rietveld et al., 2014; Rietveld et al., 2013). This is a measure of convenience and future studies would benefit from better assessments using latent variables based on multiple measures. This is so because latent measures are more robustly correlated with the *g*-factor, with correlations of 0.81 and 0.54, respectively, in samples of over 70,000 students (Deary et al., 2007), and over 105,000 combined participants from a meta-analysis of 240 samples (Roth et al., 2015). Of course, direct assessments of intelligence from multiple measures would also be optimal, if feasible, given the constraints of data collection in these collaborative projects. Importantly, proxy variables like years of education introduce more noise than latent variables, so the correlations with genetic assessments likely underestimate any actual relationships.

The collaborative success of these multinational consortia is a major story in the history of science. Because genes always work through biology mechanisms (even if there are some environmental influences on gene expression), the more we learn about the genetics of *g*, the more insights we get about the biological basis of intelligence. These mechanisms may be influenced by nongenetic factors, and the field of epigenetics hopes to identify such interactions. So far in humans, there is no compelling or independently replicated epigenetic findings related to intelligence. Partly, this is because advances will depend on identifying specific genes so that how they function can be determined along with any factors that influence their function. Because it is now believed that hundreds of genes influence intelligence (Krapohl et al., 2018; Plomin & von Stumm, 2018; Selzam et al., 2018), DNA from very large samples of people is required to detect the tiny effects of individual genes. Very large samples are also required to study individual differences in all aspects of intelligence because not all brains work the same way.

Simultaneously, a second major science story is based on advances in neuroimaging that are providing insights about intelligence and the brain. The most important methodological advance in recent years is the use of neuroimaging to assess connectivity of brain structure and function. Combined with large samples from consortia, these neuroimaging studies can test hypotheses about how the brain is organized and how it works during all kinds of cognition, including memory, learning, and general problem-solving/reasoning – all aspects of intelligence.

The genetic story and the neuroimaging story are intertwined, but there is another chapter in this handbook on genetic advances (Chapter 6). The reader is also referred to an excellent review of the genetic/DNA advances and polygenetic score predictions of intelligence (Plomin & von Stumm, 2018), which are creating a nascent molecular biology of intelligence. This chapter focuses on developments in neuroimaging studies, especially attempts to predict intelligence test scores from imaging data and what they may tell us about the biological basis of intelligence.

Before discussing these studies, however, a brief note is in order about a conceptual breakthrough for intelligence research in general. It came in the form of an editorial in *Nature* (2017) in the context of calling attention to a compelling genetic study of over 78,000 individuals that identified important associations between DNA and intelligence (Sniekers et al., 2017). The Editorial noted that “The subject [intelligence], it seems, is dying out on campus because it has echoes of elitism – and worse, racism – that make students and university officials uncomfortable. This is a distortion that contributes to widespread and wrong ideas about intelligence and the motives of those who study it. This is especially true when it comes to the genetics of intelligence . . . ” (p. 386). It went on to say, “What most people know about intelligence must be updated” (p. 386). In short, this editorial essentially removed an undeserved long-standing stigma from intelligence research and encouraged a new generation of researchers. We are now in a new exciting era of applying the most sophisticated neuroscience technologies to understanding intelligence beyond what we have learned from psychometrics.

A Short History

Before we get to the newest research, let us briefly look at some history of biological approaches to intelligence research so that the quickening pace of discovery can be appreciated. Two concepts are particularly important, brain efficiency and localization of intelligence centers in the brain. First, with respect to efficiency, before modern neuroimaging technology, EEG (electroencephalogram) studies reported modest correlations between electrical signals in the brain and various measures of intelligence (Jensen, 1998). There were early suggestions that high intelligence scores were related to more efficient brain processing (Chalke & Ertl, 1965; Ertl & Schafer, 1969; Schafer, 1982). However, different research groups used different EEG techniques and analysis methods so replications generally were not compelling. More recent reports lend support to interpreting EEG results as evidence for brain efficiency being associated with high intelligence in some circumstances, depending on sex and task difficulty (Neubauer & Fink, 2003, 2008, 2009). A more recent application of EEG methods to study fluid intelligence expanded the concept and assessment of brain efficiency to include flexible resource allocation as it relates to higher intelligence (Euler et al., 2015). Second, regarding localization of intelligence centers, early lesion work both in animals (Lashley, 1964; Thompson, Crinella, & Yu, 1990) and in humans (Duncan et al., 1996) suggested that intelligence was not a function of one specific brain region such as the frontal lobes. Newer human lesion data also show that the salient brain areas related to intelligence are linked in networks distributed throughout the brain (Barbey et al., 2014b; Barbey et al., 2012; Glascher et al., 2010; Glascher et al., 2009).

The first phase of neuroimaging studies of intelligence expanded the brain efficiency hypothesis (Haier et al., 1988) and supported a distributed model defined mostly by a parietal-frontal network, the parieto-frontal integration theory (P-FIT; Haier, 2009; Jung & Haier, 2007), as described in more detail in the section Updates on P-FIT and Brain Efficiency. These early studies were based on positron emission tomography (PET) and magnetic resonance imaging (MRI), but most of the studies had extremely small samples and relied on rudimentary image analysis methods. A second phase was based on far larger samples and used more mathematically sophisticated image analyses (Haier, 2009). These studies generally supported the distributed network findings, especially the P-FIT (Basten, Hilger, & Fiebach, 2015; Shehzad et al., 2014; Vakhtin et al., 2014), and, to some extent, they supported various interpretations of brain efficiency (Basten, Stelzel, & Fiebach, 2013; Cole et al., 2012; Li et al., 2009) although some have argued that the efficiency concept is too broad and vague (Poldrack, 2015).

New Research Predicting Intelligence from Brain Connectivity

Prediction is a major goal in science. Sometimes, accurate empirical predictions of something can be made even if there is limited understanding of why the prediction works. Over time, predictions usually become more accurate and the "why" is discovered, often based on clues provided by the variables used to make the prediction. Attempts to predict intelligence from brain images started with the earliest neuroimaging studies, but none of the attempts resulted in much success. There were many problems. Multiple-regression equations failed independent cross-validation and replication. Most of the attempts were based on group data with relatively small sample sizes and considerable individual differences. To increase statistical power, males and females often were combined along with wide age ranges. Age and sex typically were used as nuisance variables to statistically remove, but this removal masked any age or sex differences that might exist in associations between brain characteristics and intelligence. Individual differences were treated as error variance in many of these attempts. These early failures are reviewed in detail in Haier's book (2017, pp. 118–124). That review ended with a brief optimistic note about a paper just published at the deadline for submitting the final draft of the book. It reported something quite amazing.

The new paper (Finn et al., 2015) reported data from the Human Connectome Project, a consortium with the aim of mapping all connections in the human brain using neuroimaging data and graph analysis, an important method for advanced neuroimaging analysis. I described graph analysis this way (Haier, 2017, p. 102):

> Graph analysis is a mathematical tool that is used to model brain connectivity and infer networks. The idea is to establish how each voxel in a brain image is correlated to all other voxels throughout the brain. These connections, called edges, can be computed for structural or functional images. A voxel, or a cluster of voxels, that show correlations to many other voxels is called a hub. Hubs that show correlations to many other hubs are called rich clubs. The strength of any connection is determined by the magnitude of the correlation between voxels or hubs. The efficiency of any connection can be estimated by determining its length. Most of the brain has local connectivity in that many nearby voxels are connected to each other via a neighborhood hub. This makes for efficient information transfer. Rich clubs connect more distant brain areas and this makes for faster communication . . . Psychometric test scores can be correlated to the strength of hubs and connections to indicate which brain networks are related to intelligence.

There are several early examples of using graph analysis to investigate intelligence (Cole et al., 2012; Langer et al., 2012; Santarnecchi et al., 2014; Song et al., 2009; van den Heuvel & Sporns, 2011). Please note that graph analyses come in different varieties and are quite complex – see Liao, Vasilakos, and He (2017) for an excellent overview. Only summaries of selected findings are noted here, often by quoting the original papers.

In this case, Finn and colleagues (2015) used functional magnetic resonance imaging (fMRI) data obtained from 126 participants. Each completed six imaging sessions that included four task conditions and two rest conditions. The data analyses were unusual in that they focused on individual differences. Instead of averaging all participants together and comparing functional connectivity patterns during a task condition to other task conditions and to rest conditions, these researchers computed a connectivity analysis for each person individually for each condition. These determinations were based on a set of 268 voxels (called nodes in graph analysis) that included ten predefined networks.

They then determined how stable an individual's pattern was for each of the task and rest conditions. Surprisingly, they reported that the patterns were stable across all six conditions and that these patterns were so unique to individuals, they were like fingerprints (I resist the temptation to call them brain-fingerprints so I refer to them as brain-prints). The researchers also reported these brain-prints predicted fluid intelligence test scores (up to a correlation of 0.50), with the strongest predictions for connections (edges) in frontoparietal networks, consistent with the parieto-frontal integration theory (P-FIT) of brain/intelligence relationships (Jung & Haier, 2007); more about the P-FIT in the section Updates on P-FIT and Brain Efficiency. The final good news in this paper was the inclusion of cross-validation analyses. In my book, I expressed enthusiasm for this "landmark" study and quoted the authors, "These results underscore the potential to discover fMRI-based connectivity 'neuromarkers' of present or future behavior that may eventually be used to personalize educational and clinical practices and improve outcomes." They conclude:

> Together, these findings suggest that analysis of individual fMRI data is possible and indeed desirable. Given this foundation, human neuroimaging studies have an opportunity to move beyond population-level inferences, in which general networks are derived from the whole sample, to inferences about single subjects, examining how individuals' networks are functionally organized in unique ways and relating this functional organization to behavioral phenotypes in both health and disease. (Haier, 2017, p. 182)

Was my enthusiasm for this paper warranted? I try to follow three laws: No story about the brain is simple; no one study is definitive; it takes time to establish a compelling weight of evidence. It is still early with respect to connectivity brain-prints and whether they reliably predict intelligence test scores. In a brief paper published soon after Finn and colleagues (2015), Biazoli and colleagues (2017) computed individual connectivity patterns from fMRIs obtained from 655 children. They found the patterns were reliable and changed with age (based on cross-sectional data), and children with similar patterns also had similar intelligence test scores, although the correlations were small. A more elaborate study, also using Human Connectome data, reported resting fMRI connectivity patterns in 105 individuals based on dynamic spatial network parameters (Liu et al., 2018). Consistent with Finn and colleagues, they found the patterns uniquely identified individuals and that individual patterns were stable. They reported a correlation of 0.42 ($p < 0.0001$, corrected) between predicted fluid intelligence and actual test scores. Frontoparietal connections along with the default network were key. Similarly, another group used resting state fMRI Connectome data on 309 individuals and reported that connectivity profiles also were related to intelligence (Hilger et al., 2017b). The findings emphasized specific modular connections, especially among frontal and parietal regions, rather than global brain features (e.g., number or size of modules). An even larger study of resting state fMRI obtained from 884 Human Connectome young adults found that 20 percent of variance in general intelligence was accounted for by connectivity profiles, especially among P-FIT areas (Dubois et al., 2018). These studies used different parameters for establishing connectivity patterns than Finn and colleagues, and they demonstrate that we are at the beginning of exploring how variations in the pattern analysis methods might improve predictions of intelligence scores. Importantly, there is growing appreciation and evidence that there are reliable individual differences among brains both structurally (Guntupalli, Feilong, & Haxby, 2018; Valizadeh et al., 2018) and functionally, and that these differences are related to individual differences in mental abilities as noted by many early neuroimaging studies (see Haier, 2017).

The prediction of intelligence test scores is advancing beyond multiple-regression analyses based on group data, and some optimism is warranted. Correlations of 0.5 and 0.42 from the two Human Connectome studies are impressive but more independent replication is needed. In my view, connectivity analyses that characterize individual brain patterns like fingerprints hold considerable potential. Given the advances in predicting intelligence scores from DNA-based polygenetic scores, currently about 10 percent of variance (Plomin & von Stumm, 2018), it will be of interest to see how well intelligence can be predicted from a combination of quantified brain connectivity patterns and polygenetic scores.

Updates on P-FIT and Brain Efficiency

It is also noteworthy that these Human Connectome studies reported findings consistent with the P-FIT, which identified fourteen Brodmann areas (BA) related to general intelligence. It was proposed as a framework to test hypotheses about brain networks, efficiency, and intelligence (Jung & Haier, 2007). A number of papers have tested P-FIT hypotheses and were reviewed previously (Haier, 2017). This chapter updates that review with papers published subsequently. For example, using a different approach to connectivity based on a priori regions of interest, one study used Human Connectome data on resting state fMRI obtained from 317 adults (Hearne, Mattingley, & Cocchi, 2016). Their aim was to test P-FIT predictions, and they included non-P-FIT areas that defined other networks for investigation. The results supported the P-FIT network but also a role for the default network. Here is how the authors summarized their findings:

> While our findings confirm a key role for fronto-parietal networks in supporting intelligence, they also highlight the importance of connectivity between regions associated with the fronto-parietal, default-mode and regions not strongly associated with homogeneous networks (although these regions have been identified as comprising the default-mode network before), particularly in the prefrontal cortex. More broadly, our results suggest that interactions between fronto-parietal and default-mode networks are important for explaining individual differences in intelligence in a state of rest . . . Specifically, individuals with higher intelligence deactivate the default-mode network less . . . and activate fronto-parietal and cingulo-opercular network regions more than individuals with lower intelligence . . . The functional links between transitions from diffuse resting state dynamics and more segregated task dynamics and intelligence will be an important topic for ongoing research. (p. 3)

In addition to a model of networks distributed across the brain relevant for intelligence, another aspect of the P-FIT is the general idea that individual differences in intelligence are related to efficient communication within and among salient networks. This efficiency hypothesis originated from an early PET study that showed inverse correlations between intelligence test scores and glucose metabolic rate in several brain regions (Haier et al., 1988). Some early imaging evidence supported such a relationship, although results are mixed (Haier, 2017). Stronger evidence came from a study that quantified global efficiency for the whole brain from brain images (van den Heuvel et al., 2009). One recent study, however, suggests efficiency is more related to intelligence in some specific networks. This study was based on resting state fMRI obtained in fifty-four adults (Hilger et al., 2017a). Connectivity was assessed with graph analyses that included determinations of both global and nodal efficiency. Nodal efficiency is based on the shortest distances among nodes within a network (defined a priori). Global efficiency is an average of all nodal efficiency values. Global efficiency was not correlated to the intelligence measure (Full Scale IQ). Nodal efficiency for three specific brain areas was correlated to intelligence. The anterior insula (AI) and the dorsal anterior cingulate cortex (dACC) showed positive correlations; the temporoparietal junction (TPJ) showed a negative correlation. Here is the authors' summary:

> Our findings together with other studies on intelligence and intrinsic brain connectivity . . . extend these conclusions by demonstrating that intrinsic functional connectivity properties of the brain's network organization may play a key role in understanding the neural underpinnings of intelligence. Specifically, our analyses imply that with respect to network topology, brain regions that were previously related to salience processing (AI and dACC) and the filtering of irrelevant information from further processing (TPJ), play a crucial role in explaining individual differences in intelligence. We speculate that the observed differences in network integration of these three regions may enable intelligent people to more quickly detect, evaluate, and mark salient new stimuli for further processing and to protect ongoing cognitive processing from interference of irrelevant information, ultimately contributing to higher cognitive performance and high intelligence. (p. 22)

Another important comprehensive study also failed to replicate correlations between any global efficiency measures and either general intelligence, fluid intelligence, or crystallized intelligence (Kruschwitz et al., 2018). They used a Human

Connectome sample of 999 to1096 adults (not all participants completed all measures) and graph analyses on resting state fMRI. These authors discuss their results in the context of previous efficiency/intelligence studies. They conclude, “it is likely that the low power of previous studies, together with a publication bias, may have led to publication of false positive results that have fostered the widely accepted notion of general intelligence being associated with functional global network efficiency” (p. 330). Although this is a failure to replicate the study by van den Heuvel and colleagues (2009), the efficiency hypothesis of intelligence may still be viable. There are new methods to quantify efficiency (Liao et al., 2017; Liu et al., 2018), and whether efficiency of specific networks may be related to intelligence, or whether other variables such as age or sex must be considered in separate analyses of global efficiency, is still unclear. Also, most of the connectivity studies so far have used resting state fMRI; there may be task conditions that produce efficiency relationships with intelligence scores. And, intriguingly, efficiency may have origins at the neuron level. Here are some examples.

In the case of efficiency of neurons, an impressive study by Goriounova and colleagues (2018) collected data from adults undergoing brain surgery for epilepsy or tumor ($N = 23$ to 37; not every person had each measure). There were presurgical MRIs, Full Scale IQ testing, single-cell physiology and morphology determinations, and recorded action potentials from pyramidal neurons of temporal cortical tissue. The results were summarized by the authors:

> we find that high IQ scores and large temporal cortical thickness associate with larger, more complex dendrites of human pyramidal neurons. We show in silico that larger dendritic trees enable pyramidal neurons to track activity of synaptic inputs with higher temporal precision, due to fast action potential kinetics. Indeed, we find that human pyramidal neurons of individuals with higher IQ scores sustain fast action potential kinetics during repeated firing. These findings provide the first evidence that human intelligence is associated with neuronal complexity, action potential kinetics and efficient information transfer from inputs to output within cortical neurons. (p. 1)

These are preliminary findings but they illustrate how neuroscience methods at the neuron level can be applied to intelligence research.

Another remarkable study in a much larger sample also found that dendritic density and arborization in gray matter are correlated with intelligence (Genc et al., 2018), but there is an apparent difference between these two studies regarding dendrites. These researchers used a new diffusion MRI technique called neurite orientation dispersion and density imaging (NODDI); this is its first reported use in the human brain. There were two independent samples; 259 adults were in the experimental sample and 498 adults from the Human Connectome data set were in the validation sample. Overall, the findings were summarized by the authors: “we found that higher intelligence in healthy individuals is related to lower values of dendritic density and arborization. These results suggest that the neuronal circuitry associated with higher intelligence is organized in a sparse and efficient manner, fostering more directed information processing and less cortical activity during reasoning” (p. 1). Moreover, separate analyses in both samples noted findings were in several P-FIT areas.

These two studies were available about the same time so neither one cites the other. The apparent contradiction between them – more complex dendrites related to higher intelligence in Goriounova and colleagues (2018), and lower dendrite density and arborization related to higher intelligence in Genc and colleagues (2018), demonstrates my three laws: No story about the brain is simple; no one study is definitive; it takes time to sort out inconsistent findings and establish a weight of evidence. I queried both lead authors about this and they noted differences in methodology of dendritic assessments, and how intelligence was assessed. Also, Goriounova and colleagues (2018) focused only on certain cells in temporal cortex, but Genc and colleagues did note some positive associations between dendrite variables in temporal areas and higher intelligence in their sample, although they were not significant after correction for multiple comparisons. Interestingly, in this case, despite some apparently opposite results, both findings were interpreted as evidence for the efficiency hypothesis.

In another example, a structural MRI study discussed efficiency based on correlations between greater cortical gyrification and working memory (a key component of intelligence) in P-FIT areas for forty-eight adults (Green et al., 2018). They concluded

> Cortical folding shortens the distance between white matter tracts of adjacent brain regions, which increases signaling speed throughout these regions, thus aiding in rapid working memory functions. Recent studies of whole brain connectivity have reflected similar findings in the frontoparietal network in terms of fluid intelligence (Finn et al. 2015). Thus, gyrification may enhance local connectivity and reflect a "small-world network" solution to synchronously connect widespread cortical regions; thereby improving the efficiency and robustness of information processing. (p. 307)

Efficiency is actually a popular concept/speculation/explanation for many findings across numerous studies so it will take time to sort out whether any form of an efficiency hypothesis is viable for intelligence research at the level of specific brain areas, networks, neurons, or synapses.

In the case of sex differences, one study investigated whether gray matter volume and white matter efficiency differentially predicted intelligence separately for males and females (Ryman et al., 2016). This is an interesting question because males and females generally show no average difference on IQ measures even though males have, on average, larger brains. This has led to speculation that females may have more efficient brains. This study used graph analyses and structural equation modeling in 244 adults who completed a battery of cognitive tests that assessed general cognitive ability (GCA; essentially a *g*-factor). Here is the authors' summary of results:

> Results indicated that in males, a latent factor of fronto-parietal gray matter was significantly related to GCA when controlling for total gray matter volume. In females, white matter efficiency and total gray matter volume were significantly related to GCA, with no specificity of the fronto-parietal gray matter factor over and above total gray matter volume. This work highlights that different neural characteristics across males and females may contribute to performance on intelligence measures. (p. 4006)

This also demonstrates the necessity of computing separate analyses for males and females in every study.

Information Processing

The P-FIT also characterizes four stages of information processing in networks across the specific brain areas (Colom et al., 2010) In the first stage, temporal and occipital areas process sensory information: the extrastriate cortex (BAs 18 and 19) and the fusiform gyrus (BA 37), involved with recognition, imagery, and elaboration of visual inputs, as well as the Wernicke's area (BA 22) for analysis and elaboration of syntax of auditory information. The second stage implicates integration and abstraction of this information by parietal BAs 39 (angular gyrus), 40 (supra-marginal gyrus), and 7 (superior parietal lobule). In the third stage, these parietal areas interact with the frontal lobes, which serve to problem-solve, evaluate, and hypothesis test. Frontal BAs 6, 9, 10, 45, 46, and 47 are prominent. In the final stage, the anterior cingulate (BA 32) is implicated for response selection and inhibition of alternative responses, once the best solution is determined in the previous stage. White matter, especially the arcuate fasciculus, plays a critical role for reliable communication of information among these processing units.

Ponsoda and colleagues (2017) used a different connectivity method called multivariate distance matrix regression (MDMR), which alleviates aspects of multiple comparisons (Shehzad et al., 2014). Structural MRIs were obtained on ninety-four young adults along with a battery of cognitive tests that tapped fluid and crystallized intelligence, and spatial ability along with working memory, attention control, and processing speed. They specifically tested connections among P-FIT areas and the cognitive factors, including cross-validation procedures. Overall, the results showed:

> that individuals with similar brain connectivity profiles are also closer in their cognitive level as estimated by fluid, crystallized, and spatial ability latent factors. Furthermore, we identified a subset of 36 linkages connecting distributed brain regions that increased more than twice the predictive power of cognitive performance on brain profiles. Working memory capacity, attention, and processing speed were not significantly related with similarities among individuals in their brain connectivity profiles, suggesting that the observed joint covariation between biological and psychological data cannot be simply generalized across cognitive domains. (pp. 811–812)

With respect to the P-FIT, the authors noted that their findings supported the four-stage information-processing model (see their figure 5). They concluded, "individual differences in three key latent cognitive factors estimating fluid, crystallized, and spatial ability, predicted similarities among individuals regarding a structural network defined by 36 connections among a set of brain regions. Working memory capacity, attention control, and processing speed were cognitive factors unrelated with the identified network" (pp. 814–815).

In addition to individual neuroimaging studies of intelligence inspired in part by the P-FIT, an important meta-analysis of forty-seven PET and fMRI studies addresses verbal and visuospatial components of fluid intelligence (G*f*) as well as salient stages of cognitive processing, and effects of task complexity on network activations (Santarnecchi, Emmendorfer, & Pascual-Leone, 2017a). A variety of statistical methods are used on neuroimaging data, including activation likelihood estimates (ALE) to determine brain voxels that are more active than expected by chance. The findings are extensive and the discussion is comprehensive. The authors conclude:

> Results highlight the loading of Gf components over functionally defined resting-state fMRI networks, with different degrees of overlap in both hemispheres and subcortical structures. A major role for nodes of the dorsal attention network during both verbal and visuospatial abstract reasoning tasks represents the most consistent correlate of Gf, with additional contributions by regions of the anterior salience and left frontoparietal control network. Increase in trial difficulty elicits a more pronounced engagement of the language and left fronto-parietal control networks, while inferring the rules subtending a given Gf task relies on a different anatomo-functional substrate than producing novel solutions. Current findings might allow a clearer association between Gf-related activity and brain connectivity, also providing quantitative ALE maps to be used in network-based brain stimulation and cognitive training interventions. (p. 9)

This study extends the P-FIT and, in my view, offers a rich conceptualization of brain/intelligence relationships for testing hypotheses.

In fact, the same research group also published an elaborate quantitative validation of the previously inferred overlap between resting state functional connectivity networks and the networks related to fluid intelligence (Santarnecchi, Emmendorfer, Tadayon et al., 2017b). They studied resting state fMRIs obtained from 130 adults:

> Results highlight a striking degree of similarity between the connectivity profile of the gf network and that of the dorsal attention network, with additional overlap with the left and right fronto-parietal control networks. Interestingly, a strong negative correlation with structures of the default mode network (DMN) was also identified. Results of regression models built on two independent fMRI datasets confirmed the negative correlation between gf regions and medial prefrontal structures of the DMN as a significant predictor of individual gf scores. These might suggest a framework to interpret previously reported aging-related decline in both gf and the correlation between "task-positive" networks and DMN, possibly pointing to a common neurophysiological substrate. (p. 35)

Finally, the combination of structure and functional imaging in the same subjects is an important advance. For example, if one group has more gray matter in a brain area than another group, it might be hypothesized that more efficient function would be found in that area owing to more available neurons. However, structural and functional findings do not necessarily overlap (Haier et al., 2009). Genc and colleagues (poster presented at the International Society for Intelligence Research 2017 annual meeting in Montreal) studied eighty-five adults tested on fluid intelligence and both functional and structural MRI (including diffusion tensor imaging; DTI) connectivity to examine P-FIT areas. They reported intelligence correlations in P-FIT areas for fractional anisotropy (FA, a measure of white matter integrity from DTI) in intra- and interhemispheric white matter tracts ($r = 0.22$, 0.21, respectively; $p < 0.05$). For functional connectivity, interhemispheric coherence was correlated to intelligence in four P-FIT areas ($r = 0.23$ to 0.34, $p < 0.05$), and also for intrahemispheric coherence in three pathways ($r = 0.22$ to 0.34, $p < 0.05$).

All these connectivity, efficiency, and information-processing papers provide some of the most sophisticated investigations to date of brain/intelligence relationships. They go beyond the earlier descriptive identification of brain areas and networks associated with measures of intelligence, like those in the P-FIT, and provide hints about how structure and function are related to the neural/biological basis of intelligence. As always, independent replications of these fascinating results are required, but they already define some exciting new directions for the neuroscience study of human intelligence and individual differences ever deeper into the brain.

A New Era?

As we push inexorably deeper into the brain from cortex to neurons to synapses, we are now at the threshold of developing a molecular biology of intelligence based both on gene expression related to brain development and function, and on the cascades of neurobiological events at the neuron and synapse levels, whether influenced by genes or not. Genetic data likely will provide important avenues of research as intelligence-related genetic expressions are identified. Glutamate neurotransmitter pathways and NMDA receptors, for example, have been implicated (Hill et al., 2014; Rietveld et al., 2014) and so has brain-derived neurotrophic factor (BDNF; Barbey, Colom, Paul, Forbes et al., 2014a). One form of protein, DUF1220, has been associated with high IQ scores (Davis et al., 2015), but so far all the specific genetic associations account for only tiny amounts of intelligence variance. One gene-network analysis constructed "IQ-related pathways" based on 158 genes associated with intelligence as reported in various studies (Zhao, Kong, & Qu, 2014). The resulting pathways involved dopamine and norepinephrine, neurotransmitters involved in many brain functions. Dopamine continues to be of interest (Alavash et al., 2018; Wang et al., 2018). Molecular studies of intelligence are still in a nascent stage. Understanding the molecular biology of intelligence should offer possibilities for tweaking the salient systems with the ultimate aim of increasing general intelligence and, perhaps, other specific intelligence factors. This would be finding the Holy Grail of intelligence research.

Since we are focusing on neuroimaging in this chapter, it is relevant that some hints about the neurobiology of intelligence came from early use of MRI spectroscopy (MRS). These studies found correlations between N-acetyl aspartate (NAA), a biochemical marker of neuron health based on neuron energy production and efficiency, and intelligence test scores (Jung et al., 1999; Jung et al., 2009; Jung et al., 2005). A recent study using MRI and MRS in 211 adults found that NAA concentration correlated with verbal/spatial reasoning, and brain volume correlated with quantitative reasoning and with working memory (Paul et al., 2016). The authors concluded "that NAA and brain volume are independent predictors of verbal/spatial and quantitative facets of Gf" (p. 201). In another report, this research group also found that NAA in frontoparietal areas was related to fluid intelligence, consistent with the P-FIT (Nikolaidis et al., 2017).

These papers also imply another perspective on the definition of intelligence. In all these papers, intelligence measures are the dependent variable and brain measures are the independent variables. Could this be reversed? It is noteworthy that one side of the prediction is based on cutting-edge multimillion-dollar advanced technology and teams of multidisciplinary specialists, and the other side is based on the same paper-and-pencil/psychometric tests that have been used for decades. Is it possible that intelligence might be defined by brain characteristics instead of psychometrics? Brain processing speed has been proposed as an alternative way to define intelligence with the quantitative advantage that reaction time is a ratio scale compared to the ordinal scale of psychometric measures (Jensen, 2006). This would be a technically formidable project and not much research addresses this goal. Alternatively, the idea of brain-prints and other imaging analyses offer a possibility for defining intelligence based on a brain profile or type. Even polygenetic scores or DNA profiles might provide new definitions of intelligence. These categories could be validated against learning ability, memory performance, and speed of information processing. All these possibilities are made somewhat realistic for research because of public access to the large databases collected by multisite consortia. They may even become economically realistic, because neuroimaging for individuals in many facilities is already less expensive than formal IQ testing by a psychologist and less expensive than SAT review classes. If brain-prints, types, or profiles eventually predict academic or life success as well as, or better than, traditional IQ tests, the SATs, or similar tests, we are in a new era of brain-based intelligence research and applications. There is a long way to go in this direction and all the usual reliability and validity questions will need attention. Direct assessments of brain characteristics may prove better for some applications than indirect inferences from current psychometric tests.

Speaking of a new era, as prediction advances and the biological mechanisms underlying intelligence are identified, a major step will be manipulation of those mechanisms to enhance intelligence. We already can enhance some mental performance in limited if not always healthy ways (e.g., caffeine, amphetamines) but so far there is no proven way to increase general intelligence or its specific factors (see review in Haier, 2017), although this goal is a prime topic (Colom & Roman, 2018; Santarnecchi & Rossi, 2016). As of this writing, this goal is still unreached, but all the studies reviewed here suggest that the goal is not unreachable. That is why the study of intelligence has never been more exciting, as biological approaches take us beyond psychometrics and deep into the brain. And that is worthy of a career.

References

Alavash, M., Lim, S. J., Thiel, C., Sehm, B., Deserno, L., & Obleser, J. (2018). Dopaminergic modulation of hemodynamic signal variability and the functional connectome during cognitive performance. *Neuroimage*, *172*, 341–356. https://doi.org/10.1016/j.neuroimage.2018.01.048

Barbey, A. K., Colom, R., Paul, E., Forbes, C., Krueger, F., Goldman, D., & Grafman, J. (2014a). Preservation of general intelligence following traumatic brain injury: Contributions of the Met66 brain-derived neurotrophic factor. *PLoS One*, *9*(2). https://doi.org/10.1371/journal.pone.0088733

Barbey, A. K., Colom, R., Paul, E. J., & Grafman, J. (2014b). Architecture of fluid intelligence and working memory revealed by lesion mapping. *Brain Structure and Function*, *219*(2), 485–494.

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., & Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain*, *135*(Pt 4), 1154–1164. https://doi.org/10.1093/brain/aws021

Basten, U., Hilger, K., & Fiebach, C. J. (2015). Where smart brains are different: A quantitative meta-analysis of functional and structural brain imaging studies on intelligence. *Intelligence*, *51*, 10–27. http://dx.doi.org/10.1016/j.intell.2015.04.009

Basten, U., Stelzel, C., & Fiebach, C. J. (2013). Intelligence is differentially related to neural effort in the task-positive and the task-negative brain network. *Intelligence*, *41*(5), 517–528.

Biazoli, C. E., Jr., Salum, G. A., Pan, P. M., Zugman, A., Amaro, E., Jr., Rohde, L. A., et al. (2017). Commentary: Functional connectome fingerprint: Identifying individuals using patterns of brain connectivity. *Frontiers in Human Neuroscience*, *11*, 47. https://doi.org/10.3389/fnhum.2017.00047

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Chalke, F. C., & Ertl, J. (1965). Evoked potentials and intelligence. *Life Sciences*, *4*(13), 1319–1322.

Cole, M. W., Yarkoni, T., Repovs, G., Anticevic, A., & Braver, T. S. (2012). Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *Journal of Neuroscience*, *32*(26), 8988–8999. https://doi.org/10.1523/JNEUROSCI.0536-12.2012

Colom, R., Karama, S., Jung, R. E., & Haier, R. J. (2010). Human intelligence and brain networks. *Dialogues in Clinical Neuroscience*, *12*(4), 489–501.

Colom, R., & Roman, F. J. (2018). Enhancing intelligence: From the group to the individual. *Journal of Intelligence*, *6*(11). https://doi.org/10.3390/jintelligence6010011

Davis, J. M., Searles, V. B., Anderson, N., Keeney, J., Raznahan, A., Horwood, L. J., et al. (2015). DUF1220 copy number is linearly associated with increased cognitive function as measured by total IQ and mathematical aptitude scores. *Human Genetics*, *134*(1), 67–75. https://doi.org/10.1007/s00439-014-1489-2

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21.

Dubois, J., Galdi, P., Paul, L. K., & Adolphs, R. (2018). A distributed brain network predicts general intelligence from resting-state human neuroimaging data. *Philosophical Transactios of the Royal Society B*. https://doi.org/10.1098/rstb.2017.0284

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, *30* (3), 257–303.

Nature (2017). Intelligence test (editorial). *545*, 385–386.

Ertl, J. P., & Schafer, E. W. (1969). Brain response correlates of psychometric intelligence. *Nature*, *223*(204), 421–422.

Euler, M. J., Weisend, M. P., Jung, R. E., Thoma, R. J., & Yeo, R. A. (2015). Reliable activation to novel stimuli predicts higher fluid intelligence. *Neuroimage*, *114*, 311–319. https://doi.org/10.1016/j.neuroimage.2015.03.078

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., et al. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience*, *18*(11), 1664–1671. https://doi.org/10.1038/nn.4135

Genc, E., Fraenz, C., Schluter, C., Friedrich, P., Hossiep, R., Voelkle, M. C., et al. (2018). Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence. *Nature Communications*, *9*(1), 1905. https://doi.org/10.1038/s41467-018-04268-8

Glascher, J., Rudrauf, D., Colom, R., Paul, L. K., Tranel, D., Damasio, H., & Adolphs, R. (2010). Distributed neural system for general intelligence revealed by lesion mapping. *Proceedings of the National Academy of Sciences*, *107*(10), 4705–4709. https://doi.org/10.1073/Pnas.0910397107

Glascher, J., Tranel, D., Paul, L. K., Rudrauf, D., Rorden, C., Hornaday, A., et al. (2009). Lesion mapping of cognitive abilities linked to intelligence. *Neuron*, *61*(5), 681–691. https://doi.org/10.1016/j.neuron.2009.01.026

Goriounova, N., Heyer, D. B., Wilbers, R., Verhoog, M. B., Giugliano, M., Verbist, C., et al. (2018). Large and fast human pyramidal neurons associate with intelligence. *eLife*, *7*, e41714, https://doi.org/10.7554/eLife.41714

Green, S., Blackmon, K., Thesen, T., DuBois, J., Wang, X. Y., Halgren, E., & Devinsky, O. (2018). Parieto-frontal gyrification and working memory in healthy adults. *Brain Imaging and Behavior*, *12*(2), 303–308.

Guntupalli, J. S., Feilong, M., & Haxby, J. V. (2018). A computational model of shared fine-scale structure in the human connectome. *PLoS Computational Biology*, *14*(4), e1006120. https://doi.org/10.1371/journal.pcbi.1006120

Haier, R. J. (2009). Neuro-intelligence, neuro-metrics and the next phase of brain imaging studies. *Intelligence*, *37*(2), 121–123. https://doi.org/10.1016/j.intell.2008.12.006

Haier, R. J. (2011). Biological basis of intelligence. In R. J. Sternberg & A. S. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 351–368). Cambridge, UK: Cambridge University Press.

Haier, R. J. (2014). Increased intelligence is a myth (so far). *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00034

Haier, R. J. (2017). *The neuroscience of intelligence*. New York: Cambridge University Press.

Haier, R. J. (2018). A view from the brain. In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 167–182). New York: Cambridge University Press.

Haier, R. J. (2019 in press). Biological approaches to intelligence. In R. J. Sternberg (Ed.), *Human intelligence: An introduction* (pp. 139–173). New York: Cambridge University Press.

Haier, R. J., Karama, S., Leyba, L., & Jung, R. E. (2009). MRI assessment of cortical thickness and functional activity changes in adolescent girls following three months of practice on a visual-spatial task. *BMC Research Notes*, *2*. https://doi.org/10.1186/1756–0500-2–174

Haier, R. J., Siegel, B. V., Nuechterlein, K. H., Hazlett, E., Wu, J. C., Paek, J., et al. (1988). Cortical glucose metabolic-rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence*, *12*(2), 199–217.

Hearne, L. J., Mattingley, J. B., & Cocchi, L. (2016). Functional brain networks related to individual differences in human intelligence at rest. *Scientific Reports*, *6*, 32328. https://doi.org/10.1038/srep32328

Hilger, K., Ekman, M., Fiebach, C. J., & Basten, U. (2017a). Efficient hubs in the intelligent brain: Nodal efficiency of hub regions in the salience network is associated with general intelligence. *Intelligence*, *60*, 10–25. https://doi.org/10.1016/j.intell.2016.11.001

Hilger, K., Ekman, M., Fiebach, C. J., & Basten, U. (2017b). Intelligence is associated with the modular structure of intrinsic brain networks. *Scientific Reports*, *7*(1), 16088. https://doi.org/10.1038/s41598-017-15795-7

Hill, W. D., Davies, G., van de Lagemaat, L. N., Christoforou, A., Marioni, R. E., Fernandes, C. P., et al. (2014). Human cognitive ability is influenced by genetic variation in components of postsynaptic signalling complexes assembled by NMDA receptors and MAGUK proteins. *Translational Psychiatry*, *4*, e341. https://doi.org/10.1038/tp.2013.114

Hunt, E. B. (2011). *Human intelligence*. Cambridge, UK: Cambridge University Press.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. New York: Elsevier.

Johnson, W., te Nijenhuis, J., & Bouchard, T. J. (2008). Still just 1 g: Consistent results from five test batteries. *Intelligence*, *36*(1), 81–95.

Jung, R. E., Brooks, W. M., Yeo, R. A., Chiulli, S. J., Weers, D. C., & Sibbitt, W. L. (1999). Biochemical markers of intelligence: A proton MR spectroscopy study of normal human brain. *Proceedings of the Royal Society B: Biological Sciences*, *266*(1426), 1375–1379.

Jung, R. E., Gasparovic, C., Chavez, R. S., Caprihan, A., Barrow, R., & Yeo, R. A. (2009). Imaging intelligence with proton magnetic resonance spectroscopy. *Intelligence*, *37*(2), 192–198.

Jung, R. E., & Haier, R. J. (2007). The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(2), 135–154; discussion 154–187. https://doi.org/10.1017/S0140525X07001185

Jung, R. E., Haier, R. J., Yeo, R. A., Rowland, L. M., Petropoulos, H., Levine, A. S., et al. (2005). Sex differences in N-acetylaspartate correlates of general intelligence: An H-1-MRS study of normal human brain. *Neuroimage*, *26*(3), 965–972.

Krapohl, E., Patel, H., Newhouse, S., Curtis, C. J., von Stumm, S., Dale, P. S., et al. (2018). Multi-polygenic score approach to trait prediction. *Molecular Psychiatry*, *23*(5), 1368–1374. https://doi.org/10.1038/mp.2017.163

Kruschwitz, J. D., Waller, L., Daedelow, L. S., Walter, H., & Veer, I. M. (2018). General, crystallized and fluid intelligence are not associated with functional global network efficiency: A replication study with the human connectome project 1200 data set. *Neuroimage*. https://doi.org/10.1016/j.neuroimage.2018.01.018

Langer, N., Pedroni, A., Gianotti, L. R., Hanggi, J., Knoch, D., & Jancke, L. (2012). Functional brain network efficiency predicts intelligence. *Human Brain Mapping*, *33*(6), 1393–1406. https://doi.org/10.1002/hbm.21297

Lashley, K. S. (1964). *Brain mechanisms and intelligence*. New York: Hafner.

Li, Y., Liu, Y., Li, J., Qin, W., Li, K. C., Yu, C. S., et al. (2009). Brain anatomical network and intelligence. *PLoS Computational Biology*, *5*(5). https://doi.org/10.1371/journal.pcbi.1000395

Liao, X., Vasilakos, A. V., & He, Y. (2017). Small-world human brain networks: Perspectives and challenges. *Neuroscience and Biobehavioral Reviews*, *77*, 286–300.

Liu, J., Liao, X., Xia, M., & He, Y. (2018). Chronnectome fingerprinting: Identifying individuals and predicting higher cognitive functions using dynamic brain connectivity patterns. *Human Brain Mapping*, *39*(2), 902–915. https://doi.org/10.1002/hbm.23890

Neubauer, A. C., & Fink, A. (2003). Fluid intelligence and neural efficiency: Effects of task complexity and sex. *Personality and Individual Differences*, *35*(4), 811–827.

Neubauer, A. C., & Fink, A. (2008). Intelligence and neural efficiency: A review and new data. *International Journal of Psychophysiology*, *69*(3), 168–169.

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency: Measures of brain activation versus measures of functional connectivity in the brain. *Intelligence*, *37* (2), 223–229.

Nikolaidis, A., Baniqued, P. L., Kranz, M. B., Scavuzzo, C. J., Barbey, A. K., Kramer, A. F., et al. (2017). Multivariate associations of fluid intelligence and NAA. *Cerebral Cortex*, *27*(4), 2607–2616. https://doi.org/10.1093/cercor/bhw070

Paul, E. J., Larsen, R. J., Nikolaidis, A., Ward, N., Hillman, C. H., Cohena, N. J., et al. (2016). Dissociable brain biomarkers of fluid intelligence. *Neuroimage*, *137*, 201–211. https://doi.org/10.1016/j.neuroimage.2016.05.037

Plomin, R., & von Stumm, S. (2018). The new genetics of intelligence. *Nature Reviews Genetics*, *19*(3), 148–159. https://doi.org/10.1038/nrg.2017.104

Poldrack, R. A. (2015). Is "efficiency" a useful concept in cognitive neuroscience? *Developmental Cognitive Neuroscience*, *11*, 12–17. https://doi.org/10.1016/j.dcn.2014.06.001

Ponsoda, V., Martinez, K., Pineda-Pardo, J. A., Abad, F. J., Olea, J., Roman, F. J., et al. (2017). Structural brain connectivity and cognitive ability differences: A multivariate distance matrix regression analysis. *Human Brain Mapping*, *38*(2), 803–816. https://doi.org/10.1002/hbm.23419

Rietveld, C. A., Esko, T., Davies, G., Pers, T. H., Turley, P., Benyamin, B., et al. (2014). Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proceedings of the National Academy of Sciences*, *111* (38), 13790–13794. https://doi.org/10.1073/pnas.1404623111

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., et al. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, *340*(6139), 1467–1471. https://doi.org/10.1126/science.1235488

Roth, B., Becker, N., Romeyke, S., Schafer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118–137.

Ryman, S. G., Yeo, R. A., Witkiewitz, K., Vakhtin, A. A., van den Heuvel, M., de Reus, M., et al. (2016). Fronto-parietal gray matter and white matter efficiency differentially predict intelligence in males and females. *Human Brain Mapping*, *37*(11), 4006–4016. https://doi.org/10.1002/hbm.23291

Santarnecchi, E., Emmendorfer, A., & Pascual-Leone, A. (2017a). Dissecting the parieto-frontal correlates of fluid intelligence: A comprehensive ALE meta-analysis study. *Intelligence*, *63*, 9–28.

Santarnecchi, E., Emmendorfer, A., Tadayon, S., Rossi, S., Rossi, A., & Pascual-Leone, A. (2017b). Network connectivity correlates of variability in fluid intelligence performance. *Intelligence*, *65*, 35–47.

Santarnecchi, E., Galli, G., Polizzotto, N. R., Rossi, A., & Rossi, S. (2014). Efficiency of weak brain connections support general cognitive functioning. *Human Brain Mapping*, *35* (9), 4566–4582. https://doi.org/10.1002/hbm.22495

Santarnecchi, E., & Rossi, S. (2016). Advances in the neuroscience of intelligence: From brain connectivity to brain perturbation. *Spanish Journal of Psychology*, *19*. https://doi.org/10.1017/sjp.2016.89

Schafer, E. W. (1982). Neural adaptability: A biological determinant of behavioral intelligence. *International Journal of Neuroscience*, *17*(3), 183–191.

Selzam, S., Krapohl, E., von Stumm, S., O'Reilly, P. F., Rimfeld, K., Kovas, Y., et al. (2018). Predicting educational achievement from DNA. *Molecular Psychiatry*, *23*(1), 161. https://doi.org/10.1038/mp.2017.203

Shehzad, Z., Kelly, C., Reiss, P. T., Cameron Craddock, R., Emerson, J. W., McMahon, K., et al. (2014). A multivariate distance-based analytic framework for connectome-wide association studies. *Neuroimage*, *93*(Pt. 1), 74–94. https://doi.org/10.1016/j.neuroimage.2014.02.024

Sniekers, S., Stringer, S., Watanabe, K., Jansen, P. R., Coleman, J. R. I., Krapohl, E., et al. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature Genetics*, *49*(7), 1107–1112. https://doi.org/10.1038/ng.3869

Song, M., Liu, Y., Zhou, Y., Wang, K., Yu, C. S., & Jiang, T. Z. (2009). Default network and intelligence difference. *IEEE Transactions on Autonomous Mental Development*, *1*(2), 101–109. https://doi.org/10.1109/tamd.2009.2029312

Thompson, R., Crinella, F. M., & Yu, J. (1990). *Brain mechanisms in problem solving and intelligence: A survey of the rat brain*. New York: Plenum Press.

Vakhtin, A. A., Ryman, S. G., Flores, R. A., & Jung, R. E. (2014). Functional brain networks contributing to the parieto-frontal integration theory of intelligence. *Neuroimage*, *103*, 349–354. https://doi.org/10.1016/j.neuroimage.2014.09.055

Valizadeh, S. A., Liem, F., Merillat, S., Hanggi, J., & Jancke, L. (2018). Identification of individual subjects on the basis of their brain anatomical features. *Scientific Reports*, *8*(1), 5611. https://doi.org/10.1038/s41598-018-23696-6

van den Heuvel, M. P., & Sporns, O. (2011). Rich-club organization of the human connectome. *Journal of Neuroscience*, *31*(44), 15775–15786. https://doi.org/10.1523/JNEUROSCI.3539-11.2011

van den Heuvel, M. P., Stam, C. J., Kahn, R. S., & Pol, H. E. H. (2009). Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience*, *29*(23), 7619–7624. https://doi.org/10.1523/jneurosci.1443-09.2009

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., et al. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, *21*, 860–868.

Zhao, M., Kong, L., & Qu, H. (2014). A systems biology approach to identify intelligence quotient score-related genomic regions, and pathways relevant to potential therapeutic treatments. *Scientific Reports*, *4*, 4176. https://doi.org/10.1038/srep04176

PART V

Intelligence and Information Processing

20 Basic Processes of Intelligence

Ted Nettelbeck, Oliver Zwalf, and Con Stough

Introduction

The search for basic processes that support intelligence has a long history. This endeavor rests on the assumption that there are individual differences in structures of the central nervous system (CNS) whereby information critical to decision-making is conducted more or less rapidly. Reductionist theory has linked intelligent behaviors with low-level perceptual sensitivity since Galton's (1883) explorations of individual differences in sensory discriminations and reaction times. This approach was adjudged nonproductive around the beginning of the twentieth century, because studies measuring reaction time (RT) had, to that time, failed to support the theory (Jensen, 1982). At around the same time, Binet developed a practical measure of intelligence, and behaviorism and psychoanalysis successfully captured mainstream interest within psychology (Deary, 2000). Together, these circumstances established an orthodoxy that eschewed attempts to address theory about putative biological bases to intelligence for more than half a century.

Instead, the main focus for differential psychology became the further development and validation of tests of higher-order mental abilities. This approach to defining intelligence struggled initially to avoid the circularity of using description as explanation; but, arguably, modern tests do have good construct validity for culturally valued behaviors held by consensus to require intelligence (Jensen, 1998). This is important because the vast majority of researchers following a reductionist paradigm have relied on IQ-type tests to provide an (imperfect) proxy for intelligence.

Recent Interest in Speeded Tasks

From about the 1960s there has been renewed interest in mental speed as somehow fundamental to intelligence (Eysenck, 1987). Broadly, speeded tasks have been of two kinds.

Acknowledgements: Preparation of this chapter was supported by grant DP0772346 from the Australian Research Council. The authors are grateful to Nick Burns, Tess Gregory, and Carlene Wilson for their comments on a draft. This chapter was updated by Oliver Zwalf and Con Stough from an earlier published version by Ted Nettelbeck.

In the first, generally drawn from mainstream cognitive psychology and neuropsychology, tasks have been conceived as measuring individual differences in cognitive subsystems that traditionally have been incorporated within psychometric accounts of intelligence, like attention (e.g., orientation, focused, divided, sustained) or short-term, working, and long-term memory. The reductionist account assumes that individual differences in response latencies reflect underlying stages or mechanisms essential to the specified construct. Examples are S. Sternberg's (1975) four-stage short-term memory scanning model, Posner's (1978) long-term encoding function task, and R. J. Sternberg's (1977) componential analysis of analogical reasoning, which led him to invoke metacognition to direct processing resources where most required. Such tasks have successfully discriminated between people with brain damage or an intellectual disability and those without; but results have not generally located these differences within set processing stages or convincingly demonstrated that bottom-up processing, as opposed to top-down processing, was involved (Nettelbeck & Wilson, 1997). Within the normal population correlations between such measures and IQ have typically been modest (Jensen, 2006), but stronger for more cognitively demanding tasks (Schweizer, Zimmermann, & Koch, 2000). However, Deary (2000) expressed strong reservations about the utility of these more demanding tasks to reductionist theory because of uncertainty about what they measure.

The second category of speeded performance has included tasks assumed to reflect more general basic functions, such as *perceptual speed* and *information-processing speed*. As will be clear from what follows, there is uncertainty about the precise meaning of these terms. However, perceptual speed has commonly been defined as speediness on very simple tasks (Nettelbeck, 1994), whereas speed of information processing is a generic term, referring to the rate at which hypothetical basic mechanisms within the brain and CNS operate.

Theoretical accounts for why such tasks might correlate with intelligence have postulated that the brain has limited capacity to process incoming information simultaneously, so that short-term storage is lost without rehearsal. Faster processing therefore confers advantage, particularly for complex decision-making. Jensen (1982) proposed a model of this kind that derived from "neural oscillation" whereby variability in performance, rather than central tendency, is the key to understanding timed performance. Consistent individual biological differences are held to exist in the rate at which cells in a neural network oscillate between excitatory and refractory phases. A fast rate means that, irrespective of when a response is required, excitatory potential is closer to threshold, resulting in faster and less variable reactions than are generated from a slower rate of oscillation. Capacity to encode information more quickly therefore equates with a more efficient processing system at a given point in time, because more information, critical to the integration of different essential elements of a problem, is acquired from the environment and/or existing long-term storage and is retained in working memory (WM). This account implies that processing speed is central to WM, the capacity to retrieve, manipulate, and rehearse information within a very short time frame. If information quality is degraded because of limited capacity before processing has been completed, the accumulation of task-relevant knowledge will be less effective. Extending this theory to an account of intelligence, Jensen's argument becomes that the more efficient system conveys cumulative advantage for the acquisition of knowledge over time. This theory has influenced current research directions, as described below.

Reliable correlations between speeded measures and cognitive tests are now established for children. Whether faster RTs with age are confounded by higher-order responding strategies that reflect maturing problem-solving skills (Anderson, Nettelbeck, & Barlow, 1997) or reflect more functional basic cognitive development as a child grows older (Jensen, 1982) is still not known, although recent results from Edmonds and colleagues (2008) appear to favor the latter interpretation. This distinction notwithstanding, however, the evidence is overwhelming that speed on tasks with low knowledge requirements, such as RT and inspection time (IT; see the section Defining Intelligence), improves markedly from preschool years to adolescence, in parallel with increasing cognitive ability (Edmonds et al., 2008; Fry & Hale, 2000; Kail, 1991; Nettelbeck & Wilson, 1985).

Jensen (2006) has observed that the course of cognitive decline during old age that accompanies slowing processing speed appears to be "a mirror image" (p. 97) of how improving cognitive maturity and increasing processing speed develops during childhood. Whether this is literally so remains to be tested but large bodies of cross-sectional and longitudinal research, conducted over several decades, have confirmed that slowing processing speed accounts substantially, if not entirely, for age-related changes in fluid cognitive abilities (G*f*; coping with novel situations), as opposed to crystallized abilities (G*c*; using acquired knowledge to solve problems) (Finkel et al., 2007; Salthouse, 1996; Schaie, 2005). Thus, whereas tests for vocabulary and cultural knowledge show little decline throughout adult life, tests for inductive reasoning, WM, and spatial orientation on average show very marked effects, and individual differences in these abilities become more substantial with age. Moreover, when speeded performances of elderly persons on diverse tasks, supposedly requiring different processes, are plotted against the performances of young adults on the same tasks (so-called Brinley plots), the outcome is a single function (Cerella, 1985; Madden, 2001), consistent with the theory that a general speed factor is responsible for age-related cognitive differences.

Nonetheless, there are grounds for challenging whether a general speed factor provides a sufficient account for such differences. Following Danthiir and colleagues' (2005b) finding that both a general mental speed factor and independent, specific speed factors were incrementally related to differences in higher reasoning among university students, Danthiir and colleagues (2009) have confirmed a similar multifaceted speed structure with elderly participants. Age effects on speed were general, with a strong general mental speed factor accounting substantially for age-related variance in reasoning and WM. However, there were also direct effects of age, unrelated to speed, on reasoning and WM. Moreover, the best-fitting structural model for these data included additional, specific speed factors that reflected performance on tests of RT and perceptual speed. Identifying age effects unrelated to speed and better defining the nature of specific speed influences are therefore prospects for future research.

Sheppard and Vernon (2008) compiled results from 172 studies of processing speed and intelligence conducted between 1955 and 2005, involving more than 53,500 participants. Correlations between diverse measures of processing speed (choice RT, IT, perceptual speed, more complex short-term memory processing, or long-term retrieval) and various tests of intelligence remind us that, although understanding differences in mental speed may be essential to an improved understanding of intelligence, these differences do not on current evidence provide a full account of differences in intelligence. Measures of mental speed, whether tapping more or less complex decisions, correlated reliably with intelligence, whether categorized as general, fluid, or crystallized, but the n-weighted mean coefficient overall from single speed measures was only −0.24. This is typical of RT studies, and reflects the fact from Sheppard and Vernon's review that RT measures under conditions requiring low prior knowledge have vastly outnumbered other forms of speed measurement. Moreover, as will be explored further in what follows, more substantial correlations have been found with other forms of measurement.

Current widespread interest in whether a reductionist approach utilizing speeded performance can deliver a better understanding of intelligence is a major change of direction within differential psychology. Strong skepticism three decades ago as to whether more than trivial correlation between mental speed and intelligence could be established captured the Zeitgeist at that time (see Jensen, 2006, pp. 155–158). However, the volume of ongoing research currently addressing whether and, if so, to what extent basic speed processes contribute to intelligence suggests that these questions are now recognized by researchers as future priorities.

Several reviews of the field have been published (Deary, 2000; Deary & Stough, 1996, 1997; Deluca & Kalmar, 2007; Jensen, 2006; Roberts & Stankov, 1999; Sheppard & Verson, 2008). It is clear from these reviews, however, that although mental speed is widely recognized as a facet of intelligence, there is divergent opinion about the nature of this association. Brand (1996) held speed and intelligence to be isomorphic. Eysenck (1987) gave mental speed primacy as a fundamental cognitive variable that, together with aspects of personality, was responsible for individual differences in intelligence. He also speculated that accuracy of neuronal transmission might provide the biological basis of mental speed. Jensen's position has been similar but with a focus on speed as central to his definition for Spearman's *g*, that is, the first unrotated principal component extracted from performance on a battery of ability tests. Deary (2000) considered restricting intelligence to a general factor to be too narrow a description of human abilities and allowed that mental speed could prove to be more closely aligned with some specific cognitive abilities than with a general ability.

Others have pointed out that speed-IQ correlation could reflect individual differences in attentional and memory processes applied in all tasks, rather than a basic rate of processing at a biological level (Carlson, Jensen, & Widaman, 1983; Detterman, 1987; Hunt, 1980; Mackintosh, 1998; Marr & Sternberg, 1987). Alternatively, as demonstrated by substantial practice effects on elementary cognitive tasks (ECTs, that is, tasks with low knowledge requirements; see the section

Speed of Information Processing and Elementary Cognitive Tasks), it is possible that higher IQ determines capacity to render response organization more automatic (Rockstroh & Schweizer, 2004).

Studies of cognitive development have pointed to a close association between improving processing speed and WM. Thus, Fry and Hale (2000) described this relationship as part of a "cognitive developmental cascade" whereby cognitive maturation depends on improving processing speed, which results in improved WM, which in turn influences fluid reasoning. Salthouse (1996) has expressed the same idea, but in reverse, to account for cognitive aging. There are grounds, however, to question whether the simple cascade model provides a sufficient account for cognitive performance, either in older or younger adults. Thus, Gregory and colleagues (2009a) reported a direct path between age and WM for elderly participants that excluded speed differences, and Conway and colleagues (2002) found strong support for a model wherein WM in young adults strongly predicted fluid reasoning, whereas processing speed did not. Following Engle and colleagues (1999), they suggested that strong correlation between WM and general ability may reflect executive attentional processes.

Recent research, particularly within Germany, has explored relationships between attention, WM, speed and intelligence. Buehner and colleagues (2006) provide a good example of this approach, set within debate about whether WM and intelligence are essentially isomorphic (Kyllonen & Christal, 1990) or substantially independent (Ackerman, Beier, & Boyle, 2005). Buehner and colleagues used an extensive test battery, requiring up to nine hours of testing for timing and accuracy on tests of WM, sustained attention, intelligence and two-choice RT to diverse verbal, numerical, and spatial stimuli. They found that aspects of WM responsible for brief retention of new information and for coordination/integration of operations, rather than a general speed factor, were central to reasoning, but that WM and reasoning were nonetheless distinguishable. Sustained attention was equivalent to coordination. Speed of WM operations, particularly for selective attention, conferred performance advantage but this was independent from the influence of a general factor derived from all tests of WM, G*f*, and G*c*. By this account, therefore, speed is essential, but does not provide a sufficient explanation, for intelligence. This conclusion has received strong support from a recent study by Kaufman and colleagues (2009). They found that general associative learning, WM, and a composite speed variable (derived from verbal, numerical, and figural speed tests) all had incremental validity for a general intelligence factor defined by verbal and perceptual reasoning, and mental rotation abilities.

Speed of processing may indicate higher intelligence via greater control of *vigilant attention* (VA), the ability to focus mental resources into relevant task schema over time. A meta-analysis investigating VA in fifty-five functional neuroimaging studies by Langner and Eickhoff (2013) revealed substantial overlap with the right-lateralized ventral attention network during tasks requiring VA, which has been consistently related to bottom-up reorienting of attentional processes. This finding was in opposition to most traditional top-down, goal-driven theories of VA, leading the authors to postulate that sustaining VA may require maintaining a state in which target stimuli can be optimally reoriented toward themselves. They noted the presence of time-related VA decrements, where time spent on task was highly correlated to worse performance, thus highlighting the utility of mental speed measurements.

We have already raised the possibility that mental speed is multifaceted. In fact, within psychometric theory, Horn consistently raised doubts about speed as a unitary process, distinguishing between broad group factors for speediness (G*s*; quick responding on very simple tasks) and correct decision speed (CDS; responding speed on cognitively demanding tasks) while acknowledging that CDS has been less reliably identified (Horn & Noll, 1997). Further, Danthiir, Wilhelm, and Schacht (2005a) found distinguishable but correlated CDS factors that related to G*f* and G*c*, respectively, but that resulted from confounding between speed, ability levels, and item difficulty. Similar to R. J. Sternberg (1977), Danthiir and colleagues found that, although more intelligent participants were generally quicker overall, they took longer than less intelligent participants on the most difficult items. As Danthiir and colleagues pointed out, irrespective of whether these differences reflected task characteristics like higher complexity of difficult items, or personal characteristics like more persistence among smarter participants, they did not support a simple explanation for higher reasoning in terms of faster basic processing.

Carroll's (1993) taxonomy for intelligence included G*s* as a second-stratum factor, which he distinguished from processing speed (G*t*) and psychomotor speed (G*p*) components from ECTs. It has not always been clear, however, that such theoretical distinctions have been justified by empirical evidence. Confusion about what constructs different tests represent has sometimes been the consequence of different assessment traditions, for example, neuropsychological versus psychometric. For example, Krumm and colleagues (2008) found that a latent variable, sustained attention (a neuropsychological construct), was virtually indistinguishable from psychometric G*s*, which, however, closely resembled Carroll's psychometric G*t*.

Roberts and Stankov (1999) provided detailed consideration of methodological issues that research should confront and reported a large-scale investigation of speed in relation to a hierarchical, multivariate model of intelligence. Their battery included multiple ECTs and psychometric tests representing seven of the nine broad group factors that define Horn's G*f*-G*c* theory. Roberts and Stankov concluded that mental speed is complex and described by a hierarchical model with a broad cognitive speed factor extracted from five separable less broad speed factors and located on the same level as their seven broad cognitive abilities.

Jensen's (2006) comprehensive overview of the history of "mental chronometry" is a substantial account of research in this field and it is clear that he remains convinced that an emerging "science of chronometry" can further understanding of intelligence. Both Roberts and Stankov (1999) and Jensen (2006) emphasized that, before attempting to answer how mental speed relates to intelligence, there are two fundamental theoretical issues to be addressed by future research. The first is how best to describe intelligence; the second is to determine whether speed is better represented as unitary or multifaceted.

Defining Intelligence

Many researchers correlating individual differences in ECTs with differences in intelligence have assumed that a single test like Raven's Progressive Matrices provided a sufficient account of intelligence, a practice criticized by Juhel (1991) as inadequate. However, although we have witnessed growing acceptance during the past two decades that relying on a single test as a marker for intelligence is not adequate, the definition of intelligence accepted for much of the research with ECTs still lacks consensual definition. We have seen that Jensen (1998) has argued that "intelligence" is so vague as to be scientifically useless, proposing instead that the core aspect of mental ability be represented by Spearman's *g*. Others have disagreed, arguing that, although a general factor represents commonality among whichever tests comprise the test battery, this will reflect different content across batteries, so that other aspects of intelligence, defined by hierarchical psychometric models, should be taken into account (Horn & Noll, 1997; Roberts & Stankov, 1999).

This debate also reflects uncertainty about the causal function of a general factor. Although evidence for psychometric *g* is strong (Jensen, 1998), it does not necessarily follow that there exists a single property that is invested in all mental activities. For example, as Detterman (1982) pointed out, individual differences in *g* could be the consequence of relative efficiencies within a system composed of independent functions, like executive control of attention, a perceptual register, WM, long-term memory, and a response mechanism. Although defined as separate components, Detterman conceptualized these functions as interrelated within the system because all would be necessary for the system to operate; and on this view all would be involved to varying degrees in all mental activities.

There is now at least considerable agreement among researchers in the field that the psychometric intelligence for which reductionist accounts are sought is multifaceted. Therefore, explanations for individual differences in intelligence require taking account of some nine or ten broad, relatively independent factors that nonetheless share variance that defines a substantial general factor. These broad factors are derived from a larger number of more narrowly defined ability factors, with these in turn defined by performance on a potentially limitless number of tests. Because most hierarchical models require a strong general factor to provide a comprehensive psychometric description of test variance, they accommodate both sides of the long debate about whether intelligence is better described as a single entity or as multiple abilities.

Several different versions of a hierarchical structure have been proposed but the taxonomy currently attracting widest acceptance derives from the three-stratum account of cognitive abilities advanced by Carroll (1993). Following adoption of Carroll's taxonomy as compatible with Horn's and appropriate to underpin the development of the Woodcock-Johnson III Tests of Cognitive Abilities (McGrew, 2005), it has become widely referred to as CHC theory (i.e., Cattell-Horn-Carroll). This account of intelligence has explanatory value insofar as test scores can be shown to predict important life outcomes. However, this conception, although multifaceted, does not extend to include suggestions about the importance of practical intelligence or creativity (R. J. Sternberg, 2003), or musical or bodily-kinesthetic abilities (Gardner, 1983), or emotional intelligence (Mayer & Salovey, 1993).

Speed of Information Processing and Elementary Cognitive Tasks

Different terms have described quick responding – processing speed, cognitive speed, psychometric speed, perceptual speed, and so on. As clarified above, speed of information processing is a generic term referring to putative basic processes whereby external events are registered and manipulated, so as to give rise to observable behaviors. Methodology derived from speeded tasks assumes that cognitive processes intervening between stimulus and response can at least be relatively isolated by appropriate manipulation of experimental conditions.

The term "elementary cognitive task" was first coined by John B. Carroll around 1980 to describe tests of timed performance assumed to require few cognitive processes, that could be completed satisfactorily by anyone in the absence of time constraints (see Carroll, 1993, pp. 11–13). Current acceptance that, after all, speed of information processing may be an important aspect of intelligence dates from the last four decades of the twentieth century. This research has focused on correlations between ECTs and scores on a diverse range of IQ tests, foremost among these being the Wechsler scales and matrices tests like Raven's and Cattell's.

Deary (2000) criticized use of the term "ECT" – and others like speed of information processing, perceptual speed, and mental speed – as lacking explanatory value because they have remained poorly defined. Various speed terms have been used interchangeably, implying that all mean the same thing, although this has not been established. Arguably, however, although such terms reflect limited current understanding, they do capture aspects of mental activities that are intrinsic to human nature. Moreover, they are what we currently have to work with, and theoretical formulation of some kind is a necessary first step to scientific progress. Thus, it does not follow that because a construct is poorly understood, future improvement in understanding is impossible.

It is also apparent that the complexity of content of different speeded tasks varies. Thus, Jensen (1998) has maintained that information-processing speed is different from G*s*, typically measured from pencil-and-paper psychometric tests. Jensen (2006) has raised the possibility that speed from more simple RT tasks might be distinguishable from speed on tasks developed to tap more complex cognitive processes. Detterman (1987) earlier outlined a possible way forward on issues of this kind, using factor analysis to clarify the definition of commonalities among and specificities within multiple speeded tasks and then testing these structures against multifaceted models for intelligence. Although some researchers have followed this path (Burns & Nettelbeck, 2003; Danthiir and colleagues, 2005b; Neubauer & Bucik, 1996; Roberts & Stankov, 1999), the matter is certainly not yet resolved

(Jensen, 2006). As foreshadowed above, debate continues about whether there are individual differences in different kinds of processing speed that are specific to different capacities or in a single, basic speed construct (Anderson, 1992), although most recent evidence suggests that speed is multifaceted (Danthiir et al., 2009).

Widespread use of the term "ECT" today is principally the consequence of its adoption by Jensen (e.g., 1998), the most prolific researcher on this topic of a relationship between speeded performance and intelligence. Attempts to better understand the nature of intelligence by the study of ECTs rest on the reductionist assumption that such tasks, although not strictly biological, predominantly isolate low-level processes that operate to generate and manipulate knowledge within storage and retrieval structures. This theory holds that individual differences in measures of intelligence, and therefore in real-life achievements, are to some extent the consequence of differences in ECT performance. Broadly, two different approaches to measuring processing speed have been used: reaction time (RT), whereby the time of making a detection or discrimination is measured by the duration between a presented stimulus and the registration of a reaction; and inspection time (IT), whereby the time to make a decision is inferred from accuracy of judgments under time constraints but without requiring quick reactions.

Jensen's Studies of Reaction Times

The most comprehensive body of data assembled to test the theory that processes responsible for speed on ECTs are the same as those responsible for complex intelligent actions comes from Jensen's studies of simple and choice RTs, made principally from the late 1970s through the 1980s. Jensen (1982, 1987, 1998, 2006) has provided extensive accounts of this research involving more than 2,000 participants, which has been reviewed by several authors (Carroll, 1987; Deary, 2000, 2003; Longstreth, 1984; Mackintosh, 1998; Nettelbeck, 1998; Neubauer, 1997). Although reviewers have not reached consensus about how Jensen's results should be interpreted, there is now general agreement that stronger correlations can be found between RT and intelligence tests than was previously thought.

Jensen adopted an apparatus designed to decouple a decision time (DT) from movement time (MT) in a two-stage responding process (see Figure 20.1; and see Jensen, 2006, pp. 27–29 for detailed description). Jensen's main objective was to test the hypothesis that individual differences in the slope of the linear regression of latency on the number of target alternatives (expressed as binary logarithmic transformations) are the principal source of correlations between RT and intelligence (Hick, 1952). Specifically, if DT taps processing speed then flatter slopes should reflect higher intelligence, whereas MT should be constant across degrees of choice and therefore not correlate with intelligence. This hypothesis has been tested, predominantly using scores on Raven's matrices as an index for general intelligence, by comparing groups with different average abilities and by within-group correlation between various parameters of distributions of DT and MT with intelligence scores.

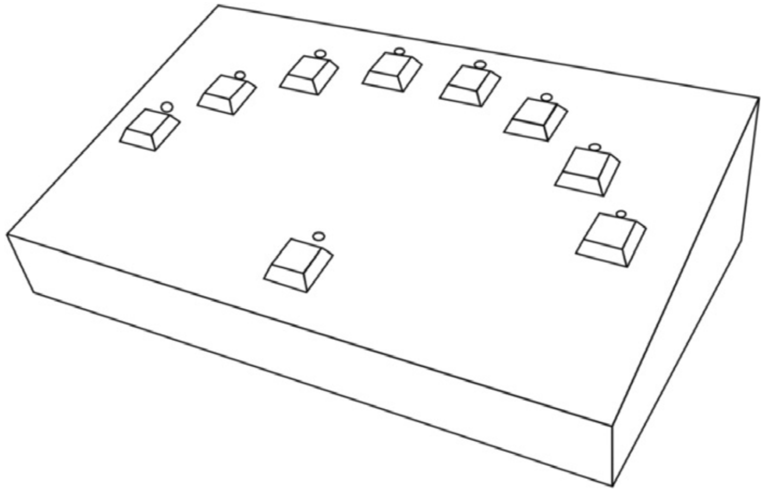

Figure 20.1 *Reaction-time apparatus (after Jensen, 1987).The eight alternative stimulus lights are equidistant from the home button. When a stimulus light is illuminated, two timers register (1) time to lift-off from the home button (decision time); and (2) time from lift-off to turning off the target (movement time).*

In some instances substantial correlations have been demonstrated between latency and intelligence measures, but results have generally not supported the hypothesis.

Although group data have generally conformed closely to Hick's theory, individual data have fitted less well. Moreover, following Longstreth (1984), several critics have challenged Jensen's interpretation, which attributes a causal function to processing speed. Subsequent consideration has probably successfully discounted alternative explanations for the observed correlations in terms of cognitive strategies reflecting sundry methodological variables (configuration of potential targets, order of presentation for choice alternatives, putative visual attentional biases linked to set size, different set sizes requiring different physical responses, opportunities for speed-accuracy trade-off). Nor were these correlations the consequence of speed constraints on intelligence items (Vernon, 1987). It is possible, nonetheless, that Jensen's procedure provided insufficient practice to discount a possibility that higher-IQ participants adapted to task requirements more efficiently (Nettelbeck, 1985).

Most critically, however, correlations involving individual regression slopes (proposed by Hick, 1952, to capture information-processing speed) were not reliably stronger than those involving other parameters of RT, like regression intercept for DT, mean or median DT, or even MT. Using multiple regression, Jensen demonstrated that different combinations of latency variables can account for as much as about 50 percent of variance in intelligence scores. However, such analyses did not identify an optimal set of parameters that might advance explanation for the correlation. Particularly troublesome have been significant correlations involving MT because the theory provides no basis for these. A likely explanation is that these have reflected confounding between DT and MT as a consequence of occasional early detection responses (i.e., an as-yet-unlocated illuminated target is detected but before the discrimination judgment has formed) (Smith & Carew, 1987).

Nonetheless, although Deary (2000, p. 181) concluded that attempts to decompose RT into underlying cognitive constructs were not convincing, he accepted that accumulated evidence established that correlations between RTs and psychometric ability were sufficiently substantial to warrant continuing interest. His subsequent comments on this matter (Deary, 2003) noted that the more complex response actions required with Jensen's apparatus may have introduced unexpected top-down strategic processes and that future work should therefore rely on the traditional apparatus (individual fingers for alternative responses). To an extent, Stough and colleagues (1995) addressed some of these strategy use/methodological issues. However, whether adopting the earlier techniques will improve prospects for advancing knowledge is still unclear. Deary's point was well made. Arguably, however, all ECTs are to some extent confounded by idiosyncratic cognitive strategies that cannot be excluded (Nettelbeck, 1998) and, although this need not be a critical obstacle to progress if acceptably robust construct validity for such tasks can be established, it may be that different kinds of apparatus will prove to be better suited to different circumstances. For example, removing or reducing motor influences from responding requirements could be more of an issue for elderly than for younger respondents.

Deary (2003) made two further points for future consideration. First, relying on untransformed data from simple and choice RT conditions, rather than continuing with parameters extracted from the Hick function, should be more tractable for theory building. Second, despite a very large body of research published in this area, the effect size of the RT-intelligence correlation had not yet been determined. Deary, Der, and Ford (2001a) addressed the second question for a large representative sample of Scottish men and women in their fifties who were participants in a large ongoing population-based study, begun in 1988. Scores on a widely used British test of general mental ability (Alice Heim Part 1; AH4) correlated with simple and four-choice RT. Corrected for test unreliability, "true" effect size was about −0.5, independent from sex, social class, and education, confirming Deary's conviction that there is a substantial relationship to be explained.

In a follow-up after thirteen years, Deary, Allehand, and Der (2009) applied cross-lagged correlational analyses to test the hypothesis that faster processing speed is responsible for more successful cognitive aging. The rationale for this design rests on the assumption that correlation between antecedent and subsequent variables establishes consequence, from former to latter. Structural equation modelling defined latent factors for processing speed from simple and four-choice RT at both baseline and time 2; and latent factors for intelligence from the AH4 tests. Correlations between latent speed and ability factors were as expected from the 2001 study (−0.49 and −0.41 for times 1 and 2 respectively). However, contrary to prediction, only the path from the first latent ability factor to the later processing speed factor was statistically significant (−0.21), leading the authors to suggest that "higher general intelligence might be associated with lifestyle and other factors that preserve processing speed" (p. 40). This may be so; but, as outlined in the section Variability of Individual Reaction Times, it does not exclude the possibility that antecedent measures of processing speed can predict subsequent cognitive integrity. It is possible too that Deary and colleagues' (2009) result owed something to the relatively low test-retest reliability of their speed construct (0.49 cf. 0.89 for the ability factor). Indeed, insofar as their RT apparatus confounded cognitive and motor responding (a problem that Jensen's apparatus tends to reduce), this outcome may have reflected deteriorating motor dexterity in sixty-nine-year-olds.

Across several narrow bands of age samples, Demetriou and colleagues (2013b) observed several cycles of differentiation and integration between WM, WM reaction times, and G*f*. Fluid intelligence evolves over several iterations of reconceptualization, where changes in the nature of mental representations alternate with changes in the command and interlinking of previously constructed representations. Using approximated age boundaries, they postulated that four reconceptualization cycles occur at ages two, six, and eleven, while within-cycle transitions occur at ages four, eight, and fourteen, finding G*f* changes were predicted by speed during the initial phase of a cycle (i.e., ages 6–8 and 11–13) and WM during the second phase (i.e., ages 4–6, 8–10, and 13–16 years). Interestingly, the stages outlined by Demetriou and colleagues correspond to Piaget's theory of cognitive development, insofar as development of logical reasoning may not be the driver of representational reorganization, but instead an observable symptom, as Demetriou and colleagues noted. Verhaegen's (2013) meta-analysis of age-related differences in processing speed supported this moderately strong association, finding a correlation of 0.53 between reasoning and mental speed.

Variability of Individual Reaction Times

Recent theoretical interest about how RT relates to intelligence has tended to shift from measures of central tendency in RT to variability in trial-to-trial performance. This has followed observations (Baumeister & Kellas, 1968; Brewer & Smith, 1984; Jensen, 1987) that, even between groups with widely disparate abilities, fastest RTs differ little and differences are captured by the extent to which individual distributions are positively skewed. When a respondent's RTs within a set condition are ranked from fastest to slowest, within-rank correlation with intelligence increases from the fastest to the slowest RTs. This finding has resulted in a focus on *worst performance* (WP; Larson & Alderton, 1990). Variability of responding also increases systematically as RTs slow, implying that it is increasing unreliability of responding that is responsible for higher correlation between intelligence and worst-performance RTs. The relationship appears to apply for cognitive abilities that have higher *g* loading but not for tasks that do not. Moreover, mean levels of WP reliably differentiate between groups with different mean IQs when RTs in these groups are measured by the same procedures, principally because more marked skewing of RT distributions is related to lower intelligence.

Coyle (2003) reviewed relevant research, including consideration of possible causes for these relationships. He acknowledged that WP could reflect psychological variables like lapses in attention or WM but argued that these can represent functioning at a fundamental biological level rather than top-down cognitive processes influenced by conceptual knowledge. He favored Jensen's theory of individual differences in rate of neural oscillations, and outlined an agenda for future WP research.

Longitudinal studies of children/adolescents by Demetriou and colleagues (Demetriou et al., 2013b; Demetriou, Spanoudis, & Shayer, 2013a) have supported Coyle's and Jensen's postulations, finding both variability of speeded performance and G*s* to be distinct factors that additively contribute to WM, and fluid abilities as indexes of processing efficiency. Variability shared a greater association with WM, which may reflect superior regulation of executive control and attentional resources across numerous speeded tasks. Further, the observed performance differences between ECTs and complex tasks shared a significant relationship with variability, potentially explaining the significant relationship shared with G*f* development.

Schmiedek and colleagues (2007) have drawn from three previously largely separate strands of research to test whether efficiency of RT performance relates to intelligence. First, they pointed out that reliability of WP analyses derived from separate RT bands is limited by small numbers of trials within bands. However, the ex-Gaussian distribution (a normal-like distribution obtained by convolving a Gaussian with an exponential distribution) provides an appropriate description for RT distributions. Specifically, in addition to mean and SD, the distribution parameter tau (τ) integrates information from all trials but predominantly reflects skewness, particularly at the extreme tail. Tau is therefore sensitive to the slowest RTs; and Schmiedek and colleagues noted evidence that linked τ with fluctuation in attention. Second, Schmiedek and colleagues considered evidence that WM and reasoning (core abilities to *g*) reflect attentional control, both over distraction and for maintaining focus. This theory therefore predicts that slower RT is the consequence of poorer executive attention, which impacts WM, which in turn impacts on reasoning ability. Third, however, Schmiedek and colleagues sought an alternative to attention as a causal explanation, drawing on the diffusion model of choice RT (Ratcliff & Smith, 2004). This is a *random walk* model for two-choice decision-making that assumes that information on which a decision is reached is accumulated sequentially over time. The two most critical parameters of this model for current discussion are the *response criterion* (i.e., level of information required before responding) and *drift rate* (mean rate of decision-making). Because drift rate is essentially an index for the quality of information processed, it should be the most sensitive to slower RTs and therefore most related to τ.

Latent trait analyses of multiple tasks for WM, reasoning, and RTs for verbal classification, quantitative decision, and spatial orientation tasks confirmed commonalities within parameters across different tasks. RT (mean, SD, τ) accounted for more than 50 percent of variance in WM and reasoning factors; but τ showed stronger correlations with the cognitive traits (around −0.7). Similarly, compared with parameters for response criterion and nondecision components of RT, drift rate extracted from a scaled-down diffusion model was by far the strongest predictor of WM (0.68) and reasoning (0.79).

These results were consistent with theory that lower intelligence reflects poorer executive control but, as Schmiedek and colleagues (2007) argued, could also mean that differences in τ, representing efficiency of information processing, can provide a more parsimonious account. They tested this idea by simulating model and distribution parameters, demonstrating that the strong correlation between τ and the WM factor was wholly accounted for by drift rate. A second simulation introduced trial-to-trial variability into drift rate, to represent occasional lapses of attention that could interrupt information accumulation. This simulation produced lower τ-WM correlation than was determined empirically, so it was improbable that the observed correlation was due to attentional fluctuations, although not excluding this possibility.

Schmiedek and colleagues' (2007) account therefore avoided introducing an attentional construct in addition to drift rate. To account for what is responsible for the efficiency construct, they proposed their theory that the function of WM is to make and maintain temporary "bindings" between stimulus and response representations. (Binding is the mechanism whereby separate elements of knowledge are accessed within memory and coordinated and synthesized as required, to produce new knowledge). Oberauer and colleagues (2008) called this concept *relational integration* and proposed its centrality to G*f*. However, debate has focused on the extent to which WM influences fluid abilities, on account of the large variation between reported associations of the two constructs. For example, WM was reported by Oberauer and colleagues (2005) to explain shared variance of 75 percent while conversely, Unsworth (2010) reported only 24 percent shared variance with fluid intelligence. Chuderski (2013) delineated these discrepancies with a meta-analysis of twenty-six studies, finding that methodological designs with a "highly speeded task" condition (i.e., participants were given a time limit to complete the Ravens' Progressive Matrices) had indistinguishable (i.e., 1.0) WM-G*f* correlations compared to less correlated "moderately speeded task" or "no speeded task" studies. Similarly, a follow-up study by Chuderski (2015) using an original sample found that under time pressure, WM accounted for 83 percent of G*f* variance, while untimed G*f* tests only shared 58 percent of WM variance. Crucially, the *relational integration* factor (aka binding) had the highest loading on WM. This theory therefore holds that efficiency of the binding mechanism located in WM, which relies on consistency in speeded performance, is central to intelligence. This work represents an advance and sets a promising future research agenda that focuses on the relevance of individual differences in response variability to improved understanding of differences in intelligence.

Inspection Time

Inspection time (IT) was conceived by Douglas Vickers around 1970 as a fundamental limitation on the rate at which external information critical to making a decision can be accumulated in temporary sensory stores. Vickers' theory was heavily influenced by earlier ideas about a "perceptual moment" (Stroud, 1956) and limitations to processing efficiency dictated by "single channel operation" (Welford, 1968) (see also Lehrl & Fischer, 1990, for their account of the history of such ideas within the German information-processing tradition).

Vickers proposed an optional-stopping, random walk model of decision-making whereby information is initially briefly stored at an early stage of visual processing by a series of discrete sequential samples ("inspections") from proximal stimulation, made against a background of "noise," both internal and external, in accordance with an internally held standard for what constitutes sufficient evidence to permit a decision. The duration of an inspection, which determined the rate at which information is accumulated, was held to be independent from the criterion for sufficient evidence. The measurement of IT was operationalized as the minimum time to accumulate sufficient information to make a decision with high reliability about which of two highly discriminable lines of different lengths was longer (or shorter).

Several challenges to the construct validity for this account of IT have been acknowledged (Deary, 2000; Nettelbeck, 2001). Here, "IT" is used to refer to the measure, not a putative sampling mechanism. Figure 20.2 illustrates a current version of this task. Alternative targets are briefly displayed, with duration varying in accordance with the viewer's accuracy. Consistent accuracy results in shortened target duration but an error results in lengthened duration. Exposure duration is set by presentation of a second figure, termed a backward pattern mask, which disrupts perception of the target. Phenomenologically the target disappears, becoming integrated with the contours of the masking figure. Based on theory advanced by Turvey (1973), Nettelbeck and Wilson (1985) demonstrated by experiment that this masking effect was located centrally, beyond the peripheral visual system.

The viewer indicates whether the shorter (or longer) line is located to left or right, but speed of this response is not relevant to the determination of IT. Instead,

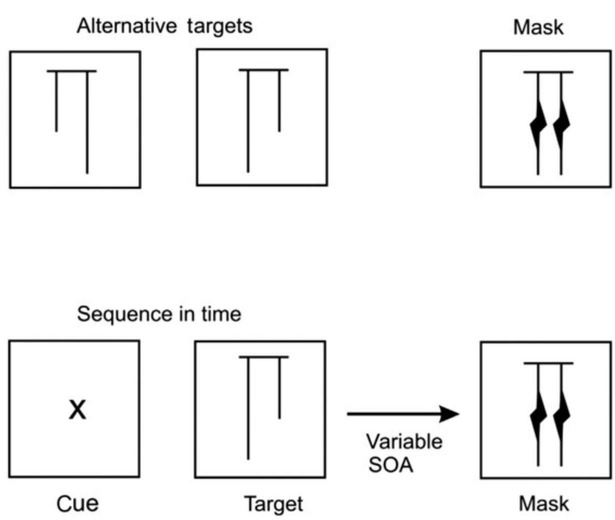

Figure 20.2 *A procedure for measuring visual inspection time.*

processing speed is inferred from accuracy of performance under conditions that limit exposure of the target to the duration between the target onset and the onset of the mask that follows (*stimulus-onset-asynchrony*; SOA). IT has been measured by different methods, with different criteria for accuracy, and using different targets and a variety of masking procedures.

There have been attempts to measure IT in other sensory modalities, on grounds that similar results across modalities would strengthen the conclusion that IT tapped central, not peripheral, processes. The first such task, developed by Brand and Deary (1982), required auditory discrimination between two tones presented for varying lengths of time as either high-low or low-high sequences. Just like in the visual IT paradigm, the critical variable was the shortest tone duration at which a listener achieved a specified high accuracy. Subsequently, other researchers devised different versions of this task that manipulated the pitch difference between the tones or that used different forms of auditory masking. However, problems in achieving effective masking, together with the realization that up to as many as 35–50 percent of participants encountered difficulty in completing the task, led Olsson and colleagues (1998) to develop a task in which loud-soft or soft-loud alternatives replaced pitch discrimination (see Deary, 2000, chapter 7 for a detailed account of this work). Parker, Crawford and Stephen (1999) developed an auditory discrimination task that requires locating a target tone in space, with tone duration at which high accuracy is achieved as the critical variable. Zajac and Burns (2007) have recently compared performance of children aged 10–12 years on both visual IT and auditory IT requiring spatial location. They concluded that both versions, together with a coding task (G*s*), shared sufficient variance to implicate common central processes. However, correlations between the three tasks were markedly stronger for children with slower ITs, implying that children with faster and slower ITs may be using different strategies. Only one study (Nettelbeck & Kirby, 1983) has sought to measure IT in the touch modality; this encountered a problem with diminishing tactile sensitivity as a consequence of direct stimulation. To summarize, only limited attempts have been made to measure IT in different sensory modalities, with most research limited to visual IT.

Correlation between IT and IQ

The first actualization of the now widely applied visual version of IT (Vickers, Nettelbeck, & Willson, 1972) was observed by Nettelbeck and Lally (1976) to correlate with IQ. The considerable body of research generated by this initial finding has been previously reviewed on a number of occasions and the interested reader is referred to Brand and Deary (1982), Deary (2000, chapter 7), Deary and Stough (1996), and Nettelbeck (1987, 2001, 2003).

Nettelbeck and Lally's assumption that IT represented early perceptual efficiency, and might therefore reveal some basic aspect of intelligence, was soon challenged by suggestions that those with higher IQ performed more effectively than those with lower IQ on simple and complex tasks alike, because they were capable of generating better learning strategies, including being prepared to try harder (Mackintosh, 1986). In addition to learning strategies there were several methodological criticisms of the IT methodology, which employed a backward mask to limit stimulus duration and therefore discrimination time. Differences in the effect of these methodologies were empirically evaluated by Stough and colleagues (2001a), who concluded that appropriate designed backward masks do not allow participants to bypass the actual processing of sensory information. Deary (2000, chapter 7) has provided a detailed review of research that has attempted to resolve these matters, concluding that there was no evidence to suppose that the relationship was principally the consequence of better learning strategies or motivation or the effects of personality. This conclusion is challenged, however, by evidence that extended practice tends to reduce range of individual differences in IT (Nettelbeck & Vita, 1992) and that even limited task experience can produce larger improvement in children's ITs than maturation (Anderson, Reid, & Nelson, 2001). Currently it remains plausible that IT taps some low-level aspect of perceptual learning (Burns et al., 2007).

Nonetheless, twenty-five years of research into the relationship between IT and IQ (Grudnik & Kranzler, 2001) has established that a moderately strong correlation exists. Grudnik and Kranzler's meta-analyses were based on more than 4,000 participants in ninety-two studies, sixty-two involving adults and thirty involving children. Ten studies involved auditory IT, but the mean correlations with IQ from auditory and visual tasks were virtually identical. Across all studies the uncorrected mean correlation was −0.3. Corrected for sampling error, attenuation, and range variation, this correlation was −0.51. The mean corrected correlation among children was slightly lower (−0.44) but still substantial. Corrected correlation for self-identified strategy users (those who acknowledged associating apparent movement cues with the shorter line when the backward mask appeared) was statistically significantly lower than that for nonusers (−0.60 and −0.77 respectively), although still substantial. Clearly, this result was consistent with Egan's (1994) conclusion that IT-IQ correlation is not explained simply by assuming that smarter people have access to smarter strategies for both easy and more challenging tasks. Reliability of IT, estimated for both test-retest and internal consistency, was good, averaging 0.8.

Inspection Time as a Lead Marker for Unfavorable Aging

Although noticeable decline in WM and fluid abilities accompanies normal aging, particularly beyond the sixth decade, chronological age (CA) is a poor predictor for individual functioning because different functions change at different rates, highly practiced skills may be relatively protected and, despite average trends, there are marked individual differences in onset and progress of age-related changes accepted as normal. Moreover, some individuals experience more severe decline, which may reflect the impact of age-related dementia-type diseases, the prevalence of which increases with old age. A major challenge is therefore to develop quantitative lead markers that can detect early preclinical signs of deterioration, before this becomes established. It is assumed that, if this could be successfully done, further decline might be slowed by appropriate intervention. Although there is currently debate about the effectiveness of available interventions (Salthouse, 2006), a considerable body of recent research has provided grounds for optimism (Hertzog et al., 2008). Related to this prospect, recent research has suggested that slower and/or slowing IT may provide a *biomarker* for less favorable aging (Gregory et al., 2009; Gregory et al., 2008; Gregory, Nettelbeck, & Wilson, 2009b).

Birren and Fisher (1992) have set out requirements for a quantitative biomarker; IT meets several of these. It is noninvasive, convenient, and reliable, with low knowledge requirements, it isolates cognitive performance from motor competence, and it monitors a process that reflects normal age-related cognitive decline, slowing steadily and appreciably across adulthood (Nettelbeck et al., 2008). It is also sensitive to abnormal cognitive decline in people with mild cognitive impairment (Bonney et al., 2006) and Alzheimer's disease (Deary et al., 1991).

Most importantly, Gregory and colleagues (2008) have shown that ITs from elderly persons aged 70–91 predict performances eighteen months later on fluid reasoning and WM, and decline in WM over this time. Moreover, slowing IT from baseline across both six and eighteen months correlated with fluid reasoning eighteen months later. These results were not found with concurrent physiological measures for grip strength, systolic blood pressure, and visual acuity. Follow-up forty-two months from baseline (Gregory et al., 2009b) showed that IT trajectories across this time were markedly different depending on whether participants at forty-two months showed incipient cognitive decline not apparent at baseline. For those with only marginally poorer recall and recognition memory, ITs had slowed appreciably at a constant rate, whereas ITs were unchanged for those without signs of memory decline. Gregory and colleagues (2009a) examined the potential relevance of slower IT for future practical, everyday functioning by comparing two samples of elderly persons matched at baseline for age, gender, education, and visual acuity, but with initial nonoverlapping distributions for faster and slower ITs. At baseline the two samples did not differ for self-reported functioning on activities of daily living like housekeeping, gardening, shopping, and moving around their communities. However, direct observations of performances forty-two months later on everyday tasks (understanding medication instructions, telephone use, managing finances, understanding instructions for food preparation) clearly confirmed that those persons with initially slower ITs now made more errors and were slower on the tasks of everyday functioning. To summarize, slowing IT in old age predicted subsequent decline in cognitive and everyday functioning, well before these changes were detectable. This result strongly suggests that IT is sensitive to changes in basic processes. What those processes are has not been determined, but the tasks of everyday functioning all relied substantially on WM. Taken together with Gregory and colleagues' (2008) finding that IT predicted WM functioning at eighteen months and decline over this time, these results raise the possibility that the IT task measures speed of some basic aspect of WM.

The Nature of Inspection Time

Crawford and colleagues (1998) were the first to attempt to locate IT within a psychometric model for intelligence. They found that IT loaded only weakly on an orthogonal general factor defined by all WAIS-R subtests, but moderately on a broad perceptual-organization factor defined by the Performance subtests. There was no relationship between IT and the group factor attention-concentration, although some research has implicated attention as responsible for IT differences (Bors et al., 1999; Fox, Roring, & Mitchum, 2009; Nettelbeck & Young, 1989). Results similar to Crawford and colleagues' (1998) were reported for children by Petrill and colleagues (2001), using WISC-R to define orthogonal broad factors for verbal, performance, and freedom-from-distractibility abilities, together with a strong psychometric general factor (*g*). Confirmatory factor analysis found that several ECTs combined to define a latent speed trait that shared substantial variance with *g*. IT shared variance with the speed factor, but predominantly contributed to performance and *g* via substantial residual paths. Thus, IT predicted *g* by two pathways; one shared variance with other ECTs but the other reflected different sources of variance unique to IT. These results are consistent with speculation that IT is psychologically complex (Nettelbeck, 2001), and also with a suggestion by Gregory and colleagues (2008) that IT is linked with WM, at least in elderly persons.

Mackintosh and Bennett (2002) tested relationships between IT and markers for G*c*, G*f*, and G*s*, concluding that IT correlated with G*s*. Similarly, Burns and Nettelbeck (2003) used a test battery selected to return broad factors from G*f*-G*c* theory of G*f*, G*c*, G*s*, G*v* (visual processing), and G*sm* (short-term memory), and included two different methods for estimating IT, as well as a backward masking task involving alphanumeric stimuli and up to four degrees of choice. All of these tasks loaded strongly on G*s*, which in turn loaded strongly on a general factor, although the strength of this association doubtless reflected speed constraints on many of the tests in this battery. Subsequent unpublished analyses have established strong commonality among these three tasks, thereby defining a latent IT variable with high loading on the general factor.

Burns and Nettelbeck (2003) also included "odd-man-out" RT (Frearson & Eysenck, 1986). For each trial, three stimulus lights on the panel of the apparatus in Figure 20.1 are illuminated so that two are adjacent and one is further away. The required response is a fast reaction to the latter. Unlike IT, performance on this task loaded strongly on G*f*, suggesting that the two tasks measure different processes. However, O'Connor and Burns (2003) obtained results that questioned this conclusion. O'Connor and Burns(2003) used exploratory and confirmatory factor analysis to locate IT within a hierarchical model for different speed factors derived from traditional perceptual-speed tasks, choice and odd-man-out RT (decoupled into DT and MT), and cognitively more demanding tasks involving evaluation and manipulation of digit and letter displays. IT correlated with the group factors visualization speed and perceptual speed, which together with decision time and movement time defined a general factor G*s*. However, the IT-perceptual speed correlation was entirely accounted for by correlation between visualization speed and perceptual speed. In short, this study found four different kinds of speed, with IT relating to only one. The correlation between IT and IQ depended on G*s* via visualization speed, defined in terms of an ability to visualize complex rules, principally about how triplets of ordinal digits were presented. However, odd-man-out DT also had its strongest loading on visualization speed, contrary to Burns and Nettelbeck's (2003) result. Thus, whether IT taps processes different from those measured by the odd-man-out task remains unresolved.

Basic Processes

Belief is now widespread that measures of IT and RT tap individual differences in a fundamental, biological property of the CNS that limits speed of information processing (Madden, 2001). Nonetheless, evidence for this theory is suggestive rather than conclusive. As Mackintosh (1998, p. 246) has pointed out, that correlation between IQ and RT principally reflects a capacity of those with higher IQ to avoid the slower responding that characterizes the performance of those with lower IQs means that RT must involve more than the speed of nerve conduction.

Event-related potential (ERP) recordings made at the scalp, of changes in cortical activity following presentation of target stimuli, have found correlations between IQ and the latency, rise time, amplitude, and complexity of wave forms, particularly the positive peaks found approximately 100–300 ms after stimulus onset (Deary, 2000). However, Deary has cautioned against accepting that such results establish direct links between intelligence and basic biological speed differences. Limits to current knowledge mean that there is uncertainty about the nature of ongoing brain activities that are captured by the ERP (Burns, Nettelbeck, & Cooper, 2000). For example, these may reflect "neural adaptability" (Schafer, 1985) – that is, the effectiveness of processing strategies, not differences in speed of neuronal transmission.

Some criticism of RT as a strong index of processing speed stems from suggestions that faster RTs are more reflective of higher-IQ individuals' superior learning techniques (Mackintosh, 1986). Methodologically, this may allow these individuals to "cheat" the myriad ECTs employing patterned and/or predictable response systems, which may actually require minimal working memory capacity (WMC). Broadly speaking, learning skills like this may potentiate higher-intelligence outcomes through learned G*c* abilities (Ritchie et al., 2013) relating to intelligence and cognitive tests, as opposed to true G*f* abilities. Meiran and Shahar (2018) attempted to eliminate this confound by increasing task complexity. They introduced "arbitrary mapping" into an ECT response system, which theoretically controlled for learning effects and required a significantly higher WMC to respond correctly, thereby increasing the predictive value of RTs in tandem with greater task complexity (Sheppard & Vernon, 2008). Specifically, they hypothesized that tau values would be strongly influenced by WM-load manipulation, in line with previous research (Schmiedek et al., 2007). As predicted, when ECTs used "arbitrary mapping" (i.e., WM load was high) it was mainly tau that predicted fluid intelligence, and when mapping was nonarbitrary the latent variables of mu and tau were almost evenly correlated with fluid intelligence, in line with Jensen's (1982) prediction.

A recent procedure developed by Sculthorpe, Stelmack, and Campbell (2009) as a variant on the widely used "oddball" ERP task may have potential for addressing this theoretically important distinction. Sculthorpe and colleagues' task differed from the parent version in a number of respects not important here but, in common, required detection of occasional deviant auditory stimuli located within a common pattern of tone sequences. Critically, their version included both an active detection condition and a passive condition (concurrent reading task with the sequence of tone stimuli presented but ignored). Electrophysiological responses to the unattended deviant stimuli were measured by "mismatch negativity" (MMN) – amplitude departures from the standard level of activity (regular tone patterns) in the time frame 110–350 ms following a deviant stimulus. As predicted by earlier research, higher-IQ participants were more effective (shorter latencies, higher amplitudes in the ERP P300 component – a positive component found around 300 msec post stimulus – and shorter, less variable RTs) at detecting pattern "violations." Most importantly, similar results held for MMN in the passive condition. The authors argued that because attention was focused on the reading task in the passive condition, these results excluded involvement of higher-level conscious processes. This argument relies on the difficult-to-confirm assumption that participants complied with instructions, but comparison of average ERP waveforms across the active and passive conditions was consistent with this interpretation and this paradigm therefore offers promise for future research of this kind.

There have been attempts to relate intelligence to more direct measures of speed of information transmission in the CNS. Thus, Vernon and Mori (1992) reported low–moderate correlations between peripheral nerve conduction velocity (NCV) in the arms, general RT extracted from several RT tasks, and general psychometric intelligence, but also found that the RT-IQ correlation did not depend on NVC. Reed and Jensen (1992; Reed, Vernon, & Johnson, 2004) tried to estimate individual differences in brain NVC and correlate these with measures of intelligence and with RT (Reed & Jensen, 1993). However, although Reed and Jensen (1993) found low but statistically significant correlations between NCV and nonverbal IQ and between nonverbal IQ and choice RT, the expected correlation between NCV and choice RT was not found. Reviews of these and similar studies have concluded that results have not been convincing (Deary, 2000; Vernon et al., 2000).

Strachan and colleagues (2001) attempted to clarify relations between NCV, psychometric speeded tasks, and ECTs by experiment. They manipulated the blood glucose levels of healthy participants, while measuring performance on RT and IT. As predicted by knowledge about the effects of hypoglycemia, lowered blood glucose resulted in significant slowing on all tasks; but this did not affect the velocity of motor nerve conduction in the arms or legs of participants. This result suggests that speed measured by these tasks is not at the level of nerve conduction. Although differences in neural transmission time may account for some small part of variance in cognitive functioning, RT and IT differences do not appear to reflect these.

Recent twin studies have reported that IT has moderate heritability (Edmonds et al., 2008; Luciano et al., 2001; Luciano et al., 2004; Posthuma, de Geus, & Boomsma, 2001). Correlation between IT and IQ has been accounted for by common genetic influences. Patterns of results have been similar for children, adolescents, young adults, and middle-aged adults and for males and females. Consistent results have been found for two-choice RT (Luciano et al., 2004).

Demonstrating that a trait is in part heritable implicates biological processes but does not of itself establish that these are low-level, as opposed to top-down strategic processes. A demonstration by Deary and colleagues (2001b) using functional magnetic resonance imaging (fMRI) technology during IT performance is similarly difficult to interpret. Deary and colleagues found that areas of brain activation during a difficult discrimination condition (short SOA) and deactivation during an easy condition (long SOA) overlapped with areas in the lateral frontal cortex that Duncan and colleagues (2000) proposed are the basis for *g*. These results are consistent with the theory that IT and abstract problem-solving share common processes but do not reveal the direction of causality. Luciano and colleagues (2004) acknowledged that their results would be equally well explained by a top-down explanation involving attention. Similarly, Edmonds and colleagues (2008) noted substantial correlations between IT and neuropsychological functions, including attention/executive, language, and memory, all of which were substantially correlated with IQ. However, Posthuma and colleagues (2001) have considered a bottom-up account more likely. Drawing on research into conduction velocity in early visual pathways in the monkey brain, they concluded that "genes related to CNS axonal conduction velocity constitute good candidate genes for intelligence" (p. 601). Similarly, both Luciano and colleagues (2004) and Edmonds and colleagues (2008) have speculated that processing speed may be related to basic brain characteristics, like the quality of axonal myelination.

A promising line of enquiry, supporting the theory that IT does measure basic processes underpinning intelligence, has been pointed by Stough and colleagues (reviewed by Stough et al., 2001b). Their research derived from initial observation that acute nicotine dosage improves speed of processing, vigilance, attention, and memory. Pharmacological theory has implicated nicotine in enhanced synaptic transfer of acetylcholine. By systematically testing changes in IT coincident with neurochemical interventions, Stough and others have demonstrated that administering nicotine enhances IT whereas blocking nicotinic receptors impairs IT. Other neurotransmitters – serotonin, noradrenaline, and dopamine – which also contribute to effective cognitive performance were found not to influence IT (Thompson et al., 2000; Nathan, Stough, & Siteram, 2000). Stough and colleagues (2001b) have therefore proposed that IT is specifically a marker for the integrity of the cholinergic system (for detailed studies see Hutchison et al., 2001; Nathan & Stough, 2001), which uses acetylcholine to transmit nerve impulses, and is involved in regulation of memory and learning. These ideas align with the suggestion that processing speed provides a necessary but insufficient condition for intelligence (Nettelbeck & Wilson, 1985) and with Detterman's (1982) model for intelligence as a system of different cognitive functions. A relatively recent biological study by Chevalier and colleagues (2015) reported that increased myelination in several areas of the brain was associated with better IT or faster processing speed in early childhood, again supporting a biological basis for IT/processing speed and contributing to the idea that developmental changes in brain maturation or neurotransmitter activity could underlie changes in cognition with development.

Furthermore, a study of children with attention deficit hyperactivity disorder (ADHD) by Galloway-Long & Huang-Pollock (2018) found longer/slower SD and tau on IT tasks, compared to non-ADHD controls, reflecting previously observed differences of other intelligence indices, such as executive control and academic performance (Biederman et al., 2004).

Conclusions

After more than a century beyond Galton's speculations about the bases of intelligence, a growing body of evidence provides support for his ideas. An improved understanding of processing speed will prove fundamental to an understanding of intelligence, but current evidence suggests that speed constructs will not provide a sufficient explanation and, moreover, the influence of speed may be manifest by different pathways. Although the extent to which IT and RT measure the same or different processes is still an open question, there is, however, compelling evidence that correlation between IQ and processing speed estimated by IT or choice RT reflects shared genetic influences. Although these influences might implicate higher-order strategic-based processing, the current balance of opinion appears to favor a role for basic perceptual processes. These may rely on the quality of brain white matter communication systems, perhaps even at the level of chemical neurotransmitters responsible for specific functions, although this has not been established, and currently there is uncertainty about the influence of white matter abnormalities, which increase with normal aging, on cognitive functioning among healthy elderly persons. There is considerable evidence that white matter lesions are associated with slower processing speed and poorer performance on tests of attention and memory (Gunning-Dixon & Raz, 2000). However, whereas some researchers have found no evidence to link the extent of lesions to intelligence (Gunning-Dixon & Raz, 2000; Rabbitt et al., 2007), others have (Deary et al., 2006; Deary et al., Whalley, 2003). Deary's studies are persuasive because they have controlled for prior IQ. They found that both IQ measured at age eleven and contemporaneous white matter integrity independently accounted for variance in general cognitive ability in elderly participants, with the latter mediated by standard deviation for simple RT. Moreover, IQ at age eleven predicted both general cognitive ability and white matter integrity some seventy years later. By this account, cognitive integrity throughout life reflects white matter integrity, which determines efficiency of information processing. This is an intriguing scenario but, clearly, further research is required that better defines more comprehensive models for processing speed, psychometric intelligence, and white matter structures. Additionally, new studies examining cellular processes may also be complementary in terms of brain function.

For instance, a new study by Camfield and colleagues (2019) has revealed interesting relationships between biomarkers of oxidative stress (F2 isoprostanes) and processing speed measured by the Jensen apparatus, suggesting that an understanding of the functional activity at the cellular level may also be important in understanding processing speed. Future work examining other key cellular processes such as inflammatory processes mediated by glial cells as well as other biomarkers known to influence cognition (e.g., amyloid load) could also provide worthwhile insights into individual differences in processing speed and cognitive function and better establish causal biological mechanisms for the relationship between processing speed and cognitive ability.

Future Directions

The forgoing account has identified the major questions that future research should attempt to address. An important next step is to determine whether different kinds of speed are required to account for differences in intelligence. It is possible that different ECTs tap different processes underlying different components, all of which contribute to individual differences in intelligence. However, identifying different kinds of speed would not rule out the possibility that there are also individual differences in a general speed factor that reflects some fundamental biological constraint and that has some important explanatory value for understanding differences in higher-level abilities. Thus, a clearer definition of basic processes requires that commonalities and specificities within batteries of speeded tasks that encompass a range of cognitive demands from simple to more complex are first identified. On current evidence, there should be a focus on response variability, rather than relying on measures of central tendency. These endeavors should be theory driven and based on more comprehensive, multivariate models for intelligence than have typically been applied in the past and should attempt, moreover, to encourage closer collaboration between the cognitive, neurological, and psychometric traditions.

Promising directions have been pointed by attempts to establish links between speeded performance and biochemical and neurophysiological features of the brain. Attempts to test the adequacy of statistical models that include the independent contribution of both higher-order cognitive constructs and speed variables to intelligence also have potential to improve understanding in reductionist terms. And if it can be established that prior levels of speed and/or changes in speed precede subsequent cognitive changes, this finding would provide powerful evidence for a causal relationship. Research that addresses developmental cascade theory across a longitudinal timeframe, both with children and with elderly adults, would contribute to knowledge here. Of course, it is possible that changing processing speed during childhood and old age has a different role in relation to intelligence than is the case for middle life. Moreover, although improving processing speed during normal childhood development may be the consequence of increasingly complex brain structures, which later deteriorate during normal adult aging, it is also possible that declining processing speed reflects, at least in part, different biological states than those associated with improving speed.

Finally, the major challenge is to ascertain whether the speed of bottom-up processes is primarily responsible for developmental trends and individual differences in higher reasoning abilities, as opposed to whether speed differences are the consequence of top-down strategic functions, or whether both mechanisms interact. These are open questions that so far have proved difficult to resolve, but it is already clear that the potential utility of bottom-up explanation does not exclude the possibility that higher-order functions influenced by responding strategies can have a nontrivial explanatory role. In fact, future confirmation that the brain's neural structures have potential to change in response to idiosyncratic behaviors and experience (Doidge, 2007) would point toward the theory that bottom-up and top-down processes are inextricably linked. Advances in functional imaging with functional magnetic resonance imaging or magnetoelectroencephalography (MEG) are obvious new methods with appropriate spatial and temporal resolution that may shed some light on to this question.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence. *Psychological Bulletin*, *131*, 30–60.

Anderson, M. (1988), Inspection time, information processing and the development of intelligence. *British Journal of Developmental Psychology*, *6*, 43–57. https://doi.org/10.1111/j.2044-835X.1988.tb01079.x

Anderson, M., Nettelbeck, T., & Barlow, J. (1997). Reaction time measures of speed of processing: Speed of response selection increases with age but speed of stimulus categorization does not. *British Journal of Developmental Psychology*, *15*, 145–157.

Anderson, M., Reid, C., & Nelson, J. (2001). Developmental changes in inspection time: What a difference a year makes. *Intelligence*, *29*, 475–486.

Baumeister, A. A., & Kellas, G. (1968). Reaction time and mental retardation. In N. R. Ellis (Ed.), *International review of research in mental retardation* (vol. 3, pp. 163–193). New York: Academic Press.

Biederman, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F., et al. (2004). Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of Consulting and Clinical Psychology*, *72*(5), 757–766. https://doi.org/10.1037/0022-006X.72.5.757

Birren, J. E., & Fisher, L. M. (1992). Aging and slowing of behavior: Consequences for cognition and survival. In T. B. Sonderegger (Ed.), *Nebraska Symposium on Motivation 1991* (pp. 1–37). Lincoln, NE: University of Nebraska Press.

Bonney, K. R., Almeida, O. P., Flicker, L., Davies, S., Clarnette, R., Anderson, M., et al. (2006). Inspection time in non-demented older adults with mild cognitive impairment. *Neuropsychologia*, *44*, 1452–1456.

Bors, D. A., Stokes, T. L., Forrin, B., & Hodder, S. L. (1999). Inspection time and intelligence: Practice, strategies and attention. *Intelligence*, *27*, 111–129.

Brand, C. R. (1996). *The g factor: General intelligence and its implications*. Chichester, UK: Wiley.

Brand, C. R., & Deary, I. J. (1982). Intelligence and "inspection time." In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 133–148). New York: Springer-Verlag.

Brewer, N., & Smith, G. A. (1984). How normal and retarded individuals monitor and regulate speed and accuracy of responding in serial choice tasks. *Journal of Experimental Psychology: General*, *113*, 71–93.

Buehner, M., Krumm, S., Ziegler, M., & Pluecken, T. (2006). Cognitive abilities and their interplay: Reasoning, crystallized intelligence, working memory components, and sustained attention. *Journal of Individual Differences*, *27*, 57–72.

Burns, N. R., & Nettelbeck, T. (2003). Inspection time in the structure of cognitive abilities: Where does IT fit? *Intelligence*, *31*, 237–255.

Burns, N. R., Nettelbeck, T., & Cooper, C. J. (2000). Event-related potential correlates of some human cognitive ability constructs. *Personality and Individual Differences*, *29*, 157–168.

Burns, N. R., Nettelbeck, T., McPherson, J., & Stankov, L. (2007). Perceptual learning on inspection time and motion perception. *Journal of General Psychology*, *134*, 83–100.

Camfield, D. A., Nolidin, K., Savage, K., Timmer, J., Croft, K., Simpson, T., et al. (2019). Higher plasma levels of F2-isoprostanes are associated with slower psychomotor speed in healthy older adults. *Free Radical Research*, *53*(4), 377–386. https://doi.org/10.1080/10715762.2018.1513133

Carlson, J. S., Jensen, C. M., & Widaman, K. (1983). Reaction time, intelligence and attention. *Intelligence*, *7*, 329–344.

Carroll, J. B. (1987). Jensen's mental chronometry: Some comments and questions. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy* (pp. 297–301, 310–311). New York: Falmer.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge, UK: Cambridge University Press.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*, 67–83.

Chevalier, N., Kurth, S., Doucette, M. R., Wiseheart, M., Deoni, S. C. L., Dean, D. C., et al. (2015). Myelination is associated with processing speed in early childhood: Preliminary insights. *PLoS One*, 10, e0139897.

Chuderski, A. (2013). When are fluid intelligence and working memory isomorphic and when are they not? *Intelligence*, *41*(4), 244–262. https://doi.org/10.1016/j.intell.2013.04.003

Chuderski, A. (2015). The broad factor of working memory is virtually isomorphic to fluid intelligence tested under time pressure. *Personality and Individual Differences*, 85, 98–104. https://doi.org/10.1016/j.paid.2015.04.046

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.

Coyle, T. R. (2003). A review of the worst performance rule: Evidence, theory, and alternative hypotheses. *Intelligence*, *31*, 567–587.

Crawford, J. R., Deary, I. J., Allan, K. M., & Gustafsson, J. E. (1998). Evaluating competing models of the relationship between inspection time and psychometric intelligence. *Intelligence*, *26*, 27–42.

Danthiir, V., Burns, N. R., Nettelbeck, T., Wilson, C., & Wittert, G. (2009). Relationships between age, processing speed, working memory, inhibition and fluid intelligence in older adults. Paper presented at the International Society for the Study of Individual Differences, Chicago, July18–22.

Danthiir, V., Wilhelm, O., & Schacht, A. (2005a). Decision speed in intelligence tasks: Correctly an ability? *Psychology Science*, *47*, 200–229.

Danthiir, V., Wilhelm, O., Schulze, R., & Roberts, R. D. (2005b). Factor structure and validity of paper-and-pencil measures of mental speed: Evidence for a higher-order model? *Intelligence*, *33*, 491–514.

Deary, I. J. (2000). *Looking down on human intelligence: From psychophysics to the brain*. Oxford: Oxford University Press.

Deary, I. J. (2003). Reaction time and psychometric intelligence: Jensen's contributions. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 53–75). Amsterdam: Pergamon.

Deary, I. J., Allerhand, M., & Der, G. (2009). Smarter in middle age, faster in old age: A cross-lagged panel analysis of reaction time and cognitive ability over 13 years in the West of Scotland Twenty-07 study. *Psychology and Aging*, *24*, 40–47.

Deary, I. J., Bastin, M. E., Pattie, A., Clayden, J. D., Whalley, L. J., Starr, J. M., et al. (2006). White matter integrity and cognition in childhood and old age. *Neurology*, *66*, 505–512.

Deary, I. J., Der, G., & Ford, G. (2001a). Reaction times and intelligence differences: A population-based cohort study. *Intelligence*, *29*, 389–399.

Deary, I. J., Hunter, R., Langan, S. J., & Goodwin, G. M. (1991). Inspection time, psychometric intelligence and clinical estimates of cognitive ability in pre-senile Alzheimer's disease and Korsakoff's psychosis. *Brain*, *114*, 2543–2554.

Deary, I. J., Leaper, S. A., Murray, A. D., Staff, R. T., & Whalley, L. J. (2003). Cerebral white matter abnormalities and lifetime cognitive change: A 67-year follow-up of the Scottish Mental Survey of 1932. *Psychology and Aging*, *18*, 140–148.

Deary, I. J., Simonotto, E., Marshall, A., Marshall, I., Goddard, N., & Wardlaw, J. M. (2001b). The functional anatomy of inspection time: A pilot fMRI study. *Intelligence*, *29*, 497–510.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time: Achievements, prospects and problems. *American Psychologist*, *51*, 599–608.

Deary, I. J., & Stough, C. (1997). Looking down on intelligence. *American Psychologist*, 52, 1148–1150.

Deluca, J., & Kalmar, J. H. (2007). *Information processing speed in clinical populations*. New York: Psychology Press.

Demetriou, A., Spanoudis, G., & Shayer, M. (2013a). Developmental intelligence: From empirical to hidden constructs. *Intelligence*, *41*(5), 744–749. https://doi.org/10.1016/j.intell.2013.07.014

Demetriou, A., Spanoudis, G., Shayer, M., Mouyi, A., Kazi, S., & Platsidou, M. (2013b). Cycles in speed-working memory-G relations: Towards a developmental–differential theory of the mind. *Intelligence*, *41*(1), 34–50. https://doi.org/10.1016/j.intell.2012.10.010

Detterman, D. K. (1982). Does g exist? *Intelligence*, *6*, 99–108.

Detterman, D. K. (1987). What does reaction time tell us about intelligence? In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 177–200). Norwood, NJ: Ablex.

Doidge, N. (2007). *The brain that changes itself*. New York: Viking Press.

Duncan, J., Seitz, R. J., Koldny, J., Bor, D., Herzog, H., Ahmed, A., et al. (2000). A neural basis for general intelligence. *Science*, *289*, 457–460.

Edmonds, C. J., Isaacs, E. B., Visscher, P. M., Rogers, M., Lanigan, J., Singhal, A., et al. (2008). Inspection time and cognitive abilities in twins aged 7 to 17 years: Age-related changes, heritability and genetic covariance. *Intelligence*, *36*, 210–255.

Egan, V. (1994). Intelligence, inspection time and cognitive strategies. *British Journal of Psychology*, *85*, 305–316.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Eysenck, H. J. (1987). Speed of information processing, reaction time, and the theory of intelligence. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 21–67). Norwood, NJ: Ablex.

Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2007). Age changes in processing speed as a leading indicator of cognitive aging. *Psychology and Aging*, *22*, 558–568.

Fox, M. C., Roring, R. W., & Mitchum, A. L. (2009). Reversing the speed-IQ correlation: Intra-individual variability and attentional control in the inspection time paradigm. *Intelligence*, *37*, 76–80.

Frearson, W., & Eysenck, H. J. (1986). Intelligence, reaction time (RT) and a new "odd-man-out" RT paradigm. *Personality and Individual Differences*, *7*, 807–817.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*, *54*, 1–34.

Galloway-Long, H., & Huang-Pollock, C. (2018). Using inspection time and ex-Gaussian parameters of reaction time to predict executive functions in children with ADHD. *Intelligence*, 69, 186–194. https://doi.org/10.1016/j.intell.2018.06.005

Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Harper and Row.

Gregory, T., Callaghan, A., Nettelbeck, T., & Wilson, C. (2009). Inspection time predicts individual differences in everyday functioning among elderly adults: Testing discriminant validity. *Australasian Journal on Ageing*, *28*, 87–92.

Gregory, T., Nettelbeck, T., Howard, S., & Wilson, C. (2008). Inspection time: A biomarker for cognitive decline. *Intelligence*, *36*, 664–671.

Gregory, T., Nettelbeck, T., Howard, S., & Wilson, C. (2009a). A test of the cascade model in the elderly. *Personality and Individual Differences*, *46*, 71–73.

Gregory, T., Nettelbeck, T., & Wilson, C. (2009b). Within-person changes in inspection time predict memory. *Personality and Individual Differences*, *46*, 741–743.

Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence*, *29*, 523–535.

Gunning-Dixon, F. M., & Raz, N. (2000). The cognitive correlates of white matter abnormalities in normal aging: A quantitative review. *Neuropsychology*, *14*, 224–232.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2008). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest*, *9*, 1–65.

Hick, W. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11–26.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53–91). New York: Guilford Press.

Hunt, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology*, *71*, 449–474.

Hutchison, C. W., Nathan, P. J., Mrazek, L., & Stough, C. (2001). Cholinergic modulation of speed of early information processing: The effect of donepezil on inspection time. *Psychopharmacology*, *44*, 440–442.

Jensen, A. R. (1982). Reaction time and psychometric g. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93–132). New York: Springer-Verlag.

Jensen, A. R. (1987). Individual differences in the Hick paradigm. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 101–175). Norwood, NJ: Ablex.

Jensen, A. R. (1998). *The g factor: the science of mental ability*. New York: Praeger.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam: Elsevier.

Juhel, J. (1991). Relationships between psychometric intelligence and information-processing speed indexes. *European Bulletin of Cognitive Psychology*, *11*, 73–105.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin*, *109*, 490–501.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Brown, J., & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*. https://doi.org/10.1016/j.intell.2009.03.004

Krumm, S., Schmidt-Atzert, L., Michalczyk, K., & Danthiir, V. (2008). Speeded paper-pencil sustained attention and mental speed tests. *Journal of Individual Differences*, *29*, 205–216.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity? *Intelligence*, *14*, 389–433.

Langner, R., & Eickhoff, S. B. (2013). Sustaining attention to simple tasks: A meta-analytic review of the neural mechanisms of vigilant attention. *Psychological Bulletin*, 139(4), 870–900. https://doi.org/10.1037/a0030694

Larson, G. E., & Alderton, D. L. (1990). Reaction time variability and intelligence: A "worst performance" analysis of individual differences. *Intelligence*, *14*, 309–325.

Lehrl, S., & Fischer, B. (1990). A basic information psychological parameter (BIP) for the reconstruction of concepts of intelligence. *European Journal of Personality*, *4*, 259–286.

Longstreth, L. E. (1984). Jensen's reaction time investigations of intelligence: A critique. *Intelligence*, *8*, 139–160.

Luciano, M., Smith, G. A., Wright, M. J., Geffen, G. M., Geffen, L. B., & Martin, N. G. (2001). On the heritability of inspection time and its covariance with IQ: A twin study. *Intelligence*, *29*, 443–457.

Luciano, M., Wright, M. J., Geffen, G. M., Geffen, L. B., Smith, G. A., & Martin, N. G. (2004). A genetic investigation of the covariation among inspection time, choice reaction time, and IQ subtest scores. *Behavior Genetics*, *34*, 41–50.

Mackintosh, N. J. (1986). The biology of intelligence? *British Journal of Psychology*, *77*, 1–18.

Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford: Oxford University Press.

Mackintosh, N. J., & Bennett, E. S. (2002). IT, IQ and perceptual speed. *Personality and Individual Differences*, *32*(4), 685–693.

Madden, D. J. (2001). Speed and timing in behavioral processes. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 288–312). San Diego, CA: Academic Press.

Marr, D. B., & Sternberg, R. J. (1987). The role of mental speed in intelligence: A triarchic perspective. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 271–294). Norwood, NJ: Ablex.

Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, *17*, 433–442.

McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (2nd ed., pp. 156–182). New York: Guilford.

Meiran, N., & Shahar, N. (2018). Working memory involvement in reaction time and its contribution to fluid intelligence: An examination of individual differences in reaction-time distributions. *Intelligence*, *69*, 176–185. https://doi.org/10.1016/j.intell.2018.06.004

Nathan, P. J., & Stough, C. (2001). Inspection time: A neuropsychophysiological test for measuring the functional integrity of the cholinergic system. *Medical Hypotheses*, 57, 759–760.

Nathan, P. J., Stough, C., & Siteram, G. (2000). Serotonin and information processing: A pharmacodynamic study on the effects of citalopram on cognitive and psychomotor function. *Human Psychopharmacology: Clinical and Experimental*, *15*, 306–307.

Nettelbeck, T. (1985). What reaction times time. *Behavioral and Brain Sciences*, *8*, 193–263.

Nettelbeck, T. (1987). Inspection time and intelligence. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 295–346). Norwood, NJ: Ablex.

Nettelbeck, T. (1994). Speediness. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence*. (pp. 1014–1019). New York: Macmillan.

Nettelbeck, T. (1998). Jensen's chronometric research: Neither simple nor sufficient but a good place to start. *Intelligence*, *29*, 233–241.

Nettelbeck, T. (2001). Correlation between inspection time and psychometric abilities: A personal interpretation. *Intelligence*, *29*, 459–474.

Nettelbeck, T. (2003). Inspection time and g. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 77–91). Amsterdam: Pergamon.

Nettelbeck, T., Gregory, T., Wilson, C., Burns, N., Danthiir, V., & Wittert, G. (2008). Inspection time: A marker for less successful ageing. Paper presented at the Ninth Annual Conference of the International Society for Intelligence Research (ISIR), Decatur, Georgia, December 11–13.

Nettelbeck, T., & Kirby, N. H. (1983). Measures of timed performance and intelligence. *Intelligence*, *7*, 39–52.

Nettelbeck, T., & Lally, M. (1976). Inspection time and measured intelligence. *British Journal of Psychology*, *67*, 17–22.

Nettelbeck, T., & Vita, P. (1992). Inspection time in two childhood age cohorts: A constant of a developmental function? *British Journal of Developmental Psychology*, *10*, 189–198.

Nettelbeck, T., & Wilson, C. (1985). A cross-sequential analysis of developmental differences in speed of visual information processing. *Journal of Experimental Child Psychology*, *40*, 1–22.

Nettelbeck, T., & Wilson, C. (1997). Speed of information processing and cognition. In W. E. J. Maclean (Ed.), *Ellis' handbook of mental deficiency, psychological theory and research* (3rd ed., pp. 245–274). Mahwah, NJ: Lawrence Erlbaum.

Nettelbeck, T., & Young, R. (1989). Inspection time and intelligence in 6 year old children. *Personality and Individual Differences*, *10*, 605–614.

Neubauer, A. C. (1997). The mental speed account to the assessment of intelligence In J. S. Carlson, J. Kingma, & W. Tomic (Eds.), *Advances in cognition and educational practice: Reflections on the concept of intelligence* (vol. 4, pp. 149–173). Greenwich, CT: JAI Press.

Neubauer, A. C., & Bucik, V. (1996). The mental speed-IQ relationship: Unitary or modular? *Intelligence*, *22*, 23–48.

Oberauer, K., Süβ, H.-M., Wilhelm, O., & Wittmann, W. W. (2008). Which working memory functions predict intelligence? *Intelligence*, *36*(6), 641–652. https://doi.org/10.1016/j.intell.2008.01.007

Oberauer, K., Wilhelm, O., Schulze, R., & Süß, H.-M. (2005). Working memory and intelligence – their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 61–65. https://doi.org/10.1037/0033-2909.131.1.61

O'Connor, T. A., & Burns, N. R. (2003). Inspection time and general speed of processing. *Personality and Individual Differences*, *35*, 713–724.

Olsson, H., Björkman, C., Haag, K., & Juslin, P. (1998). Auditory inspection time: On the importance of selecting the appropriate sensory continuum. *Personality and Individual Differences*, *25*, 627–634.

Parker, D. M., Crawford, J. R., & Stephen, E. (1999). Auditory inspection time and intelligence: A new spatial localization task. *Intelligence*, *27*, 131–139.

Petrill, S. A., Luo, D., Thompson, L. A., & Detterman, D. K. (2001). Inspection time and the relationship among elementary cognitive tasks, general intelligence, and specific cognitive abilities. *Intelligence*, *29*, 487–496.

Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Lawrence Erlbaum.

Posthuma, D., de Geus, E. J. C., & Boomsma, D. I. (2001). Perceptual speed and IQ are associated through common genetic factors *Behavior Genetics*, *31*, 593–602.

Rabbitt, P., Scott, M., Lunn, M., Thacker, N., Lowe, C., Pedleton, N., et al. (2007). White matter lesions account for all age-related declines in speed but not in intelligence. *Neuropsychology*, *21*, 363–370.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367.

Reed, T. E., & Jensen, A. R. (1992). Conduction velocity in a brain nerve pathway of normal adults correlates with intelligence level. *Intelligence*, *16*, 259–272.

Reed, T. E., & Jensen, A. R. (1993). Choice reaction time and visual pathway nerve conduction velocity both correlate with intelligence but appear not to correlate with each other: Implications for information processing. *Intelligence*, *17*, 191–203.

Reed, T. E., Vernon, P. A., & Johnson, A. M. (2004). Confirmation of correlation between brain nerve conduction velocity and intelligence level in normal adults. *Intelligence*, *32*, 563–572.

Ritchie, S. J., Bates, T. C., Der, G., Starr, J. M., & Deary, I. J. (2013). Education is associated with higher later life IQ scores, but not with faster cognitive processing speed. *Psychology and Aging*, *28*(2), 515–521. https://doi.org/10.1037/a0030820

Roberts, R. D., & Stankov, L. (1999). Individual differences in speed of mental processing and human cognitive abilities: Toward a taxonomic model. *Learning and Individual Differences*, *11*, 1–120.

Rockstroh, S., & Schweizer, K. (2004). The effect of retest practice on the speed-ability relationship. *European Psychologist*, *9*, 24–31.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428.

Salthouse, T. A. (2006). Mental exercise and mental aging. *Perspectives on Psychological Science*, *1*, 68–87.

Schafer, E. P. W. (1985). Neural adaptability: A biological determinant of *g* factor intelligence. *Behavioral and Brain Sciences*, *8*, 240–241.

Schaie, K. W. (2005). *Developmental influences on adult intelligence*. Oxford: Oxford University Press.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*, 414–429.

Schweizer, K., Zimmermann, P., & Koch, W. (2000). Sustained attention, intelligence, and the crucial role of perceptual processes. *Learning and Individual Differences*, *12*, 271–287.

Sculthorpe, L. D., Stelmack, R. M., & Campbell, K. B. (2009). Mental ability and the effect of pattern violation discrimination on P300 and mismatch negativity. *Intelligence*, *37*, 405–411.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*, *44*, 535–551.

Smith, G. A., & Carew, M. (1987). Decision time unmasked: Individuals adopt different strategies. *Australian Journal of Psychology*, *39*, 339–351.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum.

Sternberg, R. J. (2003). *Wisdom, intelligence and creativity synthesized*. Cambridge, UK: Cambridge University Press.

Sternberg, S. (1975). Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology*, *27*, 1–32.

Stough, C., Bates, T., Mangan, G. L., & Colrain, I. (2001a). Inspection time and intelligence: Further attempts at reducing the apparent motion strategy. *Intelligence*, 29, 219–230.

Stough, C., Nettelbeck, T., Cooper, C., & Bates, T. (1995). Strategy use in Jensen's RT Paradigm: Relationships to intelligence? *Australian Journal of Psychology*, *47*, 61–65.

Stough, C., Thompson, J. C., Bates, T. C., & Nathan, P. J. (2001b). Examining neurochemical determinants of inspection time: Development of a biological model. *Intelligence*, *29*, 511–522.

Strachan, M. W. J., Deary, I. J., Ewing, F. M. E., Ferguson, S. S. C., Young, M. J., & Frier, B. M. (2001). Acute hypoglycemia impairs the functioning of the central but not peripheral nervous system. *Physiology and Behavior*, *72*, 83–92.

Stroud, J. M. (1956). The fine structure of psychological time. In H. Quastler (Ed.), *Information theory in psychology*. Glencoe, UK: The Free Press.

Thompson, J. C., Stough, C., Nathan, P. J., Ames, D., & Ritchie, C. (2000). Effects of the nicotinic antagonist mecamylamine on inspection time. *Psychopharmacology*, *150*(1), 117–119.

Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. *Psychological Review*, *80*, 1–52.

Unsworth, N. (2010). Interference control, working memory capacity, and cognitive abilities: A latent variable analysis. *Intelligence*, *38*(2), 255–267. https://doi.org/10.1016/j.intell.2009.12.003

Verhaeghen, P. (2013). *The elements of cognitive aging: Meta-analyses of age-related differences in processing speed and their consequences*. Oxford: Oxford University Press.

Vernon, P. A. (1987). New developments in reaction time research. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 1–20). Norwood, NJ: Ablex.

Vernon, P. A., & Mori, M. (1992). Intelligence, reaction times, and peripheral nerve conduction velocity. *Intelligence*, *16*, 273–288.

Vernon, P. A., Wickett, J. C., Bazana, P. C., & Stelmack, R. M. (2000). The neuropsychology and psychophysiology of human intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence*. Cambridge, UK: Cambridge University Press.

Vickers, D., Nettelbeck, T., & Willson, R. J. (1972). Perceptual indices of performance: The measurement of "inspection time" and "noise" in the visual system. *Perception*, *1*, 263–295.

Welford, A. T. (1968). *Fundamentals of skill*. London: Methuen.

Zajac, I. T., & Burns, N. R. (2007). Measuring auditory inspection time in primary school children. *Journal of Individual Differences*, *28*, 45–52.

21 Working Memory and Intelligence

Andrew R. A. Conway and Kristof Kovacs

> We want to understand intelligence, not only map its network of correlations with other constructs. This means to reveal the functional – and ultimately, the neural – mechanisms underlying intelligent information processing. Among the theoretical constructs within current theories of information processing, [working memory capacity] WMC is the one parameter that correlates best with measures of reasoning ability, and even with g_f and g. Therefore, investigating WMC, and its relationship with intelligence, is psychology's best hope to date to understand intelligence.
>
> Oberauer and colleagues (2005, p. 64)

Working memory (WM) is a construct developed by cognitive psychologists to characterize and help further investigate how human beings maintain access to goal-relevant information in the face of concurrent processing and/or distraction. For example, suppose you are conducting an internet search to find information about an intelligence researcher. To conduct the search you need to remember the researcher's name and institution, and perhaps a few keywords about their work. WM is required to keep these pieces of information in mind while typing and then navigating the search results. Many important cognitive behaviors, beyond searching the Internet – such as reading, reasoning, and problem-solving – require WM because for each of these activities, some information must be maintained in an accessible state while new information is processed and potentially distracting information is ignored.

Working memory is a limited-capacity system. That is, there is only so much information that can be maintained in an accessible state at one time. There is also substantial variation in WM capacity (WMC) across individuals: Older children have greater capacity than younger children, the elderly tend to have lesser capacity than younger adults, and patients with certain types of neural damage or disease have lesser capacity than healthy adults. There is even a large degree of variation in WMC within healthy adult samples of subjects, such as within-college student samples.

It is important to clarify at the outset the distinction between working memory and working memory capacity. Working memory refers to the cognitive system required to maintain access to information in the face of concurrent processing and/or distraction (including mechanisms involved in stimulus representation, maintenance, manipulation, and retrieval), while working memory capacity refers to the maximum amount of information an individual can maintain in a particular task that is designed to measure some aspect(s) of WM.

The focus of this chapter is on the relationship between WMC and fluid intelligence (G*f*) in healthy young adults. Two meta-analyses, conducted by different groups of researchers, estimate the correlation between WMC and G*f* to be somewhere between $r = 0.72$ (Kane, Hambrick, & Conway, 2005) and $r = 0.85$ (Oberauer et al., 2005). These estimates are remarkably consistent with a recent large sample study ($N = 2,200$) that found a correlation of $r = 0.77$ between WMC and G*f* (Gignac, 2014). Thus, according to these analyses, WMC accounts for at least half the variance in G*f*. This is impressive, yet for this line of work to truly inform theoretical accounts of intelligence, we need to better understand the construct of WM and discuss the various ways in which it is measured.

The emphasis here is on fluid intelligence rather than crystallized intelligence, general intelligence (*g*), or intelligence more broadly defined, because most of the research linking WM to the concept of intelligence has focused on fluid abilities and reasoning rather than on acquired knowledge or skill (however, see Hambrick, 2003; Hambrick & Engle, 2002; Hambrick & Meinz, 2011; Hambrick & Oswald, 2005).

Fluid intelligence is defined as "the use of deliberate and controlled mental operations to solve novel problems that cannot be performed automatically" (McGrew, 2009, p. 5.). This is a natural place to focus our microscope because WM is most important in situations that do not allow for the use of prior knowledge and less important in situations in which skills and strategies guide behavior (Ackerman, 1988; Engle et al., 1999).

This chapter begins with a brief review of the history of working memory, followed by our own contemporary view of WM, which is largely shaped by Cowan's model (1988, 1995, 2001, 2005) but also incorporates ideas from individual-differences research (for a review, see Unsworth & Engle, 2007), neuroimaging experiments (for a review, see Jonides et al., 2008), and computational models of WM (Ashby et al., 2005; Oberauer et al., 2012; O'Reilly & Frank, 2006). We then discuss the measurement of WMC. These initial sections allow for a more informed discussion of the empirical work that has linked WMC and G*f*. We then consider various theories of the relationship between WMC and G*f*, with an emphasis on our new view, which we refer to as process overlap theory (Kovacs & Conway, 2016).

Historical Perspective on Working Memory

The *concept* of WM was first introduced by G. A. Miller, Galanter, and Pribram (1960) in their influential book, *Plans and the Structure of Behavior*. They proposed a dynamic and flexible short-term memory system that is necessary to structure and execute a plan. They referred to this short-term memory system as a type of "working memory" and speculated that it may be dependent upon the prefrontal cortex.

The *construct* WM was introduced in the seminal chapter by Baddeley and Hitch (1974). Prior to their work, the dominant theoretical construct used to explain short-term memory performance was the short-term store (STS), epitomized by the so-called modal model of memory popular in the late 1960s (e.g., Atkinson & Shiffrin, 1968). According to these models, the STS plays a central role in cognitive behavior, essentially serving as a gateway to further information processing. It was therefore assumed that the STS would be crucial for a range of complex cognitive behaviors, such as planning, reasoning, and problem-solving. The problem with this approach, as reviewed by Baddeley and Hitch, was that disrupting the STS with a small memory load had very little impact on people's performance on a range of complex cognitive tasks, particularly reasoning and planning (cf. Crowder, 1982). Moreover, patients with severe STS deficits – for example, a digit span of only two items – functioned rather normally on a wide range of complex cognitive tasks (Shallice & Warrington, 1970; Warrington & Shallice, 1969). This would not be possible if the STS were essential for information processing, as proposed by the modal model.

Baddeley and Hitch therefore proposed a more complex construct, *working memory*, that could maintain information in a readily accessible state, consistent with the STS, but could also engage in concurrent processing, as well as maintain access to more information than the limited capacity STS could purportedly maintain. According to this perspective, a small amount of information can be maintained via "slave" storage systems, akin to the STS, but more information can be processed and accessed via a central executive, which was poorly described in the initial WM model but has since been refined and will be discussed in more detail in the section Contemporary View of Working Memory.

Baddeley and Hitch argued that WM but not the STS plays an essential role in a range of complex cognitive tasks. According to this perspective, WMC should be more predictive of cognitive performance than the capacity of the STS. This prediction was first supported by an influential study by Daneman and Carpenter (1980), which explored the relationship between the capacity of the STS, WMC, and reading comprehension, as assessed by what then was called the Verbal section of the Scholastic Aptitude Test (SAT-V). STS capacity was assessed using a word span task, in which a series of words was presented, one word per second, and at the end of a series, the subject was prompted to recall all the words in correct serial order. Daneman and Carpenter developed a novel task to measure WMC. The task was designed to require short-term storage, akin to word span, but also to require the simultaneous processing of new information. Their *reading span* task required subjects to read a series of sentences aloud and remember the last word of each sentence for later recall. Thus, the storage and recall demands of reading span are the same as for the word span task, but the reading span task has the additional requirement of reading sentences aloud while trying to remember words for later recall. This type of task is thought to be an ecologically valid measure of the WM construct proposed by Baddeley and Hitch.

Consistent with the predictions of WM theory, the reading span task correlated more strongly with SAT-V ($r = 0.59$) than did the word span task ($r = 0.35$). This may not seem at all surprising, given that both the SAT-V and reading span involve *reading*. However, subsequent work by Turner and Engle (1989) and others showed that the processing component of the WM span task does not have to involve reading for the task to be predictive of SAT-V. They had subjects solve simple mathematical operations while remembering words for later recall and showed, consistent with Daneman and Carpenter (1980), that this task – called operation span – predicted SAT-V more strongly than did the word span task. More recent research has shown that a variety of WM span tasks, all demanding parallel processing and storage but with diverse content, are strongly predictive of a wide range of complex cognitive tasks. This suggests that the relationship between WM span performance and complex cognition is largely domain-general (e.g., Kane et al., 2004).

In sum, WM is a relatively young construct in the field of psychology. It was proposed as an alternative conception of short-term memory performance in an attempt to account for empirical evidence that was inconsistent with the modal model of memory that included an STS to explain short-term memory. Complex memory span tasks, such as reading span and operation span, were shown to be more strongly correlated with measures of complex cognition, including intelligence tests, than are simple span tasks, such as digit span and word span.

Contemporary View of Working Memory

Delineating the exact characteristics of WM and accounting for variation in WMC continues to be an extremely active area of research. There are, therefore, several current theoretical models of WM and several explanations of WMC variation. In this section we introduce just one view of WM, Cowan's model (1988, 1995, 2001, 2005), simply to provide the proper language necessary to explain WM measurement and the empirical data linking WMC to intelligence. Later in the chapter we will consider alternative theoretical accounts.

Cowan's model (see Figure 21.1) assumes that WM consists of activated long-term memory representations (see also Anderson, 1983; Atkinson & Shiffrin, 1971;

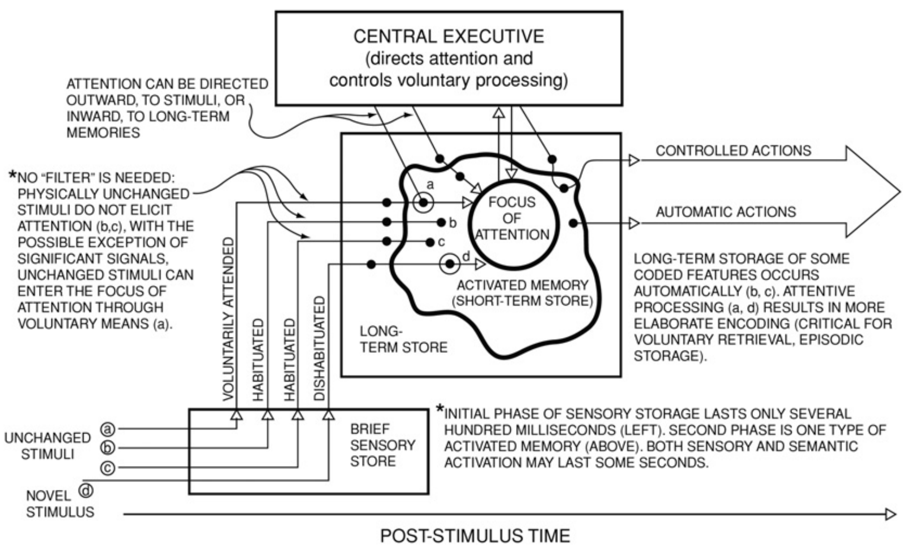

Figure 21.1 *Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system (from Cowan, 1988). Reprinted with permission of the American Psychological Association.*

Hebb, 1949) and a central executive responsible for cognitive control (for work that explains cognitive control without reference to a homuncular executive, see O'Reilly & Frank, 2006). Within this activated set of representations, or "short-term store," there is a focus of attention that can maintain approximately four items in a readily accessible state (Cowan, 2001). In other words, we can "think of" approximately four mental representations at one time.

Our own view is quite similar to the model in Figure 21.1. However, we make three modifications. First, we prefer "unitary store" models of memory rather than multiple-store models and therefore do not think of the activated portion of long-term memory (LTM) as a "store." The reason for this distinction is that there is very little neuroscientific evidence to support the notion that there is a neurologically separate "buffer" responsible for the short-term storage of information (see Postle, 2006). We acknowledge that there are memory phenomena that differ as a function of retention interval (for a review, see Davelaar et al., 2005), but we argue that these effects do not necessitate the assumption of a short-term store (for a review see Sederberg, Howard, & Kahana, 2008).

Second, recent work has shown that the focus of attention may be limited to just one item, depending on task demands (Garavan, 1998; McElree, 2001; Nee & Jonides, 2008; Oberauer, 2002). We therefore adopt Oberauer's view that there are actually three layers of representation in WM: (1) the focus of attention, limited to one item; (2) the region of direct access, limited to approximately four items; and (3) representations active above baseline but no longer in the region of direct access. To avoid confusion over Cowan versus Oberauer's terminology, we will use the phrase "scope of attention" to refer to the limited number of items that are readily accessible, recognizing that one item may have privileged access.

Third, and most important for the current chapter, we argue that Cowan's view of WMC is too limited to account for complex cognitive activity. Complex cognitive behavior, such as reasoning, reading, and problem-solving, requires rapid access to more than four items at one time. WM therefore must also consist of a retrieval mechanism that allows for the rapid retrieval of information from LTM. This notion has been referred to as long-term WM (Ericsson & Kintsch, 1995).

Thus, we view WM as consisting of at least three main components: (1) cognitive control mechanisms (or the central executive), which are most likely governed by the prefrontal cortex (PFC), anterior cingulate cortex (ACC), and subcortical structures including the basal ganglia and thalamus (Ashby et al., 2005; Botvinick, 2007; E. K. Miller & Cohen, 2001; O'Reilly & Frank, 2006); (2) one to four representations in the scope of attention, which are most likely maintained via activity in a frontoparietal network (Todd & Marois, 2004; Vogel & Machizawa, 2004); and (3) a retrieval mechanism responsible for the rapid retrieval of information from LTM. This process is most likely achieved via cortical connections from the PFC to the medial temporal lobe (MTL), including the hippocampus (Chein, Moore, & Conway, 2011; Nee & Jonides, 2008; Ranganath, 2006; O'Reilly & Norman, 2002; Unsworth & Engle, 2007).

Measurement of Working Memory Capacity

Several different WM tasks are used in contemporary research. These tasks vary in extremely important ways, which we discuss. Also, the extent to which WMC predicts G*f* is largely dependent upon which set of tasks one uses to measure WMC. Thus, a detailed discussion of various WM tasks is essential here. We mainly consider WM tasks that have shown strong correlations with measures of G*f* in a domain-general fashion, for example, a verbal WM task predicting a spatial-reasoning task and vice versa.

Complex Span Tasks

As discussed, complex span tasks, such as reading span (Daneman & Carpenter, 1980) and operation span (Turner & Engle, 1989), were designed from the perspective of the original WM model. Other complex span tasks include the counting span task (Case, Kurland, & Goldberg, 1982), as well as various spatial versions (see Kane et al., 2004; Shah & Miyake, 1996). Complex span tasks require participants to engage in some sort of simple processing task (e.g., reading unrelated sentences aloud or completing a math problem, as in reading span and operation span, respectively) between the presentations of to-be-remembered items (e.g., letters, words, digits, spatial locations). After several items have been presented, typically between two and seven, the subject is prompted to recall all the to-be-remembered items in correct serial order. A common characteristic of all complex span tasks is that they require access to information (the digits) in the face of concurrent processing (for a review of these tasks see Conway et al., 2005).

As mentioned earlier, complex span tasks reveal strong correlations with the SAT-V (r approximately 0.5; see Daneman and Carpenter, 1980, 1983; Turner and Engle, 1989) and other measures of reading comprehension (r ranging from 0.50 to 0.90 depending on the comprehension task). Complex span tasks also correlate highly with each other regardless of the processing and storage task (Turner & Engle, 1989). For example, Kane and colleagues (2004) administered several verbal and several spatial complex span tasks and the range of correlations among all the tasks was $r = 0.39$ to $r = 0.51$. Moreover, the correlation between latent variables representing spatial complex span and verbal complex span was $r = 0.84$ and the correlation between a latent variable representing all complex span tasks and G*f* was $r = 0.76$. These results suggest that complex span tasks tap largely domain-general mechanisms, which makes them good candidates for exploring the relationship between WMC and G*f*.

Simple Span Tasks

Simple span tasks (e.g., digit span, word span, letter span), in contrast to complex span, do not include an interleaved processing task between the presentation of to-be -remembered items. For example, in digit span, one digit is presented at a time, typically one per second, and after a series of digits, the subject is asked to recall the digits in correct serial order. Simple span tasks are among the oldest tasks used in memory research – for example, digit span was included in the first intelligence test (Binet, 1903) – and continue to be popular in standardized intelligence batteries (e.g., WAIS, WISC).

As discussed earlier, simple span tasks like digit span correlate less well with measures of complex cognition than complex span tasks (Conway et al., 2002; Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Engle et al., 1999; Kane et al., 2004). Also, simple span tasks are thought to be more domain-specific than complex span tasks, such that within-domain correlations among simple span tasks are higher than cross-domain correlations among simple span tasks (Kane et al., 2004).

These results would suggest that simple span tasks are not ideal candidates for exploring the relationship between WMC and G*f*. However, recent research has shown that in some situations, simple span tasks correlate as well with measures of G*f* as do complex span tasks, and in some cases they tap domain-general WM processes. We discuss three of these situations here: (1) simple span with very rapid presentation of items, known as running span; (2) simple span with spatial stimuli, known as spatial simple span; and (3) simple span with long lists of items, known as long-list simple span.

In a running memory span task (Pollack, Johnson, & Knaff, 1959), subjects are rapidly presented with a very long list of to-be-remembered items, the length of which is unpredictable. At the end of the list, the subject is prompted to recall as many of the last few items as possible. Cowan and colleagues (2005) found that running span correlates well with various measures of cognitive ability in children and adults (see also Mukunda & Hall, 1992). Cowan and colleagues argued that the rapid presentation (e.g., four items per second as compared to one item per second in digit span) prevents verbal rehearsal and that any WM memory task that prevents well-learned maintenance strategies, such as rehearsal and chunking, will serve as a good predictor of complex cognition, including G*f*.

This same explanation may demonstrate why simple span tasks with spatial stimuli tend to show strong correlations with measures of G*f* (Kane et al., 2004; Miyake et al., 2001). For example, in a computerized version of the Corsi blocks task, subjects are presented with a 4 × 4 matrix and a series of cells in the matrix flash, one location at a time, typically at a rate of one location per second. At the end of a series, the subject is required to recall the flashed locations in correct serial order. Kane and colleagues (2004) found that a latent variable derived from three spatial simple span tasks correlates as well with G*f* as a latent variable derived from three spatial complex span tasks.

Simple span tasks are also strong predictors of G*f* when only trials with long lists are considered. Reanalyzing data from Kane and colleagues (2004), Unsworth and Engle (2006a) showed that the correlation between simple span and G*f* increased as the number of to-be-remembered items in the span task increased. In contrast, the correlation between complex span and G*f* remained stable as the number of items in the complex span task increased. Also, the correlation between simple span and G*f* was equivalent to the correlation between complex span and G*f* for lists of four or more items. Unsworth and Engle therefore argued that controlled retrieval of items is needed when the number of items exceeds the scope of attention, that is, approximately four items. According to this perspective, simple span tasks with long lists require the same retrieval mechanism as complex span tasks because in each type of task, some information is lost from the scope of attention and must be recovered at the recall prompt. In the case of long-list simple span, some items are lost because the scope of attention is full and in the case of complex span, items are lost because attention is shifted to the processing component of the task.

Scope of Attention Tasks

Running-memory span and spatial simple span tasks with short lists, discussed earlier, might also be considered “scope of attention” tasks. Cowan (2001) reviewed evidence from a variety of tasks that prevent simple maintenance strategies such as rehearsal and chunking, and found that for most of these tasks, the number of items that could be maintained was about four. As mentioned above, other researchers have shown that, in some tasks, one item in the focus of attention has privileged access (Garavan, 1998; McElree, 2001; Nee & Jonides, 2008; Oberauer, 2002) but according to Cowan’s (2001) review, the *scope* of attention is approximately four items. While running span and spatial simple span may be considered part of this class, they are not ideal measures of the scope (and control) of attention because the to-be-remembered items must each be recalled and therefore performance is susceptible to output interference. In other words, it’s possible that more than four items are actively maintained but some representations are lost during recall.

For this reason, the visual-array comparison task (Luck & Vogel, 1997) is considered a better measure of the scope of attention. There are several variants of the visual-array comparison task, but in a typical version, the subject is briefly presented (e.g., 100 ms) with an array of several items that vary in shape and color. After a short retention interval (e.g., 1 s), the subject is then presented with another array and asked to judge whether the two arrays are the same or different. On half of the trials, the two arrays are the same and on the other half, one item in the second array is different. Thus, if all items in the initial array are maintained, then subjects will be able to detect the change. Most subjects achieve 100 percent accuracy on this task when the number of items is fewer than four, but performance begins to drop as the number of items in the array increases beyond four.

Tasks that are designed to measure the scope of attention, like visual-array comparison tasks, have not been used in studies of WM and G*f* as often as in complex and simple span tasks, but research shows that scope of attention tasks account for nearly as much variance in cognitive ability as complex span tasks (Awh et al., 2009; Cowan et al., 2005; Cowan et al., 2006).

Coordination and Transformation Tasks

All of the above-mentioned tasks require subjects to recall or recognize information that was explicitly presented. In some WM tasks, which we label “coordination and transformation" tasks, subjects are presented with information and required to manipulate and/or transform that information to arrive at a correct response. We include in this class backward span, letter-number sequencing, and alphabet recoding, as well as more complex tasks used by Kyllonen and Christal (1990) and Oberauer and colleagues (Oberauer, 2004; Oberauer et al., 2003; Süß et al., 2002).

Backward span tasks are similar to simple span tasks except that the subject is required to recall the items in reverse order. Thus, the internal representation of the list must be transformed for successful performance. In letter-number sequencing, the subject is presented with a sequence of letters and numbers and required to recall first the letters in alphabetical order and then the numbers in chronological order. In alphabet recoding, the subject is required to perform addition and subtraction using the alphabet, for example, C – 2 = A. The subject is presented with a problem and required to generate the answer. Difficulty is manipulated by varying the number of letters presented, such as CD – 2 = AB.

Kyllonen and Christal (1990) found very strong correlations between WMC and reasoning ability, using a variety of WM tasks that can all be considered in this "coordination and transformation" class . Also, Oberauer and colleagues (2003) showed that the correlation between WMC and G*f* does not depend upon whether WM is measured using complex span tasks or these types of transformation tasks, suggesting that coordination and transformation tasks tap the same mechanisms as complex span tasks, suggesting that the dual-task nature of complex span tasks (i.e., processing and storage) is not necessary for a WM task to be predictive of G*f*.

N-Back Tasks

In an n-back task, the subject is presented with a series of stimuli, one at a time, typically one every two to three seconds, and must determine if the current stimulus matches the one presented n-back. The stimuli may be verbal, such as letters or words, or visual objects, or spatial locations. N-back tasks have been used extensively in functional magnetic resonance imaging (fMRI) experiments, and more recently in WM training experiments. Gray, Chabris, and Braver (2003) showed that a verbal n-back task was a strong predictor of a matrix reasoning task (Raven's Advanced Progressive Matrices), making n-back a class of WM tasks to consider as we discuss the relationship between WMC and G*f*.

Empirical Evidence Linking WMC and Gf

Now that we have considered various measures of WMC, we turn to a review of the empirical evidence linking WMC and G*f*. As mentioned, two recent meta-analyses, conducted by two different groups of researchers, estimated the correlation between WMC and G*f* to be somewhere between $r = 0.72$ (Kane et al., 2005) and $r = 0.85$ (Oberauer et al., 2005). Kane and colleagues summarized the studies included in their meta-analysis in a table, which is reproduced here (see Table 21.1). Each of the studies included in the meta-analysis administered several tests of

Table 21.1 *Correlations between WMC and Gf/reasoning factors derived from confirmatory factor analyses of data from latent-variable studies with young adults*

Study	WMC tasks	G*f*/reasoning tasks	r (95% CI)
Kyllonen & Christal (1990) Study 2: *N* = 399	ABC numerical assignment, mental arithmetic, alphabet recoding	Arithmetic reasoning. AB grammatical reasoning, verbal analogies, arrow grammatical reasoning, number sets	0.91 (0.89, 0.93)
Study 3: *N* = 393	Alphabet recoding, ABC	Arithmetic reasoning, AB grammatical reasoning, ABCD arrow, diagramming relations, following instructions, letter sets, necessary arithmetic operations, nonsense syllogisms	0.79 (0.75, 0.82)
Study 4: *N* = 562	Alphabet recoding, mental math	Arithmetic reasoning, verbal analogies, number sets, 123 symbol reduction, three term series, calendar test	0.83 (0.80, 0.85)
Engle et al. (1999); *N* = 133	Operation span, reading span, counting span, ABCD, keeping track, secondary memory/ immediate free recall	Raven, Cattell culture fair	0.60 (0.48, 0.70)
Miyake et al. (2001); *N* = 167	Letter rotation, dot matrix	Tower of Hanoi, random generation, paper folding, space relations, cards, flags	0.64 (0.54, 0.72)
Ackerman, Beier, & Boyle (2002); *N* = 135	ABCD order, alpha span, backward digit span, computation span, figural-spatial span, spatial span, word-sentence span	Raven, number series, problem-solving, necessary facts, paper folding, spatial analogy, cube comparison	0.66 (0.55, 0.75)
Conway et al. (2002); *N* = 120	Operation span, reading span, counting span	Raven, Cattell culture fair	0.54 (0.40, 0.66)
Süß et al. (2002); *N* = 121[a]	Reading span, computation span, alpha span, backward digit span, math span, verbal span, spatial working memory, spatial short-term memory, updating numerical, updating spatial, spatial coordination, verbal coordination	Number sequences, letter sequences, computational reasoning, verbal analogies, fact/opinion, senseless inferences, syllogisms, figural analogies, Charkow, Bongard, figure assembly, surface development	0.86 (0.81, 0.90)
Hambrick (2003); *N* = 171)	Computation span, reading span	Raven, Cattell culture fair, abstraction, letter sets	0.71 (0.63, 0.78)
Mackintosh & Bennett (2003); *N* = 138b	Mental counters, reading span, spatial span	Raven, mental rotations	1.00

Table 21.1 (*cont.*)

Study	WMC tasks	G*f*/reasoning tasks	r (95% CI)
Colom et al. (2004) Study 1: *N* = 198	Mental counters, sentence verification, line formation	Raven, surface development	0.86 (0.82, 0.89)
Study 2: *N* = 203	Mental counters, sentence verification, line formation	Surface development, cards, figure classification	0.73 (0.82, 0.89)
Study 3; *N* = 193	Mental counters, sentence verification, line formation	Surface development, cards, figure classification	0.41 (0.29, 0.52)
Kane et al. (2004); *N* = 236)	Operation span, reading span, counting span, rotation span, symmetry span, navigation span	Raven, WASI matrix, BETA III matrix, reading comprehension, verbal analogies, inferences, nonsense syllogisms, remote associates, paper folding, surface development, form board, space relations, rotated blocks	0.67 (0.59, 0.73)

WMC = working memory capacity; G*f* = general fluid intelligence; 95% CI = the 95% confidence interval around the correlations; WASI = Wechsler Abbreviated Scale of Intelligence.

[a] *N* with the complete data set available (personal communication, K. Oberauer, July 7, 2004).

WMC and several tests of G*f*, and latent variable analysis was used to determine the strength of the relationship between the two constructs. A variety of WM tasks was used in these studies, including complex span, simple span, and coordination and transformation tasks. None of the studies referenced in Table 21.1 used tests designed to measure the scope of attention, such as visual-array comparison, or n-back tasks. One finding that has emerged from these studies is that complex span tasks are a stronger predictor of G*f* than is a simple span (Conway et al., 2002; Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Engle et al., 1999; Kane et al., 2004).

These recent findings have important implications for theories of the relationship between WMC and G*f*. However, it is imperative to emphasize that, in each of these cases – simple span with spatial stimuli, and simple span with long lists – the variance explained in G*f* is not entirely the same as the variance explained by complex span. To illustrate this, we reanalyzed data from Kane and colleagues (2004). We conducted a series of hierarchical regression analyses to determine the variance in G*f* that is either uniquely or commonly explained by complex span and simple span (cf. Chuah & Maybery, 1999). The results of this analysis are presented in Figure 21.2, panel (a). As the figure illustrates, simple span with spatial stimuli accounts for a substantial portion of variance in G*f*, and some of that variance is shared with complex span but some of it is unique to simple span with spatial stimuli. At first glance, this finding indicates that spatial simple span is tapping a mechanism that is important to G*f* but is not common to complex span. However, the battery of reasoning tasks used by Kane and colleagues to derive the G*f* factor had a slight bias toward spatial reasoning tests. When we model G*f* from only the verbal reasoning tests, we observe a different result (see Figure 21.2, panel (b)). This suggests that spatial simple span does *not* account for any domain-general variance in G*f* above and beyond complex span.

Unsworth and Engle (2006a) conducted a similar analysis with respect to the relationship between complex span, simple span with short and long lists, and G*f*. The results of their analysis are reproduced here in Figure 21.3. As with simple span with spatial stimuli, simple span with long lists (5–7 items) accounts for a substantial percentage of variance in G*f* (22.5%). However, most of that variance is shared with complex span (79%). This suggests that simple span with long lists and complex span tap similar mechanisms.

As mentioned, none of the studies in the meta-analyses conducted by Kane and colleagues (2005) included tasks specifically designed to measure the scope of

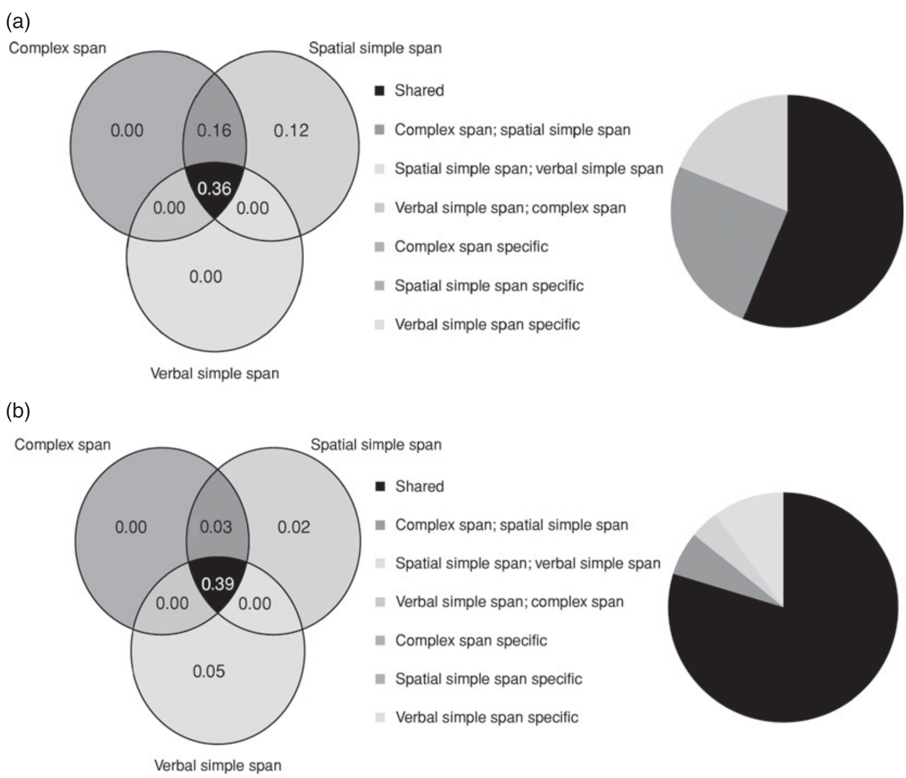

Figure 21.2 *Reanalysis of Kane et al. (2004). Reprinted with permission of the American Psychological Association.*
Panel (a): Complex span, spatial simple span, and verbal simple span predicting G*f* indexed by verbal reasoning, spatial reasoning, and figural matrix tasks. Panel (b): Complex span, spatial simple span and verbal simple span predicting verbal reasoning.

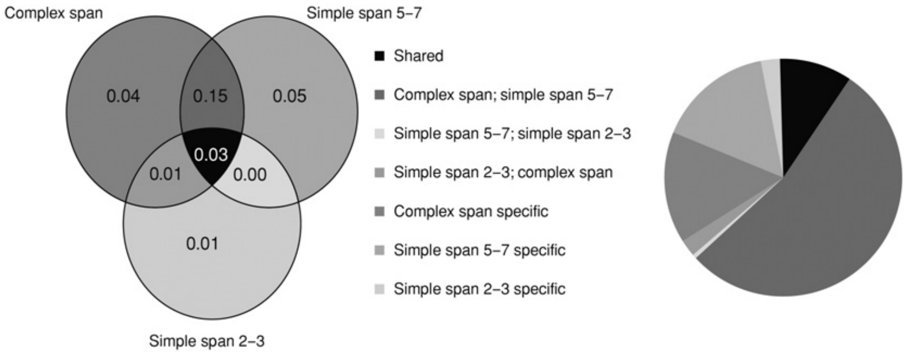

Figure 21.3 *Reanalysis of Unsworth and Engle (2006a). Reprinted with permission of Elsevier.*

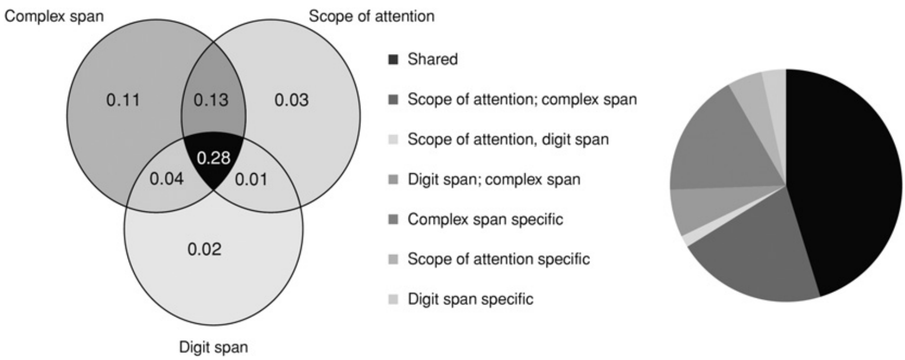

Figure 21.4 *Reanalysis of Cowan et al. (2005). Reprinted with permission of Elsevier.*

attention. However, Cowan and colleagues (2005) have conducted several recent studies to explore the relationship among scope of attention tasks, complex span, and cognitive ability in both children and adults. The results from just one of these studies are reproduced in Figure 21.4. Here we see that the variance in G*f* accounted for by scope of attention tasks is largely shared by complex span tasks but that complex span tasks account for variance in G*f* above and beyond scope of attention tasks. This result suggests that complex span and scope of attention tasks tap some overlapping mechanisms but complex span taps something that is important to G*f* that is not required by scope of attention tasks.

Finally, studies by Jeremy Gray and colleagues have considered the relationship among complex span, G*f*, and n-back. An important feature of Gray's n-back task is the inclusion of lure trials, which are trials in which the current stimulus matches a recently presented stimulus, but not the one n-back (e.g., n − 1 or n + 1 back). Accuracy to lure trials is lower than accuracy to non-lure foils, and accuracy to lure trials correlates more strongly with complex span tasks and with tests of G*f* than

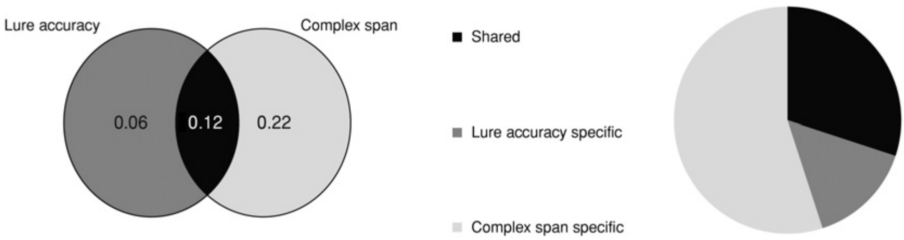

Figure 21.5 *Reanalysis of Burgess et al. (2011).*

accuracy to non-lure trials (Burgess et al., 2011; Gray et al., 2003; Kane et al., 2007). Burgess and colleagues examined the relationship between lure accuracy, complex span, and G*f*. The results of their analyses are reproduced in Figure 21.5. Here again, n-back and complex span account for much of the same variance in G*f*, but complex span accounts for a substantial portion of variance in G*f* that is not explained by n-back (see also Kane et al., 2007). As with the scope of attention tasks, this suggests that complex span and n-back tap some mechanisms that are common and important to G*f* but that they also tap some mechanisms that are unique and important to G*f*.

Theoretical Accounts of the Link between WM and Gf

Several theoretical accounts have been offered to account for the strong relationship between WMC and G*f*. It should be stated at the outset that these different accounts vary more in terms of emphasis and approach than they do in terms of the data they explain or the predictions they make. Furthermore, we believe that these various accounts can be encompassed by one theory, our multi-mechanism view, which we discuss in the section Process Overlap Theory: A Multi-Mechanism View.

Executive Attention

The first comprehensive theoretical account of the relationship between WMC and G*f* was offered by Engle and colleagues, and particularly in the work of Engle and Kane (Engle & Kane, 2004; Kane & Engle, 2002). This view has been referred to as the "controlled attention" or "executive attention" theory. According to this perspective, individuals with more effective cognitive control mechanisms, such as goal maintenance, selective attention, and interference resolution (inhibition), will perform better on a variety of tasks, including measures of WMC and tests of G*f*. There is a great deal of support for this theory, and an exhaustive review is not possible here. Instead, we will highlight a few important findings. First, performance on various WM tasks has been linked to mechanisms of cognitive control, such as inhibition. For example, individuals who perform better on complex span tasks do so in part because they are better at resolving proactive interference from previous trials (Bunting, 2006; Unsworth & Engle, 2007). Similarly, individuals who perform better on complex span tasks are also more accurate on lure trials in the n-back task and lure trials predict G*f* better than non-lure trials (Burgess et al., 2011; Gray et al., 2003; Kane et al., 2007). Also, tasks that place heavy demands on cognitive control but little demand on memory predict G*f* (Dempster & Corkill, 1999).

Perhaps most striking, the correlation between complex span and G*f* increases as a function of the amount of proactive interference (PI) in the task (Bunting, 2006). Bunting had subjects perform a complex span task and manipulated the category from which the to-be-remembered items were drawn (words or digits). The category was repeated for three items (to build PI) and then switched on the fourth item (to release PI). The correlation between complex span and Raven's Progressive Matrices, a marker of G*f*, increased linearly as PI increased and dropped significantly when PI was released.

While executive attention theory has enjoyed considerable support, a fair criticism is that the empirical evidence is overly reliant on studies using complex span tasks. This is problematic because complex span tasks are, as the name suggests, complex. Thus, while Engle and colleagues have argued that "executive attention" is the primary source of variation in these tasks, other researchers have emphasized the fact that other sources of variance are at play as well, such as domain-specific abilities required to perform the processing component of the task (e.g., mathematical ability in the case of operation span, or verbal ability in the case of reading span; Bayliss, Jarrold, Gunn, & Baddeley, 2003; Daneman & Carpenter, 1983; Shah & Miyake, 1996). Also, performance of complex span tasks can be influenced by strategy deployment, such that a person may perform above average on a complex span task because they implement an effective strategy, not because the person actually has superior WMC (Dunlosky & Kane, 2007; McNamara & Scott, 2001; Turley-Ames & Whitfield, 2003).

Scope and Control of Attention

According to Cowan's approach, the scope of attention is limited to about four items, and individual differences in the scope and control of attention are what drive the correlation between measures of WMC and G*f* (for a similar perspective on capacity limitations, see Drew & Vogel, 2009). The difference between Cowan's approach and that of Engle and colleagues, however, may be just one of emphasis. Cowan's recent work has emphasized the scope of attention while Engle's recent work, particularly that of Unsworth and Engle, has emphasized retrieval of information that has been lost from the focus of attention. Thus, we do not see these views as necessarily incompatible and we incorporate both into our multi-mechanism view, articulated in the section Process Overlap Theory: A Multi-Mechanism View. One issue of debate, however, is whether scope of attention tests of WMC, like visual-array comparison, account for the same variance in G*f* as complex span tasks. The results of Cowan and colleagues (2005), reproduced here in Figure 21.4, suggest that complex span tasks have something in common with G*f* that scope of attention tasks do not. However, Cowan and colleagues reported confirmatory factor analyses indicating that a two-factor model of the WM tasks, dissociating scope of attention and complex span, did *not* fit the data better than a single-factor model. Also, more recent work has demonstrated correlations between scope of attention tasks and G*f* that are as strong as correlations typically observed between complex span tasks and G*f* (Awh et al., 2009; Cowan et al., 2006). More research is needed to further investigate the relationship among scope of attention tasks, complex span tasks, and G*f*.

Binding Limits

Oberauer and colleagues characterize the relationship between WMC and G*f* as one of "binding limits" rather than one of attention (Oberauer et al., 2012). Oberauer argues that memory requires the binding of features into objects and the binding of objects into episodes. There is a limit to the number of bindings that can be actively maintained at once and this causes WMC. Importantly, more complex tasks require more bindings, and Oberauer has shown that more complex WM tasks tend to show stronger correlations with tests of G*f*, which themselves are complex tasks. Of particular importance is the finding, mentioned in the section Coordination and Transformation Tasks, that WM tasks that require multiple bindings, such as coordination and transformation tasks, predict G*f* just as well as do complex span tasks, and account for largely the same variance in G*f* as complex span tasks (Oberauer et al., 2003; Süß et al., 2002). This suggests that the dual-task nature of complex span tasks is not necessary to predict G*f* and calls into question a basic tenet of executive attention theory, that is, that cognitive control mechanisms are responsible for the relationship between WMC and G*f*. That said, an unresolved issue is the relationship between attention and binding. Hence, it isn't clear if Oberauer's view is incompatible with Engle and/or Cowan's view.

Active Maintenance and Controlled Retrieval

Unsworth and Engle (2007) argue that there are two dissociable domain-general mechanisms that influence WMC: (1) a dynamic attention component that is responsible for maintaining information in an accessible state; and (2) a probabilistic cue-dependent search component, which is responsible for searching for information that has been lost from the focus of attention. For example, as a subject performs a complex span task, the dynamic attention component is necessary to coordinate the processing and storage demands of the task and to maintain the to-be-remembered items in an accessible state. The search component is necessary at the recall prompt to recover to-be-remembered items that may have been lost from the focus of attention because of the demands of the processing component of the task.

Empirical support for this theory comes from simple span tasks with long lists and from serial free recall tasks designed to assess primacy and recency effects. As mentioned, Unsworth and Engle (2006a, 2007) have shown that simple span tasks with long lists correlate as well with G*f* as measures of complex span tasks and much of the variance explained by simple span with long lists is shared with complex span (see Figure 21.4). They argue that simple span with long lists taps the same controlled retrieval mechanism as complex span because the focus of attention is overloaded and items displaced from the focus of attention must be recovered during recall. More recent work demonstrates that individual differences in the primacy portion of free recall account for different variance in G*f* than individual differences in the recency portion (Unsworth, Spillers, & Brewer, 2010). Unsworth and colleagues (2010) argue that variance in the primacy effect is driven by individual differences in controlled retrieval, and variance in the recency effect is driven by individual differences in active maintenance via attention.

While Unsworth and Engle (2007) do not provide a neural model of their theory, the dynamic attentional processes implicated in their account are consistent with recent computational models of WM that implicate PFC, ACC, and parietal cortex as regions involved in the active maintenance, updating, and monitoring of information in WM (Botvinick et al., 2001; Frank, Loughry, & O'Reilly, 2001; E. K. Miller & Cohen, 2001; O'Reilly & Frank, 2006). Indeed, neuroimaging studies of complex span tasks show that PFC, ACC, and parietal areas are more strongly recruited in complex span tasks than during simple span tasks (Bunge et al., 2000; Chein et al., 2011; Kondo et al., 2004; Osaka et al., 2003; Osaka et al., 2004; Smith et al., 2001).

Unsworth and Engle (2007) further speculate that the medial temporal lobes (MTL) are also important for WM performance, which is a relatively novel prediction (but see Ranganath, 2006). In particular, they argue that the cue-dependent search process implicated during recall relies on coordinated activity between PFC and MTL. This view is also consistent with computational models that examine the interaction between PFC and MTL in a variety of memory tasks (O'Reilly & Norman, 2002). Indeed, a recent fMRI study indicates greater PFC and hippocampal activity during recall in complex span tasks than during recall in simple span tasks (Chein et al., 2011).

Process Overlap Theory: A Multi-Mechanism View

We argue that there are multiple domain-general cognitive mechanisms underlying the relationship between WMC and G*f*. Our view is shaped by Unsworth and Engle's account discussed in the section Active Maintenance and Controlled Retrieval, but also by computational models and neuroimaging data that similarly fractionate WM into dissociable mechanisms. Most important among these are the scope and control of attention, updating and conflict monitoring, interference resolution, and controlled retrieval. These mechanisms have been linked to neural activity in specific brain regions: PFC-parietal connections for the scope and control of attention (Todd & Marois, 2004; Vogel & Machizawa, 2004); a PFC-ACC-basal ganglia-thalamus network for updating and conflict monitoring (Ashby et al., 2005; Botvinick, 2007; O'Reilly & Frank, 2006); inferior frontal cortex for interference resolution (Aron, Robbins, & Poldrack, 2004); and PFC-hippocampal connections for controlled retrieval (Chein et al., 2011; Nee & Jonides, 2008; Ranganath, 2006).

This multi-mechanism view of the relationship between WMC and G*f* is consistent with process overlap theory, a recent account of the general factor of intelligence

(Kovacs & Conway, 2016). The primary aim of the theory is to explain the finding that cognitive ability tests with diverse content all correlate positively. This finding, called the positive manifold, is the basis of the general factor, *g*, that explains 40–50 percent of the entire variance in IQ tests.

The multi-mechanism view in general and the idea of overlapping processes determining mental test performance in particular is not new. In fact, it dates back to one of the earliest criticisms of Spearman's *g* (Thompson, 1916). Spearman described the underlying source of variance in *g* as a unitary construct, reflecting some sort of cognitive resource, or "mental energy." However, Thomson demonstrated that the positive manifold could be caused by multiple processes as long as a battery of tests tap these various processes in an overlapping fashion. This is the basis of so-called sampling theories (Thomson, 1916; Thorndike, 1927).

Thomson (1916) provided a mathematical proof of this, showing that the correlation between any two tests can be described as the function of the ratio of processes in common, that is, the number of processes sampled by both tests relative to the total number of processes sampled by each. Thus, *g* may not reflect a unitary construct; instead, it may emerge from a battery of tasks that sample overlapping domain-general mechanisms. It has since been reinforced with more elaborate mathematical methods that it is impossible to select between Spearman's and Thomson's explanation on a purely statistical basis (Bartholomew, Deary, & Lawn, 2009).

Besides subscribing to a multi-process, sampling approach to intelligence, process overlap theory also draws heavily on the concept of working memory capacity in explaining the positive manifold in intelligence. The theory postulates an overlap of cognitive processes activated by various mental ability tests and working memory tasks. In particular, it is hypothesized that any item or task requires a number of domain-specific as well as domain-general cognitive processes. Domain-general processes responsible for executive attention and cognitive control are central to performance on both mental tests and working memory tasks since they are activated by a large number of items, alongside domain-specific processes tapped by specific types of items/tests only.

The theory actually focuses on limitations. That is, the central processes that are tapped by a large numbers of tasks limit performance in a general way and make errors more likely regardless of the domain-specific processes that are also tapped by the same tasks. This way, executive processes function as a bottleneck and can potentially mask individual differences in specific processes. Hence process overlap theory, contrary to traditional models of sampling, proposes a nonadditive interaction of processes: Instead of simply adding scores on sampled processes, the mathematical model behind process overlap theory proposes that each individual dimension of a task has to be completed in order for someone to arrive at a correct solution. A single process can cancel the effect of all other processes and be the cause of error on its own.

Importantly, process overlap theory provides an explanation of the general factor of working memory capacity as well as *g*. It proposes that the same pool of domain-general executive resources is tapped by different working memory tasks as different psychometric tests of cognitive ability, especially the ones that measure fluid reasoning. According to the theory, that is why the general factors of working memory and fluid intelligence correlate so strongly.

Conclusions

Working memory has emerged as a very useful construct in the field of psychology. Various measures of WMC have been shown to correlate quite strongly with measures of intelligence, accounting for at least half the variance in G*f*. We argue that these correlations exist because tests of WMC and tests of G*f* tap multiple domain-general cognitive mechanisms required for the active maintenance and rapid controlled retrieval of information. This argument is more formally expressed in a framework we refer to as process overlap theory (Kovacs & Conway, 2016).

More research is also needed to better specify the various mechanisms underlying performance of WM and reasoning tests. Neuroimaging studies on healthy adults and neuropsychological tests of patients with various neurological damage or disease will be especially fruitful. For example, fMRI studies have illustrated that individual differences in activity in PFC during a WM task partly account for the relationship between WMC and G*f* (Burgess et al., 2011; Gray et al., 2003). One intriguing possibility is that individual differences in activity in different brain regions (or network of regions) account for *different* variance in G*f*. For example, based on the work of Unsworth and Engle (2007), it may be possible to demonstrate that individual differences in activity in the PFC, ACC, and parietal cortex, reflecting active maintenance during a WM task, account for different variance in G*f* rather than individual differences in activity in PFC and hippocampus, reflecting controlled retrieval during a WM task.

The multi-mechanism view also has implications for research on WM training and for cognitive therapy for the elderly and patients with neural damage or disease. That is, rather than treat WM as a global construct, training and remediation could be tailored more specifically. Instead of "WM training" we envisage mechanism-specific training. That is, training a specific domain-general cognitive mechanism should result in improved performance across a variety of tasks. There is some research supporting this idea (Dahlin et al., 2009; Karbach & Kray, 2009) but again, more work is needed to confirm the reliability and durability of these results.

In sum, WMC is strongly correlated with G*f*. We argue that the relationship between these constructs is driven by the operation of multiple domain-general cognitive mechanisms that are required for the performance of tasks designed to measure WMC and for the performance of test batteries designed to assess fluid intelligence, consistent with process overlap theory (Kovacs & Conway, 2016). Future research in cognitive psychology and neuroscience will hopefully refine our understanding of these underlying mechanisms, which will in turn sharpen the multi-mechanism view.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, *117*, 288–318.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, *131*, 567–589.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences*, *8*, 170–177.

Ashby, F. G., Ell, S. W., Valentin, V. V., & Casale, M. B. (2005). FROST: A distributed neurocomputational model of working memory maintenance. *Journal of Cognitive Neuroscience*, *17*, 1728–1743.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (vol. 2, pp. 89–195). New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *225*, 82–90.

Awh, E., Fukuda, K., Vogel, E. K., & Mayr, U. (2009). Quantity not quality: The relationship between fluid intelligence and working memory capacity. Paper presented at the fiftieth annual meeting of the Psychonomic Society, Boston, November.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (vol. 8, pp. 47–89). New York: Academic Press.

Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review*, *116*(3), 567–579. http://doi.org/10.1037/a0016262

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General*, *132*, 71–92.

Binet, A. (1903). *Etude expérimentale de l'intelligence* (The experimental study of intelligence). Paris: Schlecher.

Botvinick, M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective and Behavioral Neuroscience*, *7*, 356–366.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.

Bunge, S. A., Klingberg, T., Jacobsen, R. B., & Gabrieli, J. D. E. (2000). A resource model of the neural basis of executive working memory. *Proceedings of the National Academy of Sciences*, *97*, 3573–3578.

Bunting, M. F. (2006). Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 183–196.

Burgess, G. C., Braver, T. S., Conway, A. R. A., & Gray, J. R. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology: General*, *140*, 674–692.

Case, R., Kurland, M. D., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386–404.

Chein, J. M., Moore, A. B., & Conway, A. R. A. (2011). Domain-general mechanisms of active maintenance and serial recall in complex working memory span. *Neuroimage*, *54*, 550–559.

Chuah, Y. M. L., & Maybery, M. T. (1999). Verbal and spatial short-term memory: Common sources of developmental change? *Journal of Experimental Child Psychology*, *73*, 7–44.

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, *32*, 277–296.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D., & Minkoff, S. (2002). A latent variable analysis of working memory capacity, short term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, *12*(5), 769–786.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, *104*, 163–191.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Cowan, N. (2005). *Working memory capacity*. Hove, UK: Psychology Press.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100.

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory and Cognition*, *34*, 1754–1768.

Crowder, R. G. (1982). The demise of short-term memory. *Acta Psychologica*, *50*, 291–323.

Dahlin, E., Bäckman, L., Neely, A. S., & Nyberg, L. (2009). Training of the executive component of working memory: Subcortical areas mediate transfer effects. *Restorative Neurology and Neuroscience*, *27*(5), 405–419.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Behavior and Verbal Learning*, *19*, 450–466.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 561–584.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, *3*, 422–433.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review*, *112*, 3–42.

Dempster, F. N., & Corkill, A. J. (1999). Interference and inhibition in cognition and behavior: Unifying themes for educational psychology. *Educational Psychology Review*, *11*, 1–88.

Drew, T., & Vogel, E. K. (2009). Working memory capacity limitations. In L. R. Squire (Ed.) *Encyclopedia of neuroscience* (vol. 10, pp. 523–531). Amsterdam: Elsevier.

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy-assessment methods. *Quarterly Journal of Experimental Psychology*, *60*, 1227–1245.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 145–199). New York: Academic Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211–245.

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience*, *1*, 137–160.

Garavan, H. (1998). Serial attention within working memory. *Memory and Cognition*, *26*, 263–276.

Gignac, G. E. (2014). Fluid intelligence shares closer to 60% of its variance with working memory capacity and is a better indicator of general intelligence. *Intelligence*, *47*, 122–133. http://doi.org/10.1016/j.intell.2014.09.004

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322.

Hambrick, D. Z. (2003). Why are some people more knowledgeable than others? A longitudinal study of real-world knowledge acquisition. *Memory & Cognition*, *31*, 902–917.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: *An investigation of the knowledge-is-power hypothesis. Cognitive Psychology*, *44*, 339–387.

Hambrick, D. Z., & Meinz, E. J. (2011). Limits on the predictive power of domain-specific experience and knowledge in skilled performance. *Current Directions in Psychological Science*, *20*(5), 275–279. http://doi.org/10.1177/0963721411422061

Hambrick, D. Z., & Oswald, F. L. (2005). Does domain knowledge moderate involvement of working memory capacity in higher-level cognition? A test of three models. *Journal of Memory and Language*, *52*, 377–397.

Hebb, D. O. (1949). *Organization of behavior*. New York: Wiley.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, *59*, 193–224.

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the *n*-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 615–622.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review*, *9*, 637–671.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66–71.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*, 189–217.

Karbach, J., & Kray, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science*, *12*(6), 976–990.

Kondo, H., Morishita, M., Osaka, N., Osaka, M., Fukuyama, H., & Shibasaki, H. (2004). Functional roles of the cingulo-frontal network in performance on working memory. *Neuroimage*, *21*, 2–14.

Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, *27*(3). http://doi.org/10.1080/1047840X.2016.1153946

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*, 389–433.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

Mackintosh, N. J., & Bennett, E. S. (2003). The fractionation of working memory maps onto different components of intelligence. *Intelligence*, *31*, 519–531.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37(1), 1–10. http://doi.org/10.1016/j.intell.2008.08.004

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 817–835.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory and Cognition*, *29*, 10–17.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General*, *130*, 621–640.

Mukunda, K. V., & Hall, V. C. (1992). Does performance on memory for order correlate with performance on standardized measures of ability? A meta-analysis. *Intelligence*, *16*, 81–97.

Nee, D. E., & Jonides, J. (2008). Neural correlates of access to short-term memory. *Proceedings of the National Academy of Sciences*, *105*, 14228–14233.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition 2002*, *28*, 411–421.

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin and Review*, *19*(5), 779–819. http://doi.org/10.3758/s13423-012-0272-4

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. M. (2005). Working memory and intelligence – their correlation and their relation: A comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 61–65.

Oberauer, K., Süß, H. M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*, 167–193.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*, 283–328.

O'Reilly, R. C., & Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, *6*(12), 505–510.

Osaka, M., Osaka, N., Kondo, H., Morishita, M., Fukuyama, H., Aso, T., et al. (2003). The neural basis of individual differences in working memory capacity: An fMRI study. *Neuroimage*, *18*, 789–797.

Osaka, N., Osaka, M., Kondo, H., Morishita, M., Fukuyama, H., & Shibasaki, H. (2004). The neural basis of executive function in working memory: An fMRI study based on individual differences. *Neuroimage*, *21*, 623–631.

Pollack, I., Johnson, I. B., & Knaff, P. R. (1959). Running memory span. *Journal of Experimental Psychology*, *57*, 137–146.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*, 23–38.

Ranganath, C. (2006). Working memory for visual objects: Complementary roles of inferior temporal, medial temporal, and prefrontal cortex. *Neuroscience*, *139*(1), 277–289.

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*, 893–912.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*, 4–27.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology*, *22*, 261–273.

Smith, E. E., Geva, A., Jonides, J., Miller, A., Reuter-Lorenz, P., & Koeppe, R. A. (2001). The neural basis of task-switching in working memory: Effects of performance and aging. *Proceedings of the National Academy of Sciences*, *98*, 2095–2100.

Süß, H. M., Oberauer, K., Wittman, W. W., Wilhelm, O., & Schulze, R. (2002). Working memory capacity explains reasoning ability – and a little bit more. *Intelligence*, *30*, 261–288.

Thompson, G. (1916). A hierarchy without a general factor. *British Journal of Psychology*, *8*, 271–281.

Thorndike, E. L. (1927). *The measurement of intelligence*. New York: Teachers College, Columbia University.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*, 751–754.

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language*, *49*, 446–468.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Unsworth, N., & Engle, R. W. (2006a). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, *54*, 68–80.

Unsworth, N., & Engle, R. W. (2006b). A temporal-contextual retrieval account of complex span: An analysis of errors. *Journal of Memory and Language*, *54*, 346–362.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104–132.

Unsworth, N., Spillers, G. J., & Brewer, A. (2010). The contributions of primary and secondary memory to working memory capacity: An individual differences analysis of immediate free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 240–247.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*, 784–775.

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain*, *92*, 885–896.

22 Intelligence and Reasoning

Joni M. Lakin and Harrison J. Kell

To define reasoning often means to emphasize what it is *not*, and then to find its definition in the gaps. In this chapter, we will begin with an attempt to affirmatively define reasoning through an ontological definition drawn from the philosophy of science. We will then give a more epistemological definition based on the psychometric and cognitive psychological literature, which defines reasoning via the discovery of interindividual variability in task performance and by manipulating item characteristics to see their effects on reasoning load and errors. Finally, we will attempt to disambiguate the concept by saying what it is *not*, including that it is not the same as working memory, critical thinking, or processing speed.

Defining Reasoning

Reasoning refers to the process of drawing conclusions or inferences from information. Bruner's (1957) helpful definition of reasoning is that it means to go "beyond the information given." A common distinction in analyzing reasoning processes is between inductive and deductive reasoning. In logic, an inference is called *inductive* if the truth of the premises makes the conclusion probable but not certain. An inference is called *deductive* if the truth of the initial information (or premises) guarantees the truth of the conclusion. Distinctions between deductive and inductive reasoning can be important in understanding logic, but in practice, these distinctions may exist more in the mind of the researcher developing a task than in the performance of examinees on that task. Researchers have found that performance on deductive and inductive tests are strongly related and distinguishing between these facets may not be critical to the definition of reasoning (Wilhelm, 2005).

These caveats aside, it is helpful at the outset to consider a more nuanced definition of these two aspects of reasoning. Individuals attempt to *infer* (either automatically or deliberately) concepts, patterns, or rules that best (i.e., most uniquely or exhaustively) characterize the relationships or patterns they perceive among all the elements (e.g., words, symbols, figures, sounds, movements) in a stimulus set. Better reasoning is characterized by the use of concepts or rules that simultaneously satisfy the opposing needs for abstraction (or generalization) and specificity. Such concepts or rules tend to be at least moderately abstract yet precisely tuned. Put differently, a poor inference is often vague and captures only a subset of the relationships among the elements in the set. Inference becomes more challenging when the contrast in ideas becomes greater, more complex, or more abstract.

Individuals use deduction when a rule, set of premises, or statements using warrants (i.e., claims) are tested for logical consistency with the information that is either given in the problem or assumed to be true within the community of discourse. Deduction maintains semantic information (confirming information) while induction increases semantic information (expanding what we know; Johnson-Laird & Khemlani, 2013). Humans often seem to reason by creating and manipulating mental models of the situation that apply inductive or deductive logic and check the consistency of the logic with the various models (Johnson-Laird & Byrne, 1991). For example, Johnson-Laird and Khemlani (2013) offer the premise of "A triangle is on the right of a circle" (p. 6), which has a single mental model that captures the rule [●▲]. The difficulty of the problem is increased when more plausible mental models can fit a given set of premises (e.g., "a square and a triangle are on the right of a circle"; two models [●▲■, ●■▲]). Better reasoning involves providing warrants that are more plausible or consistent with the rules of logic or the conditions embodied in a comprehensive mental model. More advanced deductive reasoning involves providing either multiple (possibly divergent) warrants for a single claim or requiring an increasingly sophisticated chain of logically connected and separately warranted assertions.

Increasing the complexity of reasoning demands in a problem can lead to failures of logical or accurate reasoning. The focus of psychometric research on reasoning is to explore these failures and to identify individuals who are able to maintain logical solutions while the reasoning problems increase in difficulty. Differentiating levels of skill leads to the interindividual variability in performance and response time that supports empirical models of reasoning. Cognitive psychological approaches focus more on the steps or process of reasoning (including response time) and generating hypotheses related to specific errors in reasoning.

Reasoning skills are not theorized to be innately born or narrowly defined. The consensus of the field is that reasoning is a developed skill that is more general than specific with respect to formal learning (e.g., when it is developed), the content or context of reasoning (e.g., what is reasoned about, including test tasks), and which regions of the brain are activated in neurological research. Given that it is a broad set of skills and recruits many neural processes, there is some debate over how its subfacets should be defined. It may also be isomorphic with fluid reasoning (*Gf*) and even general intelligence (*g*). These claims are explored in more depth in the following sections.

Psychometric Definitions of Reasoning

The most widely cited psychometric model of reasoning comes from Carroll's (1993) seminal factor-analytic work, where he analyzed over a century of psychological research to form an empirical model of human abilities. This model is usually referred to as the Cattell-Horn-Carroll (CHC) model of human ability and is the most prominent model of individual differences in human cognitive abilities (Schneider & McGrew, 2012).[1] It is derived from the covariation of measures of cognitive skills. At the apex is general intelligence, with six or more broad abilities (including crystallized intelligence and fluid intelligence/reasoning), and finally a large number of narrow factors representing narrow skills or abilities.

The first important implication of the CHC model is that human abilities are organized hierarchically. This means that some cognitive competencies are more broadly useful than others. It also means that theories that postulate an independent set of abilities (Gardner, 1983; Thurstone, 1938) or only one ability of any consequence (Jensen, 1998) are fundamentally flawed because they do not account for positive manifold (small to medium correlations between almost all ability measures) as well as the divergence of broad abilities (i.e., less than perfect correlations).

The hierarchy that Carroll proposes starts with *g* (general mental ability) at the topmost level. Although the broadest factor in the model, *g* is also the least psychologically transparent. Eight broad group factors that are somewhat more psychologically transparent define the second level. These factors vary in their closeness or association with *g*. The closest is an ability factor that Cattell (1963) called G*f* (general fluid ability). Other broad factors closely related to *g* at this level include G*c* (general verbal crystallized ability), G*v* (general spatial visualization ability), and G*m* (general memory ability). Finally, a longer list of primary factors that are even more psychologically transparent defines the third level. These factors include such abilities as verbal comprehension, verbal fluency, inductive reasoning, spatial visualization, perceptual speed, and number facility. Most of these specific abilities have quite narrow demands and predictive ranges. For example, the classic "cancellation" task asks examinees to find and mark certain letters, numbers, or shapes (i.e., "cross out all of the Ps on this page"). It has narrow utility to predict real-world outcomes, but when averaged with other similar tasks, provides a measure of a broader processing speed ability.

The second critical finding is that the topmost factor in the hierarchy (*g*) is virtually synonymous (i.e., highly correlated) with the factor called G*f* (general fluid ability) at the second level. And G*f* is in turn highly correlated with the primary factor called inductive reasoning (IR). Fluid reasoning (G*f*) was originally hypothesized by Cattell (1943) to contrast more process-oriented abilities needed for learning and to solve novel problems, which he termed fluid intelligence, from the more content-focused abilities associated with maturation and education, that is, those skills and pieces of knowledge derived from experience and forming adult expertise, which he termed crystallized intelligence (or G*c*).

In exploring the facets of *g* and, in turn, G*f,* Carroll found that the most powerful influence on *g* came from the tests associated with G*f* and the strongest determinants of the G*f* factor were the abstract-reasoning tasks, including figural matrices, which are classified as inductive reasoning (IR) tasks. Gustafsson (1988; Kvist & Gustafsson, 2008) claimed that the three factors are in fact identical (i.e., *g* = G*f* =

[1] Factor-analytic approaches have dominated the past forty years of reasoning research (if not the entire history of the field). However, attempts to approach the field from other perspectives are gaining ground and are discussed in the chapter.

IR). Others would describe the relationship between *g* and G*f* as more of an approximation than an identity (Carroll, 1993; Horn & Blankson, 2005) or as just highly correlated (Süß & Beauducel, 2005). In any case, we are left with the important insight that reasoning abilities are at the core of human cognitive competence. In other words, the least psychologically transparent dimension (*g*) is in large measure isomorphic with one of the most psychologically transparent dimensions (e.g., IR).

Finally, a third critical finding in the CHC literature on human abilities is that the general reasoning factor (G*f*) may be decomposed into subfactors: (1) deductive reasoning (termed sequential reasoning by Carroll and largely comprising verbal tasks), (2) quantitative reasoning (inductive or deductive reasoning with quantitative concepts), and (3) inductive reasoning (often measured with figural tasks; Carroll, 1993). A good reasoning test, then, should probably measure all three of these reasoning factors – or at least not be strongly biased toward one (Wilhelm, 2005). This fact is commonly overlooked in studies that represent fluid reasoning abilities with a single figural reasoning test, such as Raven's Progressive Matrices test (Raven, Court, & Raven, 1977).

It is this third finding of Carroll's that is most contentious in understanding reasoning abilities. In his comprehensive review of this literature, Wilhelm (2005) argued that the study of reasoning must go further in understanding the degree to which characteristics of tasks in tests influence the reasoning construct measured. Most importantly, Wilhelm argued that the *content* of reasoning may be more critical to defining facets of reasoning (such as verbal vs. figural reasoning) than the specific processes used (such as inductive vs. deductive reasoning). In his small study, he found that inductive and deductive factors correlated almost perfectly, while factors reflecting verbal, quantitative, and figural content fit the data better than a model that assumes either a single factor or two-process factors.

Wilhelm's work aligns with a competing model of intelligence that seeks to unseat the primacy of the CHC model: the Berlin Intelligence Structure (BIS; Beauducel & Kersting, 2002). In defining and measuring general intelligence, the BIS places reasoning as one of four types of processes defining an operations facet, alongside creativity, memory, and processing speed. Test content (figural, verbal, and numerical) then define a content facet, yielding a 3 × 4 grid of tasks. Thus, the BIS gives content, not processes, the primary role of defining reasoning skills (Beauducel, Brocke, & Liepmann, 2001; Wilhelm, 2005). The content facet has shown consistency across age in several studies of the reasoning dimension (Beauducel et al., 2001; Beauducel & Kersting, 2002). The model is agnostic as to whether the operation-content cells form abilities that reflect a biological reality or distinct neural processes (Beauducel & Kersting, 2002; Süß & Beauducel, 2005).

Like the CHC model, the BIS model was developed using factor analyses, but began with the review and sorting of the large number of tasks in test found in the literature (Beauducel & Kersting, 2002). Much of this research uncovers a content factor as expected (Beauducel & Kersting, 2002), but because the operations include reasoning as a unitary operation, they do not directly test whether G*f* facets from Carroll's (1993) work (including deductive and inductive reasoning) would also be identified in a broad enough sampling of reasoning tasks. As with work on the CHC model, the BIS research often fails to clearly uncover the expected number of operations. Overlap of the psychometric tasks may be to blame, as in Süß and Beauducel (2005), where numerical speed tests correlated unexpectedly well with verbal and figural reasoning tasks, suggesting unhypothesized overlapping of the cognitive abilities measured by each. This finding leads us to wonder if the fault lies in the model or the tasks that have been developed. Given the substantial overlap of content and reasoning tasks even in Carroll's model (e.g., deductive reasoning being almost exclusively defined by verbal tasks), it seems difficult to disentangle content and processes within reasoning or to develop psychometric tasks that attempt to sample each combination of content and process.

One important implication of the BIS model is that measuring reasoning requires tasks that sample across the main content areas. This implication is consistent with research on the *g*-factor, which argues that a "sufficiently diverse" sample of tasks will yield the "same" *g* because construct-irrelevant variance can be factored out (Floyd, Bergeron, McCormack, Anderson, & Hargrove-Owens, 2005). The BIS model suggests that the diversity of content required to measure G*f* consistently will require going beyond the figural tasks (e.g., Raven's Progressive Matrices) so often used as single measures of reasoning (Beauducel et al., 2001). Süß and Beauducel (2005) go so far as to suggest "it is more important to use several independent measures with limited reliability than only one task with strong reliability" (p. 329), if it achieves wider sampling of content.

A second model of cognitive abilities that currently competes with CHC is *g*-verbal-perceptual-image rotation (*g*-VPR; Johnson & Bouchard, 2005a, 2005b), which also emphasizes the role of content in the reasoning process. Based on Vernon's (1950) model, which centered on a general factor and two group factors (verbal-educational & spatial-practical-mechanical), the *g*-VPR model features four ability strata; its defining characteristics are a general ability factor constituting the fourth stratum and verbal, perceptual, and image-rotation factors constituting the third stratum. The verbal factor represents the ability to solve problems with verbal content, the perceptual factor represents the ability to solve problems featuring static images, and the image-rotation factor represents the ability to solve problems by mentally manipulating visual images (Hunt, 2010). *g*-VPR does contain mathematical factor(s), but they exist at the second stratum and vary from study to study in the extent to which they are associated with the verbal or perceptual factors, or both.

The role of reasoning in *g*-VPR is complex. *g*-VPR does not emphasize a factor dedicated explicitly to reasoning because it is content-focused and because Vernon (1965), the originator of the approach *g*-VPR is based on, usually equated reasoning with *g*. Nonetheless, when tests labeled "inductive reasoning" and "abstract reasoning" have been analyzed in the course of *g*-VPR research, their scores have loaded directly or indirectly on the perceptual factor (Johnson & Bouchard, 2005b; Johnson, te Nijenhuis, & Bouchard, 2007; Major, Johnson, & Deary, 2012). Further complicating matters is the fact that the *g*-VPR approach is agnostic as to whether the *g*-factor is reflective (i.e., derives from a latent construct with biological bases) or formative (i.e., is simply an average of lower-order abilities–Johnson & Deary, 2011; Major et al., 2012). If treated as a formative construct, the *g*-factor would not reflect a fundamental reasoning capacity that influences performance on many tests, but instead is simply a summary of performance on those tests that arise from many more narrow reasoning abilities.

The Cognitive-Psychological Approach to Reasoning

Researchers following the *cognitive-psychological approach* to the study of reasoning typically study the responses of a small number of participants to logical tasks such as syllogisms or formal logic tasks. Researchers analyze how features of the problem influence the types of errors that participants make and often base their generalizations on the proportion of participants making certain errors (Stanovich, 1999). One source of debate in the cognitive approach is whether humans are fundamentally rational, as Aristotle assumed, or whether consistent demonstrations of irrational behaviors in the laboratory mean that humans function with pervasive biases that impede or prevent rational decision-making. Researchers who conclude that humans operate with biases cite instances showing that people are swayed by personal testimony that is contrary to data and readily accept believable conclusions that are based on unlikely premises. However, critics of this research argue that the abstract structure of the problems can influence how they are solved and participants' misunderstandings of the format may explain some of these apparent failures in logical reasoning (Leighton, 2004). In some cases, illogical behavior on artificial tasks can disappear when the task is framed in a more meaningful way (Evans & Feeney, 2004; Stenning & Monaghan, 2004).

Followers of the cognitive-psychological approach have debated how best to explain variation in performance across tasks: Although some have argued that failures of logical reasoning are caused by random errors, others have shown that these errors are correlated across tasks. The observation that some people make more errors than others suggests computational limitations that vary systematically across individuals (Stanovich, 1999). That such a finding could be controversial would astonish most researchers coming from the psychometric approach.

Mental Rules or Mental Models?

Two theories have dominated psychological theorizing about reasoning: mental rules and mental models. Both theories were first applied to the study of deductive reasoning tasks such as syllogisms and then later applied to a broader range of reasoning tasks. The mental rules theory of deductive reasoning (Rips, 1994) posits mental processes common to all typically developed adults that operate directly on the representations of the premises. Humans are assumed to be natural logicians who are sometimes fallible because of errors in processing or because of limitations of the human cognitive system. According to mental rules theory, the basic processes involved in solving deductive reasoning problems are (1) encoding the premises into representations stored in working memory, (2) applying abstract, rule-based schemas to these representations to derive a conclusion, and (3) applying other rules to check the contents of working memory for incompatibilities. Although the model posits several sources of error, the number of steps to be executed in applying rules is the major source of difficulty. Errors in performance are thus primarily attributable to working memory overload (Gilhooly, 2004).

The mental models theory (Johnson-Laird, 2004; Johnson-Laird & Khemlani, 2013) of deductive reasoning posits that the individual first transforms the premises of an argument into another representation (i.e., a mental model) that is consistent with the premises. Importantly, multiple mental models that are consistent with the premises must often be constructed and then compared in order for a valid conclusion to be reached. Each mental model represents a possible state of affairs that must be evaluated. Bara, Bucciarelli, and Johnson-Laird (1995) identified the following factors that affect syllogistic inference in the mental models approach: (1) assembling a propositional representation of premises; (2) constructing models that integrate information from premises; (3) formulating a conclusion that integrates relationships not expressed in the premises; (4) searching for alternative models to refute conclusions; and (5) recognizing similarities between models. All these processes require working memory resources. Limitations of working memory are considered especially important in this theory in understanding individual differences in reasoning, because working memory limits the number of mental models that can be held in mind at once. Individuals with limited working memory capacity can fail to generate enough models to evaluate the validity of a conclusion (Stanovich, Sá, & West, 2004).

In his recent work, Johnson-Laird has incorporated Kahneman's System 1 and 2 theory into his explanation of reasoning failures. Put simply, Kahneman's System 1 (2011) is a heuristic-based, intuitive reasoning process that is used automatically in everyday problem-solving. Its definition is similar to *tacit* processing (Evans & Over, 1996; Stanovich, 1999), which are processes that facilitate reasoning without conscious intervention and outside awareness. System 2 is the rule-based, effortful, intentional reasoning process that takes deliberate effort to engage in. Johnson-Laird and Khemlani (2013) equate System 1 with a simplified analysis where only one model is checked against the premises. System 2 thinking requires analyzing all possible models implied by the premises and checking the conclusions against the models. In their review, Johnson-Laird and Khemlani (2013) summarized their research suggesting that logical premises that led to more mental models led to slower responses and higher error rates. Importantly, they also found that the most difficult logic problems led to errors consistent with a heuristic approach (System 1) rather than deliberative thought (System 2), suggesting that participants fall back on heuristics and faster processing when the item complexity exceeds their ability (or motivation).

The mental rules and mental models theories of reasoning propose universal but somewhat contradictory mechanisms for deductive reasoning (M. J. Roberts, 1993). Furthermore, advocates of both theories have been able to marshal considerable evidence in support of their position. Research that explicitly attempts to account for individual differences in reasoning offers a possible explanation for this paradox: On some problems, the behavior of some reasoners is more consistent with the mental models theory, whereas the behavior of other reasoners is more consistent with the predictions of a mental rules theory (Stanovich et al., 2004). In addition to stable individual differences in propensity to solve reasoning problems in one way or another, how the problem is presented can encourage individuals to change their strategies across items (Galotti, Baron, & Sabini, 1986). Therefore, what a task measures cannot be determined by simple inspection. Rather, what is measured depends on a complex interaction between the characteristics of the examinee, the task, and the situation. This does not mean, however, that one cannot know what tasks typically measure when they are attempted by individuals of known characteristics, but what tasks measure and for whom and under what circumstances are inferences that must be supported by other data – not merely presumed to be the case.

Measuring Reasoning Abilities: Considerations and Validity Evidence

Inferences about the psychological constructs that a test measures in any particular application require multiple sources of evidence. The two major aspects of construct validation are nicely captured in Embretson's (1983) distinction between *construct representation* and *nomothetic span*. *Construct representation* refers to the identification of psychological constructs (e.g., component processes, strategies, structures) that individuals typically use in responding to items on a test. The cognitive psychological research on families of reasoning tests or tasks summarized in the section Cognitive Psychological Approach to Reasoning and in other sources (Gentner & Markman, 1997; Sternberg, 1986) provides the foundation for this aspect of construct validation.

Nomothetic span, on the other hand, concerns evidence on the nature of a construct that derives from its relationships with other constructs. For constructs that are grounded in individual differences, these inferences are based on the complex web of relationships among scores on tests that are designed to measure different constructs. Since the patterns of individual differences on a test depend both on the characteristics of the sample of test-takers and the number and nature of other tests included in the study, inferences about the nomothetic span of a test gain credence only after the test has been used in many different studies. The aspect of construct validation captured by nomothetic span affirms the importance of understanding individual differences on families of reasoning tasks, not simply on one or two tasks that have sparked interest among researchers. It follows that using a test in which all items follow the same format to define *reasoning* (or even worse, to define *intelligence*) reflects a fundamental misunderstanding of psychological measurement.

Construct Definition and Breadth

Performance on one item measures mostly idiosyncratic variance and provides little information about individual differences compared to the consistent variance from a test composed of similar items. Likewise, just one test provides less information than the broader ability construct defined by performance on several tests. Research on reasoning requires a method for measuring reasoning abilities. Although a single test task is often used in experimental research, the term "ability" implies consistency in performance across some defined class of tasks (Carroll, 1993). Indeed, some of the confusions and controversies in the field stem from equating performance on a particular task with the broader psychological construct. Psychological tests are simply organized collections of such tasks. However, typically less than half of the variation on well-constructed, reliable tests is shared with other tests that measure the same construct using somewhat different kinds of test tasks. An early but still reasonable rule in psychological measurement is that when measuring any ability, one should combine performances across at least three different measures that use different formats to reduce the specific effects of individual tasks (Süß & Beauducel, 2005).

Although many different tasks have been used to measure reasoning, a few are used much more commonly than others: analogies, matrix problems, series completions, and classification tasks. Some test batteries also measure verbal reasoning through sentence completion tests, sentence comprehension tests, and even vocabulary[2]. Others include more specific spatial tasks, such as form boards or paper-folding tests. And others use quantitative tests that require examinees to make relational judgments (such as *greater than* or *less than*) between quantitative concepts or to determine how numbers and mathematical operators can be combined to generate a product.

Complexity and Nomothetic Span

Hundreds of studies have estimated relationships between reasoning tests and other kinds of ability tests and show that reasoning tests are good measures of general ability. But evidence of construct representation is needed to explain *why* reasoning tests are such good measures and what essential processes they tap into that could explain this relationship. Two-dimensional scalings of the correlations among large batteries of tests reveal complex tests that load heavily on *g* (or G*f*) fall near the center of the plot, whereas simpler tasks are distributed around the periphery (e.g., Snow, Kyllonen, & Marshalek, 1984).

Several hypotheses have been advanced to explain how processing complexity increases along the various spokes that run from the periphery to *g*: (1) an increase in the number of component processes; (2) an accumulation of differences in speed of component processing; (3) an increase in the involvement of one or more critically important performance components, such as the inference process; (4) an increase in demands on limited working memory or attention; and (5) an increase in demands on adaptive functions, including assembly, control, and monitoring functions. Clearly

[2] A wide-ranging vocabulary test can be a surprisingly good measure of reasoning. The ability to infer meanings of unknown words and retain that information from fewer exposures results from stronger verbal reasoning skills.

these explanations are not independent. For example, it is impossible to get an accumulation of speed differences over components (Hypothesis 2) without also increasing the number of component processes required (Hypothesis 1). Despite this overlap, these hypotheses provide a useful way to organize the research.

Item Process: Manipulating Reasoning Load

Analysis of task complexity and correlations to intelligence, including fluid reasoning, have repeatedly shown that the more central or *g*-loaded tasks require subjects to do more than the more peripheral tests (Larson, Merritt, & Williams, 1988; Marshalek, Lohman, & Snow, 1983; Primi, 2001). Spilsbury (1992) argued that the crucial manipulation was an increase in the factorial complexity of a task (that is, the number of different abilities required). However, increases in the number or difficulty of task steps beyond a certain point can decrease the correlation with *g* (Crawford, 1988; Raaheim, 1988; Swiney, 1985). Thus, one does not automatically increase the relationship with *g* simply by making problems harder, or even by increasing the factorial complexity of a task. Indeed, there are many hard problems (e.g., memorizing lists of randomly chosen numbers or words) that are not particularly good measures of *g*. Furthermore, even for problems that do require the type of processing that causes the test to measure *g*, problems must be of the appropriate level of difficulty for subjects.

Several investigators have attempted to manipulate the extent to which items require assembly and control processes and thereby alter their relationship with *g*. For example, Swiney (1985) sought to test the hypothesis that correlations between performance on geometric analogies and *g* would increase as more flexible adaptation was required, at least for easy and moderately difficult problems. Correlations with *g* were expected to decline if task difficulty was too great. Adaptation was manipulated by grouping items in different ways. In the blocked condition, inter-item variation was minimized by grouping items with similar processing requirements (estimated by the number of elements, and the number and type of transformations). In the mixed condition, items were grouped to be as dissimilar as possible, requiring maximally flexible adaptation. Results showed that low-ability students were more adversely affected by mixing items than were high-ability students, regardless of treatment order. Relationships between task accuracy and *g* varied systematically as a function of item difficulty and task requirements. The strongest relationships were observed for items that required students to identify or apply difficult rules. Retrospective reports supported the conclusion that high-*g* subjects were better able to adapt their strategies flexibly to meet changing task demands.

Chastain (1992) reported three similar studies contrasting blocked versus mixed item presentations and found small relationships consistent with Swiney's (1985) hypotheses that mixed items would show greater *g* loading. An opposite finding, however, was reported in a study by Carlstedt, Gustafsson, and Ullstadius (2000). Three kinds of inductive reasoning problems were administered to groups of Swedish military recruits. Carlstedt and colleagues unexpectedly found that *g* loadings were higher in the blocked condition than in the mixed condition; they argued that the homogeneous arrangement affords better possibilities for learning and transfer across items. However, the items were extremely difficult, and so generalization is limited.

Criterion Evidence from School Learning

Information on the nomothetic span of a test also comes from the sorts of criterion behaviors that the test predicts. Measures of general reasoning ability (or G*f*) are good predictors of success in learning a broad range of tasks. Correlations are generally highest for the early phases of learning new, especially open-ended skills (Ackerman, 1988, 2005) and for learning the sorts of organized systems of meaningful concepts that are commonly required in formal schooling. Population correlations with measures of school success range from $r = 0.4$ to 0.8, depending on the criterion measure (e.g., grades, achievement tests) and of content of reasoning test (e.g., verbal, quantitative, figural).

Reasoning tests correlate with academic success because school learning requires reasoning abilities. Successful learning of all kinds require reasoning: understanding a story, inferring the meaning of an unfamiliar word, detecting patterns and regularities in information, abstracting the information given to form more general rules or principles, applying mathematical concepts to solve a problem, etc. Perhaps paradoxically, the best way to develop reasoning abilities is through challenging instruction that requires students to exercise old reasoning strategies and to invent or learn new ones (Martinez, 2000; Nickerson, 2004).

These important reasoning skills are captured even by what some would consider narrow measures of achievement, like vocabulary tests. Individual differences on vocabulary tests may arise from variance in how well learners use certain metacognitive or performance processes when learning – such as systematically testing alternative interpretations of a word when it is used in unfamiliar contexts – that then lead to a richer and more usefully organized knowledge base to guide new learning (Robinson & Hayes, 1978). Marshalek (1981) concluded that the ability to infer word meanings from the contexts in which they occur is the cause of the high correlations typically observed between vocabulary and reasoning tests. But there is also a synergism in that vocabulary knowledge allows comprehension and expression of a broader array of ideas, which in turn facilitate the task of learning new words and concepts. Thus, language functions as a vehicle for the expression, refinement, and acquisition of thought, and the humble vocabulary test masks an enormous amount of reasoning and remembering.

Relationship of Reasoning to Other Constructs

Reasoning and Expertise

Reasoning well in domains of nontrivial complexity depends importantly on knowledge. Expertise is rooted in knowledge, and experts reason differently about problems than do novices (Feltovich, Prietula, & Ericsson, 2006; Markman & Gentner, 2001). Because of this, some have erroneously assumed that good reasoning is nothing more than good knowledge. This does not take into account the importance of good reasoning in the acquisition of a well-ordered knowledge base. Everyday reasoning depends heavily on the efficacy of past reasoning processes (stored as knowledge) as well as the efficacy of present reasoning processes. An increasingly sophisticated knowledge base supports increasingly sophisticated forms of reasoning. A more sophisticated knowledge base has richer, more abstract associative connections between concepts and more metacognitive knowledge that links strategies to goals. The ability to efficiently process information (e.g., chunking based on expert understandings) frees working memory resources for problem-solving (Feltovich et al., 2006; Gobet & Waters, 2003; Horn & Masunaga, 2006; Proctor & Vu, 2006).

Studies of tasks modeled after item types on intelligence tests often ignore these contributions of knowledge – particularly domain-specific knowledge – to reasoning. The loss is probably most obvious in the domain of verbal reasoning. The verbal reasoning skills of lawyers or scientists go well beyond the sorts of decontextualized reasoning abilities assessed on most mental tests. A rich understanding of a domain and of the conventions of argumentation in that domain are needed to identify relevant rather than irrelevant information when understanding the problem, to decide which alternatives are most plausible and need to be considered, and then to decide how best to marshal evidence in support of a position. Strong warrants for an argument are considered highly plausible by those evaluating it. Plausibility judgments reflect both the beliefs of listeners and their assessment of the logical consistency of the argument. Standards for evaluating arguments are thus necessarily somewhat subjective. Nevertheless, some types of arguments are widely recognized as logically unsound. Toulmin, Rieke, and Janik (1984) classify these as (1) missing grounds (e.g., begging the question); (2) irrelevant grounds (e.g., red herring); (3) defective grounds (e.g., hasty generalization); (4) unwarranted assumptions; and (5) ambiguities.

Careful studies of reasoning in knowledge-rich contexts also show processes that generalize across domains. Newell and Simon's (1972) distinction between strong and weak methods of reasoning is especially helpful here. *Strong methods* of reasoning rely heavily on knowledge within a particular domain, whereas *weak methods* depend less on content and context. That is, strong (or domain-specific) methods describe what people do when they *do know* what to do; weak (or domain-general) methods describe what people do when they *do not know* what to do. Therefore, children and novices are more likely to use domain-general methods. Strong methods are closer to the construct of fluid reasoning ability whereas weak methods are closer to the construct of crystallized ability, at least as Cattell (1963) originally defined these constructs. Note, however, that evidence showing transfer of strong problem-solving methods concurs with the finding that fluid reasoning abilities are developed, not fixed.

Relationship with Education and Achievement

One really does not know what abilities are unless one knows how they develop. Reasoning abilities are not only critical aptitudes for learning but they are also among its most important outcomes. Instructional interventions that explicitly require and succeed in developing students' reasoning abilities comprise one of the best sources of evidence on the construct validity of reasoning tests (Snow & Lohman, 1989).

The nature of the statistical interaction between instructional treatments and reasoning abilities is straightforward. Instructional methods that place the burden of making inferences and deductions on the student increase the relationship between reasoning abilities and achievement. Instructional methods that scaffold, remove, or otherwise reduce this burden reduce the relationship between reasoning abilities and achievement. The relationship is moderated by other variables, particularly anxiety, but reasoning abilities and prior knowledge in the domain are clearly the most important aptitudes for learning from instruction. Put differently, those who hope to enhance the probability of successful completion of school by offering different instructional opportunities are most likely to succeed if the adaptations are based on the developed broad reasoning abilities of students rather than narrow cognitive styles.

Gambrell (2013) explored the relative effects of schooling on a variety of tasks arrayed from emphasizing reasoning (G*f*) to emphasizing achievement (G*c*) skills. Gambrell concluded that tasks that were less directly influenced by schooling (relative to aging) were those that emphasized transfer and generalization, integrating or synthesizing a large amount of information, and/or were presented in ways that differed from typical school presentations. In other words, the more general the skill and removed the task was from specific content, the more reasoning is required. As a result, when reasoning demands increased, the effects of age were more important than the amount of schooling.

Relationship to Workplace Outcomes

The substantial association between *g* and workplace success (e.g., career accomplishment, job performance, occupational attainment) is well established and has been extensively reviewed elsewhere (e.g., Chernyshenko, Stark, & Drasgow, 2011; Ng et al., 2005; Ones, Dilchert, & Viswesvaran, 2012; Ones et al., 2010; Strenze, 2015). Here, we highlight several strands of workforce-related research that bear upon the interpretation of *g* as a dimension describing differences between individuals in their ability to reason effectively in complex environments.

Perhaps because of the frequent identification of *g* with fluid ability, explicit examinations of the relationship between G*f* and job performance (as opposed to *g* and job performance) have been rare. The focal point of a recent dissertation (Postlethwaite, 2011) was a comprehensive meta-analysis of the ability-job performance association that explicitly categorized ability tests into measures of G*f* or G*c*. Surprisingly, the author was only able to identify twelve studies (5.38% of the total sample) that featured tests that could be classified as "pure" measures of G*f*. Across all the jobs surveyed the raw correlation between G*f* and performance was 0.14 (0.29 when corrected for statistical artifacts) – as opposed to uncorrected and corrected correlations of 0.23 and 0.54 for Gc, respectively. However, the validity of fluid ability for predicting performance increased with job complexity. These findings align well with meta-analytic results for *g* and job complexity (Salgado, 2017; Schmidt & Hunter, 2004). They also shed light on research showing that the validity of *g* (so often isomorphic with G*f*) for predicting job performance decreases over time in less complex jobs but remains a strong predictor of performance in more complex jobs (Farrell & McDaniel, 2001; Kanfer & Ackerman, 2004). This has practical implications for the workforce given that fluid ability declines with age (Horn, 1998).

To further complicate matters, some occupations are very similar in their overall complexity but differ greatly in their reasoning demands across specific content domains. For example, subject-matter experts (SMEs) rate "surgeon" and "chief executive" about the same in terms of overall complexity but the job of surgeon is rated as placing high demands on visuospatial ability while chief executive is rated as placing high demands on mathematical ability (US Department of Labor, Employment and Training Administration, 2018). Although it is always important to assess reasoning abilities across multiple content domains (Süß & Beauducel, 2005), it is especially critical to consider specific facets of reasoning when attempting to make accurate predictions about what occupations people will likely be attracted to, perform well in, and be satisfied with.

Role of Speed or Efficiency

Can differences in reasoning skill be attributed entirely to neural efficiency as measured by reaction time? This hypothesis has taken several forms. In its strongest form, the assertion has been that individuals differ in the general speed or efficiency with which they process information, possibly as a result of more efficient brain structures (Jensen, 1998). Although disattenuated correlations between reaction time (RT) and *g* can be substantial when samples vary widely in ability (even, for example, including participants with an intellectual disability), samples more typical of those used in other research on abilities yield correlations between RT and *g* in the $r = -0.1$ to -0.4 range (Deary & Stough, 1996; Jensen, 1982; R. D. Roberts & Stankov, 1999; Sternberg, 1985). In principle, processing speed could be estimated on any elementary cognitive task that minimizes the importance of learning, motivation, strategy, and other confounding variables. In fact, response latencies on many tasks show a pattern of increasing correlation with an external estimate of *g* as task complexity decreases. In other words, response latencies for simpler tasks typically show higher correlations with *g* than do response latencies for more complex tasks. But this is unsurprising. The more complex the task, the more room there is for subjects to use different strategies or even to be inconsistent in the execution of different components across items.

In its weak form, the hypothesis has been that although speed of processing on any one task may be only weakly correlated with more complex performances, such small differences cumulate over time and tasks. Thus, Hunt, Frost, and Lunneborg (1973) noted that although latency differences in the retrieval of overlearned name codes correlated only $r = 0.3$ with verbal ability, such small differences on individual words cumulate to substantial differences in the course of a more extended activity such as reading comprehension. Detterman (1986) emphasized the cumulation across different component processes rather than across time. He showed that although individual component processes were only weakly correlated with *g*, their combined effect on a complex task was more substantial.

Although individual differences in speed of processing are an important aspect of *g, g* is more than rapid or efficient information processing. Furthermore, the strength of the relationship between speed of processing and *g* varies considerably across domains, being strongest ($r \approx -0.4$) in the verbal domain and weakest ($r \approx -0.2$) in the spatial domain. Indeed, for complex spatial tasks, the speed with which individuals perform different spatial operations is usually much less predictive of overall performance than the richness or quality of the mental representations they create (Lohman, 1988; Salthouse et al., 1990).

Relationship to Working Memory

One of the more important controversies about reasoning abilities is the extent to which individual differences in reasoning abilities overlap with individual differences in working memory capacity. Kyllonen and Christal (1990) sparked the controversy with their finding that latent variables for working memory and reasoning factors correlated $r = 0.80$ to 0.88 in four large studies with US Air Force recruits. Other researchers also found large path coefficients between measures of working memory and measures of fluid reasoning abilities (Conway et al., 2002; Süß et al., 2002). However, critics complained that some tasks used to estimate working memory in these studies were indistinguishable from tasks used to estimate reasoning. Other critics (e.g., Fry & Hale, 1996) have argued that processing speed accounts for most of the relationship between the reasoning and working memory constructs in these studies. Ackerman, Beier, and Boyle (2002) noted that processing speed is itself a multidimensional construct. They conclude that although there is little doubt that measures of working memory are significantly associated with measures of general intelligence, the two are not synonymous. Indeed, a meta-analysis of the existing data yielded a true-score correlation of $r = 0.48$ between working memory and *g*, far below the unity some claim (Ackerman, Beier, & Boyle, 2005).

In part, this is a problem of words. The term *working memory* connotes too small a construct; *reasoning* connotes too large a construct – especially given the way each is typically measured. Consider first the reasoning construct. In the best of these studies, reasoning is estimated by performance on a series of short, puzzle-like tasks. More commonly, it is estimated by a single test such as Raven's Progressive Matrices (Raven et al., 1977) which uses a single item format. As Ackerman and colleagues (2002) note, "if the Raven is not an exemplary measure of general intelligence (or even Gf), any corroborations between experimental measures (such as [working memory]) and Raven ... are apt to miss important variance ... and result in distortion of construct validity" (p. 586). Indeed, figural reasoning tests such as the Raven are typically much poorer predictors of both real-world learning and academic achievement than measures of verbal and quantitative reasoning. For example, Lohman, Korb, and Lakin (2008) administered the Standard Progressive Matrices (Raven et al., 1977), the Naglieri Nonverbal Ability Test (Naglieri, 1996), and Form 6 of the Cognitive Abilities Test (Lohman & Hagen, 2001) to approximately 1,200 children in grades K–6. Correlations with multiple measures of reading and mathematics achievement varied from $r = 0.3$ to 0.7 for all three nonverbal reasoning tests. The corresponding correlations for the CogAT Verbal and Quantitative batteries ranged from $r = 0.7$ to 0.8. Technical manuals for ability tests that are co-normed with achievement tests provide similar information, but on large nationally representative samples of students in grades K–12 (e.g., Lohman & Lakin, 2017). Raven was well aware of the restricted construct representation of the Progressive Matrices test. Because of this, he advised never to administer the test alone when making decisions about students but always to administer a verbal reasoning test as well (Raven et al., 1977). Therefore, whether measured by one task or several short tasks, the reasoning construct is underrepresented in virtually all research studies.

On the other hand, the construct measured by the series of working memory tests is much more complex than its label suggests. These tasks generally require participants to understand and follow a sometimes complex set of directions; to assemble and then revise a strategy for performing a difficult, attention-demanding task; to maintain a high level of effort across a substantial number of trials; and then to repeat the process for a new task with a new set of directions. In addition, many working memory tasks require individuals to simultaneously process one set of ideas while remembering another set. Although the individual tasks are generally thought to be easy, they are certainly not trivial, especially when performed under memory load. These tasks elicit executive functions such as the monitoring of processes, controlling their rate and sequence of operation, inhibiting inappropriate response processes, coordinating information from different domains, and integrating ideas into a coherent mental model. Such executive functions clearly overlap with many researchers' conceptions of reasoning or even of general intelligence. This heated debate may boil down to a difference in branding caused by the parallel development of closely related constructs in both psychometric and cognitive traditions.

Relationship to Critical Thinking

Critical thinking is a construct often discussed in relation to intelligence and reasoning and is enjoying popularity in educational policy circles (Pellegrino & Hilton, 2015), although it may be more a matter of branding than a construct truly distinct from reasoning and intelligence. Supporters of critical thinking often claim that the construct is related to, but distinct from, general ability (e.g., Halpern & Butler, 2018; Stanovich, West, & Toplak, 2013). Indeed, some have claimed that the two domains are relatively independent (Stanovich & West, 2008) and that fluid ability is related to critical thinking "indirectly and to a mild extent" (Stanovich & Stanovich, 2010, p. 217).

Ascertaining the extent to which critical thinking and g or Gf are related is challenging due to definitional difficulties (Huber & Kuncel, 2016). Liu, Frankel, and Roohr (2014) reviewed multiple critical frameworks and found that several explicitly mentioned reasoning generically or deductively and/or inductively while several others did not explicitly incorporate reasoning at all. On the other hand, Stanovich and colleagues' approach to critical thinking emphasizes the tendency to purposefully engage in effortful, analytic thought, which they claim is not captured by traditional cognitive tests (Elson et al., 2018) – yet fluid ability has been characterized as "the deliberate but flexible control of attention to solve novel, 'on-the-spot' problems that cannot be performed by relying exclusively on previously learned habits, schemas, and scripts" (Schneider & McGrew, 2012, p. 111). Some studies have even used critical thinking measures *as* tests of intelligence (Arteche et al., 2008; Furnham, Crump, & Chamorro-Premuzic, 2007).

Although there appears to be no consensus as to a singular best definition or assessment of critical thinking, meta-analytic research suggests associations between critical thinking and *g* cluster around 0.40.[3] Nonetheless, some critical thinking tasks do evince very small relations with *g*: Performance on items assessing heuristics and biases involving the law of large numbers, detecting covariation, and conjunctions have been found to correlate near zero with SAT scores (West, Toplak, & Stanovich, 2008). Strong judgments about the relationship between critical thinking and fluid reasoning should still be withheld, however, given the diversity of definitions and assessments and the tendency for investigators to rely on college populations and tests given under low-stakes conditions, which can introduce statistical artifacts (e.g., range restriction) and confounding sources of variance (e.g., motivation).

Additional evidence for the discrete value of critical thinking vis-à-vis fluid intelligence comes from the incremental value of the construct beyond a measure of *g* for predicting practical criteria. Elson and colleagues (2018) developed an assessment of four critical thinking domains (causal reasoning, identifying assumptions, hypothesis evaluation, logical reasoning) and administered it to government analysts, in addition to gathering their self-report SAT/ACT scores. The criterion was a work simulation test that was rated on five dimensions, along with overall quality, by subject-matter experts. Total scores on the critical thinking test accounted for an additional 11 percent of the variance in a composite of the SMEs' dimension-specific ratings and 18 percent of the variance in their overall ratings beyond SAT/ACT scores. As far as we are aware this is the only study that has investigated the incremental validity of critical thinking for predicting applied outcomes beyond general ability. Given the ubiquity of cognitive ability tests and many different constructs grouped under the label "critical thinking," additional research of this type is recommended prior to developing and administering critical thinking tests for widespread practical use, to confirm they consistently add value beyond the many preexisting cognitive assessments (cf. Sechrest, 1963).

[3] Kuncel's (2011) meta-analysis produced an observed correlation of 0.48. Liu et and colleagues' (2014) meta-analysis reported multiple study-specific correlations that ranged from0.15 to 0.59, with many in the 0.20–0.30 range.

Conclusions

Reasoning abilities are not static. They are developed through experience and rendered easier to perform through exercise. Individual differences in reasoning are substantially correlated with the amount of information individuals can hold in working memory while performing some transformation on it. The ability to do this depends in large measure on the attentional resources individuals bring to a task, their familiarity with the to-be-remembered information, and their skill in performing the required transformations. Thus, prior knowledge and skill are critical determiners of the level of reasoning that one can exhibit both on reasoning tests and in everyday tasks. The processes that support sophisticated reasoning by experts in a knowledge-rich domain, however, appear to be largely the same as those which enable the novice to infer consistencies or deduce likely consequents in novel problem–solving.

Reasoning has occupied a central place in conceptions and tests of intelligence since Binet's original assessment (e.g., Burt, 1909; Ryans, 1938; Spearman, 1923), allowing over a century's worth of research activity to firmly establish the construct validity of reasoning tests. Nonetheless, construct validation is a never-ending process and many important avenues of research remain open for exploration. We briefly describe several interrelated future directions, although there are many others.

First, it would be worthwhile to design reasoning tests more systematically, for example by basing item features on a priori taxonomies of task complexity (e.g., Campbell, 1988; Primi, 2001; Wood, 1986) and incorporating newer psychometric models that derive from cognitive psychological models and better elucidate test-takers' response processes (e.g., Embretson, 2016). These more precise procedures may allow for a more definitive conclusion as to whether it is feasible to consistently distinguish deductive and inductive inference, especially in light of arguments that it is often not even possible to tease apart whether individual test-takers rely on knowledge or reasoning to solve problems (e.g., Johnson & Bouchard, 2005b; Johnson & Gottesman, 2006).

Second, it is important to ascertain the extent to which different forms of reasoning can be separated from different forms of content, as decades of using of the same item types to assess the same forms of reasoning may have led process and content to become confounded. How plausible is it, for instance, to measure inductive reasoning using verbal tasks and deductive reasoning using figural tasks? If process cannot be fully separated from content it has implications for test design and how human abilities are construed. Third, researchers should work toward reconciling the CHC, BIS, and *g*-VPR models because each treats reasoning in different ways, with one postulating that inductive and deductive inferences can be separated using psychometric tests, one postulating that they cannot, and one being somewhat agnostic about where reasoning lies in the hierarchy of cognitive abilities. All three cannot be fundamentally correct and they likely vary in their usefulness for guiding psychometric testing. Not only should the degree to which the three models fit data generated by large test batteries across multiple samples, they should also be compared in the extent to which they predict external criteria (e.g., grades, occupational attainment). Neurological plausibility must also be considered.

Fourth, we have focused our attention on inductive and deductive reasoning because those are the forms of reasoning that standardized tests tend to assess. However, there is a third major form of inference, *abductive reasoning*, which is similar to induction in that the conclusion drawn is uncertain but goes further than induction by positing what is held to be the best explanation for the given observations (Douven, 2017). Abduction has been argued to be a valuable scientific tool (Haig, 2005; Rozeboom, 1997) and it is worth building tasks dedicated to assessing it, in order to determine whether this form of reasoning is psychometrically separable from induction and deduction and has unique predictive and instructional value.

Our final, and perhaps most fundamental, recommendation is that more work should occur that attempts to link the reasoning processes elicited by psychometric tasks to neural activation patterns and specific brain regions. Given that debate exists as to whether inductive and deductive reasoning can be reliably disentangled, different people use different strategies to approach the same items, and the brain regions activated by the same tasks can differ according to sample characteristics (e.g., degree of expertise, sex; Bilalic & Campitelli, 2018; Haier et al., 2005), this will be a major challenge.

Reasoning ability is fundamental to human abilities and productivity. Research exploring its definition, components, and development are bound to have practical importance for education, workplace, and societal outcomes. We should continue to explore its measurement and not be satisfied with simply administering one inductive reasoning task to capture it.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, *117*, 288–318.

Ackerman, P. L. (2005). Ability determinants of individual differences in skilled performance. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence: Identifying the mechanisms of the mind* (pp. 142–159). Cambridge, UK: Cambridge University Press.

Ackerman, P. L., Beier, M., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, *131*, 567–589.

Ackerman, P. L., Beier, M., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*, 30–60.

Arteche, A., Chamorro-Premuzic, T., Furnham, A., & Crump, J. (2008). The relationship of trait EI with personality, IQ and sex in a UK sample of employees. *International Journal of Selection and Assessment*, *16*, 421–426.

Bara, B. G., Bucciarelli, M., & Johnson-Laird, P. N. (1995). Development of syllogistic reasoning. *American Journal of Psychology*, *108*, 157–193.

Beauducel, A., Brocke, B., & Liepmann, D. (2001). Perspectives on fluid and crystallized intelligence: Facets for verbal, numerical, and figural intelligence. *Personality and Individual Differences*, *30*(6), 977–994.

Beauducel, A., & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence Structure (BIS). *European Journal of Psychological Assessment*, *18*(2), 97–112.

Bilalic, M., & Campitelli, G. (2018). Studies of the activation and structural changes of the brain associated with expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 233–256). Cambridge, UK: Cambridge University Press.

Bruner, J. S. (1957). *Going beyond the information given*. New York: Norton.

Burt, C. (1909). Experimental tests of higher mental processes and their relation to general intelligence. *Journal of Experimental Psychology*, *2*, 94–177.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of Management Review*, *13*, 40–52.

Carlstedt, B., Gustafsson, J.-E., & Ullstadius, E. (2000). Item sequencing effects on the measurement of fluid intelligence. *Intelligence*, *28*, 145–160.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*(3), 153–193.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1–22.

Chastain, R. L. (1992). Adaptive processing in complex learning and cognitive performance. Unpublished doctoral dissertation, Stanford University, Stanford.

Chernyshenko, O. S., Stark, S., & Drasgow, F. (2011). Individual differences: Their measurement and validity. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology* (vol. 2, pp. 117–151). Washington: American Psychological Association.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.

Crawford, J. (1988). Intelligence, task complexity and tests of sustained attention. Unpublished doctoral dissertation, University of New South Wales, Sydney, Australia.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time: Achievements, prospects, and problems. *American Psychologist*, *51*, 599–608.

Detterman, D. K. (1986). Human intelligence is a complex system of separate processes. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition* (pp. 57–61). Norwood, NJ: Ablex.

Douven, I. (2017). *Stanford Encyclopedia of Philosophy*, s.v. Abduction. https://stanford.library.sydney.edu.au/archives/spr2017/entries/abduction

Elson, S. B., Hartman, R., Beatty, A., Trippe, M., Buckley, K., Bornmann, J., et al. (2018). Critical analytic thinking skills: Do they predict job-related task performance above and beyond general intelligence? *Personnel Assessment and Decisions*, *4*(1), 9–29.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, *93*, 179–197.

Embretson, S. E. (2016). Understanding examinees' responses to items: Implications for measurement. *Educational Measurement: Issues and Practice*, *35*, 6–22.

Evans, J. S. B. T., & Feeney, A. (2004). The role of prior belief in reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 78–102). New York: Cambridge University Press.

Evans, J. S. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove, UK: Psychology Press.

Farrell, J. N., & McDaniel, M. A. (2001). The stability of validity coefficients over time: Ackerman's (1988) model and the General Aptitude Test Battery. *Journal of Applied Psychology*, *86*, 60–79.

Feltovich, P. J., Prietula, M. J., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 41–68). New York: Cambridge University Press.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, *7*, 237–241.

Furnham, A., Crump, J., & Chamorro-Premuzic, T. (2007). Managerial level, personality and intelligence. *Journal of Managerial Psychology*, *22*, 805–818.

Galotti, K. M., Baron, J., & Sabini, J. P. (1986). Individual differences in syllogistic reasoning: Deduction rules or mental models? *Journal of Experimental Psychology: General*, *115*, 16–25.

Gambrell, J. L. (2013). Effects of age and schooling on 22 ability and achievement tests. Unpublished doctoral dissertation, University of Iowa, Iowa City, IA.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45–56.

Gilhooly, K. J. (2004). Working memory and reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 49–77). New York: Cambridge University Press.

Gobet, F., & Waters, A. J. (2003). The role of constraints in expert memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29*, 1082–1094.

Gustafsson, J.-E. (1988). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol. 4, pp. 35–71). Hillsdale, NJ: Erlbaum.

Haig, B. D. (2005). An abductive theory of scientific method. *Psychological Methods*, *10*, 371–388.

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2005). The neuroanatomy of general intelligence: Sex matters. *NeuroImage*, *25*(1), 320–327.

Halpern, D. F., & Butler, H. A. (2018). Is critical thinking a better model of intelligence? In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 183–196). New York: Cambridge University Press.

Horn, J. L. (1998). A basis for research on age differences in cognitive capabilities. In J. J. McArdle & R. W. Woodcock (Eds.), *Human cognitive abilities in theory and practice* (pp. 57–91). Mahwah, NJ: Erlbaum.

Horn, J. L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, test, and issues* (vol. 2, pp. 41–68). New York: Guilford Press.

Horn, J. L., & Masunaga, H. (2006). A merging theory of expertise and intelligence. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 587–612). New York: Cambridge University Press.

Huber, C. R., & Kuncel, N. R. (2016). Does college teach critical thinking? A meta-analysis. *Review of Educational Research*, *86*, 431–468.

Hunt, E. B. (2010). *Human intelligence*. Cambridge, UK: Cambridge University Press.

Hunt, E. B., Frost, N., & Lunneborg, C. (1973). Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), *The psychology of learning and motivation* (vol. 7, pp. 87–122). New York: Academic Press.

Jensen, A. R. (1982). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol. 1, pp. 255–310). Hillsdale, NJ: Erlbaum.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Johnson, W., & Bouchard, T. J., Jr. (2005a). Constructive replication of the visual-perceptual-image rotation model in Thurstone's (1941) battery of 60 tests of mental ability. *Intelligence*, *33*, 417–430.

Johnson, W., & Bouchard, T. J., Jr. (2005b). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.

Johnson, W., & Deary, I. J. (2011). Placing inspection time, reaction time, and perceptual speed in the broader context of cognitive ability: The VPR model in the Lothian Birth Cohort 1936. *Intelligence*, *39*, 405–417.

Johnson, W., & Gottesman, I. I. (2006). Clarifying process versus structure in human intelligence: Stop talking about fluid and crystallized. *Behavioral and Brain Sciences*, *29*, 136–137.

Johnson, W., te Nijenhuis, J., & Bouchard, T. J., Jr. (2007). Replication of the hierarchical visual-perceptual-image rotation model in de Wolff and Buiten's (1963) battery of 46 tests of mental ability. *Intelligence*, *35*, 69–81.

Johnson-Laird, P. N. (2004). Mental models and reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 169–204). New York: Cambridge University Press.

Johnson-Laird, P. N., & Byrne, R. M. (1991). *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Johnson-Laird, P. N., & Khemlani, S. S. (2013). Toward a unified theory of reasoning. *Psychology of learning and motivation*, *59*, 1–42.

Kahneman, D. (2011). *Thinking, fast and slow*, vol. 1. New York: Farrar, Straus and Giroux.

Kanfer, R., & Ackerman, P. L. (2004). Aging, adult development, and work motivation. *Academy of Management Review*, *29*, 440–458.

Kuncel, N. R. (2011). Measurement and meaning of critical thinking (Research report for the NRC 21st Century Skills Workshop). Washington: National Research Council.

Kvist, A. V., & Gustafsson, J.-E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's investment theory. *Intelligence*, *36*, 422–436.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*, 389–433.

Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence*, *12*(2), 131–147.

Leighton, J. P. (2004). The assessment of logical reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 291–312). New York: Cambridge University Press.

Liu, O. L., Frankel, L., & Roohr, K. C. (2014). Assessing critical thinking in higher education: Current state and directions for next-generation assessment. Research Report ETS RR-14-10. Princeton: Educational Testing Service.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol. 4, pp. 181–248). Hillsdale, NJ: Erlbaum.

Lohman, D. F., & Hagen, E. (2001). *Cognitive Abilities Test (Form 6*). Itasca, IL: Riverside.

Lohman, D. F., Korb, K., & Lakin, J. M. (2008). Identifying academically gifted English language learners using nonverbal tests: A comparison of the Raven, NNAT, and CogAT. *Gifted Child Quarterly*, *52*, 275–296.

Lohman, D. F., & Lakin, J. M. (2017). *Cognitive Abilities Test (Forms 7 & 8): Research handbook*. Itasca, IL: Houghton Mifflin Harcourt.

Major, J. T., Johnson, W., & Deary, I. J. (2012). Comparing models of intelligence in Project TALENT: The VPR model fits better than the CHC and extended Gf–Gc models. *Intelligence*, *40*, 543–559.

Markman, A. B., & Gentner, D. (2001). Thinking. *Annual Review of Psychology*, *52*, 223–247.

Marshalek, B. (1981). *Trait and process aspects of vocabulary knowledge and verbal ability.* Technical Report No. 15. Aptitude Research Project, School of Education, Stanford University, Stanford.

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence*, *7*(2), 107–127.

Martinez, M. E. (2000). *Education as the cultivation of intelligence*. Mahwah, NJ: Erlbaum.

Naglieri, J. A. (1996). *Naglieri Nonverbal Ability Test*. San Antonio, TX: Harcourt Brace Educational Measurement.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ng, T. W., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, *58*(2), 367–408.

Nickerson, R. S. (2004). Teaching reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 410–442). New York: Cambridge University Press.

Ones, D. S., Dilchert, S., & Viswesvaran, C. (2012). Cognitive abilities. In N. Schmitt (Ed.), *Oxford library of psychology* (pp. 179–224). New York: Oxford University Press.

Ones, D. S., Dilchert, S., Viswesvaran, C., & Salgado, J. F. (2010). Cognitive abilities. In J. Farr & N. Tippins (Eds.), *Handbook of employee selection* (pp. 235–275). Clifton, NJ: Psychology Press.

Pellegrino, J. W., & Hilton, M. L. (2015). *Education for life and work: Developing transferrable knowledge and skills in the 21st century.* Washington: National Academies Press.

Postlethwaite, B. E. (2011). Fluid ability, crystallized ability, and performance across multiple domains: A meta-analysis. Unpublished doctoral dissertation, University of Iowa, Iowa City.

Primi, R. (2001). Complexity of geometric inductive reasoning tasks contribution to the understanding of fluid intelligence. *Intelligence*, *30*, 41–70.

Proctor, R. W., & Vu, K. L. (2006). Laboratory studies of training, skill acquisition, and retention of performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, &

R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 265–286). New York: Cambridge University Press.

Raaheim, K. (1988). Intelligence and task novelty. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol. 4, pp. 73–97). Hillsdale, NJ: Erlbaum.

Raven, J. C., Court, J. H., & Raven, J. (1977). *Raven's Progressive Matrices and Vocabulary Scales*. New York: Psychological Corporation.

Rips, L. J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.

Roberts, M. J. (1993). Human reasoning: Deductive rules or mental models, or both? *Quarterly Journal of Experimental Psychology*, *46A*, 569–589.

Roberts, R. D., & Stankov, L. (1999). Individual differences in speed of mental processing and human cognitive abilities: Toward a taxonomic model. *Learning and Individual Differences*, *11*, 1–120.

Robinson, C. S., & Hayes, J. R. (1978). Making inferences about relevance in understanding problems. In R. Revlin & R. E. Mayer (Eds.), *Human reasoning* (pp. 195–206). Washington: Halsted Press.

Rozeboom, W. W. (1997). Good science is abductive, not hypothetico-deductive. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 335–391). Hillsdale, NJ: Erlbaum.

Ryans, D. G. (1938). A study of the observed relationship between persistence test results, intelligence indices, and academic success. *Journal of Educational Psychology*, *29*, 573–580.

Salgado, J. F. (2017). Using ability tests in selection. In *The Wiley Blackwell handbook of the psychology of recruitment, selection and employee retention* (pp. 113–150). Chichester, UK: John Wiley & Sons Ltd.

Salthouse, T. A., Babcock, R. L., Mitchell, D. R. D., Palmon, R., & Skovronek, E. (1990). Sources of individual differences in spatial visualization ability. *Intelligence*, *14*, 187–230.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162–173.

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 99–144). New York: Guilford Press.

Sechrest, L. (1963). Incremental validity: A recommendation. *Educational and Psychological Measurement*, *23*, 153–158.

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of learning and ability correlations. *Advances in the Psychology of Human Intelligence*, *2*, 47–103.

Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. Linn (Ed.), *Educational measurement* (vol. 3, pp. 263–331). New York: Macmillan.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition*. London: Macmillan.

Spilsbury, G. (1992). Complexity as a reflection of the dimensionality of a task. *Intelligence*, *16*, 31–45.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Erlbaum.

Stanovich, K. E., Sá, W. C., & West, R. F. (2004). Individual differences in thinking, reasoning, and decision making. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 375–409). New York: Cambridge University Press.

Stanovich, K. E., & Stanovich, P. J. (2010). A framework for critical thinking, rational thinking and intelligence. In D. D. Preiss & R. J. Sternberg (Eds.), *Educational psychology: Perspectives on learning, teaching and human development* (pp. 195–237). New York: Springer Publishing Company.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*, 672–695.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2013). Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science*, *22*, 259–264.

Stenning, K., & Monaghan, P. (2004). Strategies and knowledge representation. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 129–168). New York: Cambridge University Press.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, UK: Cambridge University Press.

Sternberg, R. J. (1986). Toward a unified theory of human reasoning. *Intelligence*, *10*, 281–314.

Strenze, T. (2015). Intelligence and success. In S. Goldstein, D. Princiotta, & J. A. Naglieri (Eds.), *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts* (pp. 405–413). New York: Springer.

Süß, H.-M., & Beauducel, A. (2005). Faceted models of intelligence. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of measuring and understanding intelligence* (pp. 313–332). Thousand Oaks, CA: Sage.

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working memory capacity explains reasoning ability – and a little bit more. *Intelligence*, *30*, 261–288.

Swiney, J. F. (1985). A study of executive processes in intelligence. Unpublished doctoral dissertation, Stanford University, Stanford.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Toulmin, S., Rieke, R., & Janik, A. (1984). *An introduction to reasoning* (2nd ed.). New York: Macmillan.

US Department of Labor, Employment and Training Administration (2018). *Standard occupational classification manual*. www.bls.gov/soc/2018/soc_2018_manual.pdf

Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen.

Vernon, P. E. (1965). Ability factors and environmental influences. *American Psychologist*, *20*, 723–733.

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*(4), 930–941.

Wilhelm, O. (2005). Measuring reasoning ability. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of measuring and understanding intelligence* (pp. 373–392). Thousand Oaks, CA: Sage.

Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*, *37*, 60–82.

23 Problem-Solving and Intelligence

David Z. Hambrick, Alexander P. Burgoyne, and Erik M. Altmann

The ability to solve complex problems is a defining feature of what most laypeople think of as *intelligence.* This is also a common theme in how intelligence researchers describe intelligence (Sternberg, 1985a; Sternberg et al., 1981). Over a century ago, the German psychologist Wilhelm Stern (1914), who introduced the formula for computing the *intelligence quotient*, defined intelligence as "a general mental adaptability to new problems and conditions of life" (p. 101). And, in his early book on aptitude testing, Bingham (1937) explained that "[w]e shall use the term 'intelligence' to mean the ability of an organism to solve new problems" (p. 36). More recently, fifty-two intelligence researchers published a letter in the *Wall Street Journal* in the wake of the *Bell Curve* controversy, defining intelligence as "a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (see Gottfredson, 1997, p. 13).

Outline of Chapter

In this chapter, we discuss the link between intelligence and problem-solving in terms of contemporary ideas concerning both. To preview, we argue that the ability to solve problems is not just an aspect or feature of intelligence – it is the essence of intelligence. The chapter is organized into five major sections. In the first section, we consider the question of what a problem is and argue that all "intelligent" behavior can be viewed as problem-solving behavior. In the second section, we briefly review evidence from psychometric research concerning the nature of individual differences in intelligence, and then review evidence for how intelligence relates to complex problem-solving. In the third section, we consider the question of what mechanisms might underlie both problem-solving and intelligence, focusing on some of our own research. In the fourth section, we briefly review evidence for the predictive validity of intelligence and practical uses of intelligence tests. In the fifth section, we consider the question of whether intelligence as problem-solving ability can be improved through training. We close with directions for future research.

What Is a Problem?

A problem is nothing more (or less) than a goal that is not immediately attainable. As Duncker (1945) wrote, "a problem exists when a living organism has a goal but does not know how this goal is to be reached" (p. 2). And as Mayer (2013) explained, "Problem solving refers to cognitive processing directed at achieving a goal when the problem solver does not initially know a solution method. A problem exists when someone has a goal but does not know how to achieve it" (p. 769). Examples of problems range from the mundane – needing to get around a traffic jam commuting to work – to the existential – finding meaning in life. From a psychological perspective, and more particularly a cognitive perspective, the goal of research on problem-solving is to describe the mental processes involved in reducing the "distance" between the problem's initial state and the goal state.

Psychologists have traditionally distinguished between two types of problems (see Mayer, 2013). In *well-structured* problems, the goal state is clearly specified, as is the "solution path" (i.e., the specific way one should go about solving the problem). The classic example of a well-defined problem from psychological research is the Tower of Hanoi problem. The initial state in this problem is that there are three disks of three different sizes on one of three pegs. The goal is to move the disks from the left peg to the right peg, moving one disk at a time and never placing a larger disk on a smaller disk. By contrast, in an *ill-structured* problem, neither the goal state nor the solution path is clearly specified; the problem-solver must generate both. Many complex real-world problems are of this type, from producing a work of art to figuring out how to make a living. Problems also differ in their complexity – the number of steps that they involve – and in their novelty to the problem-solver – the degree to which the problem-solver has knowledge and skills that can be applied to the problem.

It is easy to think of everyday tasks that differ along these dimensions. A relatively simple well-defined problem is adding two numbers; a more complex one is filling out your tax return. A relatively simple ill-defined problem is writing a personal essay for a college application; a more complex one is planning for retirement. These tasks have the "feel" of problem-solving, but other complex tasks can be regarded as problems as well, in the sense that they involve goals that are not immediately attainable. As Anderson (1985) observed, "It seems that all cognitive activities are fundamentally problem solving in nature. The basic argument . . . is that human cognition is always purposeful, directed to achieving goals and to removing obstacles to those goals" (pp. 199–200).

Consider Raven's Progressive Matrices, a widely administered test of nonverbal intelligence developed by John C. Raven in 1936. As shown in Figure 23.1, the initial state of each test item is a series of graphical elements (or patterns) arranged in a 3×3 matrix, with the element in the lower right cell missing. The goal state is a completed series, and the test-taker's task is to identify the alternative that accomplishes this. To remove the distance between the initial state and the goal state, the test-taker must develop hypotheses about how the elements change across rows and/or down columns, and then test the hypotheses to determine which alternative is correct. In the example below, inspection of the top row suggests that each row must contain

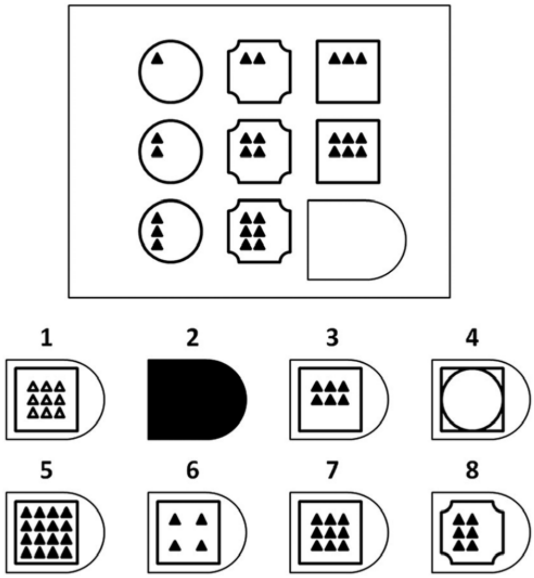

Figure 23.1 *Example of a Raven's-like problem (from Kunda et al., 2016). Reprinted with permission of Elsevier.*

one instance of each large shape, while inspection of the first column suggests that each column must contain elements of the same type. Inspection of the next rows and columns confirm these hypotheses, eliminating Alternatives 2 and 8. In turn, inspection of the small filled triangles indicates that the number of columns of triangles increases across each column, and the number of rows of triangles increases down each row. Only Alternative 7 contains both three rows and three columns of filled triangles; therefore, it is the only option that completes the series.

Even "low-level" cognitive tasks that are designed to isolate specific cognitive processes can be considered from a problem-solving perspective. Consider the Stroop task (Stroop, 1935). In the original version of Stroop, the research subject (or patient when the task is used for neuropsychological diagnosis) is given a sheet of paper with color names printed in conflicting colors (e.g., "red" printed in green), and instructed to read the words as quickly as possible. On another sheet, color names are printed in black, and again subjects are instructed to read the words as quickly as possible. As has now been replicated countless times, subjects are slower to read the words in the former condition. Stroop is often described as a test of "inhibition" (e.g., Miyake et al., 2000) – the ability to override or suppress an overlearned response, of the type that an American must exercise when driving on the left side of the road in the UK. At the same time, the task can be viewed as a problem-solving task in that it involves keeping in mind a goal to direct responses in the task ("read the word").

If all cognitive activity can be considered problem-solving, and if intelligence reflects the efficiency and effectiveness of all cognitive activity, then it follows that problem-solving ability is not just an aspect of intelligence – it is the essence of intelligence. In the next section, we briefly consider what is known about individual differences in intelligence, and then discuss evidence concerning the relationship between intelligence and complex problem-solving.

Intelligence as Problem-Solving Ability

As first documented by Spearman (1904) more than a century ago, one of the most replicated empirical findings in the entire field of psychology is that a person who performs well on one cognitive test (or task) will tend to perform well on all other cognitive tests. At the time, many psychologists subscribed to the view that the mind is a collection of a number of separate and independent abilities, or *faculties.* The assumption was that there was a faculty for practically every aspect of mental functioning – one for perception but another for memory, one for reasoning but another for intuition, and so on. To test this idea, Spearman computed correlations between grades in university courses for a sample of Columbia University students. As shown in Table 23.1, the correlations among the grades were uniformly positive. Therefore, there was a good chance that a student who did well in, say, Classics did well in all of the other subjects. Grades also correlated with a measure of pitch discrimination.

The correlations were all positive and relatively high. Pioneering the use of factor analysis in psychology, Spearman further demonstrated that each of the measures correlated very highly with the general factor implied by this "positive manifold," from a low of 0.72 for Music to a high of 0.99 for Classics. He concluded that performance on any given test of mental ability is a function of a factor reflecting something that the test shares in common with all of the others – a "common fundamental Function" (Spearman, 1904, p. 273) – along with an ability unique to that test. Spearman referred to the common factor as the general factor of intelligence, and it has since become known as "Spearman's *g*." In statistical terms, Spearman's *g* is the first principal factor in a factor analysis of cognitive ability measures; a test's "*g* loading" is the correlation between this factor and the measure of performance from the test (Jensen, 1999).

Table 23.1 *Correlation matrix (from Spearman, 1904)*

	Classics	French	English	Math	Pitch	Music
Classics	–					
French	0.83	–				
English	0.78	0.67	–			
Math	0.70	0.67	0.64	–		
Pitch	0.66	0.65	0.54	0.45	–	
Music	0.63	0.57	0.51	0.51	0.40	–

From Spearman (1904), p. 275.

Subsequent research focused on the question of whether *g* can be fractionated – that is, whether there is one intelligence or multiple intelligences. Cattell (1943) observed that certain measures of mental ability correlate more positively with each other than do others, leading him to propose that there are two factors of intelligence. "Fluid ability," he explained, "has the character of a purely general ability to discriminate and perceive relations between any fundaments, new or old," whereas "[c]rystallized ability consists of discriminatory habits long established in a particular field, originally through the operation of fluid ability, but not [sic] longer requiring insightful perception for their successful operation" (Cattell, 1943, p. 178). The theory of fluid ability (G*f*) and crystallized ability (G*c*) went on to have enormous impact in scientific thinking about intelligence.

It should be noted that Cattell's (1943) G*f*-G*c* distinction bore a remarkable similarity to Hebb's (1942) distinction between Intelligence A ("intellectual power") and Intelligence B ("intellectual products") (Hebb, 1942, p. 290). In fact, as Brown (2016) has recently documented, after attending a presentation given by Hebb at the 1941 American Psychiatric Association (APA) meeting, it appears that Cattell adopted Hebb's theory, simply relabeling Intelligence A and Intelligence B. As letters between Cattell and Hebb, and between Cattell and Hebb's department head, George Humphrey, reveal, Cattell grudgingly acknowledged Hebb's influence on his theorizing in his 1943 *Psychological Bulletin* article (Cattell, 1943). However, in later articles (e.g., Cattell, 1963), Cattell took full credit for G*f*-G*c* theory without acknowledging Hebb. At various points, Cattell also claimed that he had proposed the fluid-crystallized distinction either during the 1941 APA meeting or before it, but there is no record of him doing so.

This matter aside, later factor-analytic research with Horn supported the distinction between G*f* and G*c* factors. G*f* was found to have its highest loadings on nonverbal reasoning measures such as Raven's, and G*c* its highest loadings on measures reflecting acculturated learning, including vocabulary and general information. Subsequently, in a landmark project, Carroll (1993) compiled and reanalyzed the results of over 460 factor-analytic studies, and found that mental abilities can be arranged into a hierarchy that includes three levels, or "strata." At the highest level of the hierarchy (stratum I) is Spearman's *g*, representing what all tests of cognitive ability share in common; at the next level (stratum II) are "broad" cognitive abilities, including G*f* and G*c*; and at the lowest level (stratum III) are "narrow" cognitive abilities, representing demands unique to particular types of tests.

There has been, and continues to be, vigorous debate in the intelligence literature about the structure of cognitive abilities. McGrew and colleagues (see McGrew, 2005, 2009) merged Cattell and Horn's G*f*-G*c* model and Carroll's three-stratum model into the Cattell-Horn-Carroll (CHC) model. Meanwhile, extending Vernon's (1965) verbal-perceptual model, Johnson and Bouchard (2005) presented evidence to indicate that a model with *verbal, perceptual*, and *image rotation* factors was better fitting than either G*f*-G*c* or CHC models. Thinking about intelligence more broadly, Sternberg (1985b) proposed distinctions among three intelligences: *practical, creative*, and *analytical*.

The differences among these models are important, but from a historical perspective, it also worth pointing out a salient similarity, which reflects a scientific consensus about the nature of intelligence that has emerged through decades of intensive research and debate. All these models assume that there are far fewer factors underlying variation in scores on tests of mental ability than there are tests of mental ability. In other words, these models assume that the human intellect comprises some number of relatively general factors, and cannot be adequately (or easily) understood as a collection of independent modules as a proponent of the faculty view of the mind would have argued in 1900, stimulus-response associations as a radical behaviorist would have argued in 1950, or highly specific skills as a neobehaviorist might try to argue today. This assumption, obvious only with the benefit of hindsight, is the foundation for the whole enterprise of developing tests to measure intelligence and using scores on these tests for practical purposes such as predicting job performance. The perspective that we advance here is that, whatever factors a theory of intelligence posits, they can be considered from a problem-solving perspective.

Intelligence and Complex Problem-Solving

A different tradition of research on intelligence (the European perspective) has focused on the question of how intelligence relates to complex problem-solving (CPS), which Buchner defined as "the successful interaction with task environments that are dynamic (i.e., change as a function of user's intervention and/or as a function of time) and in which some, if not all, of the environment's regularities can only be revealed by successful exploration and integration of the information gained in that process" (Frensch & Funke, 1995, p. 14). CPS is assessed using "microworlds" that may be simulations or abstractions of real-world tasks. The best-known example of a CPS microworld is Tailorshop (see, e.g., Danner et al., 2011). The subject's task is to manage, over a period of twelve "virtual" months, a garment-manufacturing company, making decisions about hiring, pricing, advertising, and so on. Changes in these "input" variables during one cycle (month) lead to changes in "output" variables in the next cycle, such as company value, monthly sales, and customer demand. An example of a more abstract CPS task is a paradigm called MicroDYN (see Figure 23.2). The MicroDYN approach consists of a series of short (e.g., 5-minute), moderately complex tasks in which subjects must explore a novel task environment, generate knowledge about the relationships between input and output variables (e.g., how physical training methods affect outcomes for a handball team, such as power of throw), and then apply this knowledge to reach a specified goal (e.g., to combine training methods to achieve a certain level of the outcomes). Unlike in simulations of real-world tasks like Tailorshop, prior knowledge has little impact on performance in MicroDYN tasks because variable labels are either fictitious or lacking deep semantic meaning (e.g., feed imaginary animals food A, B, or C; Greiff, Wüstenberg, & Funke, 2012).

A long-standing question in this research is whether intelligence and CPS are the same or different constructs. A comprehensive review of this literature is beyond the

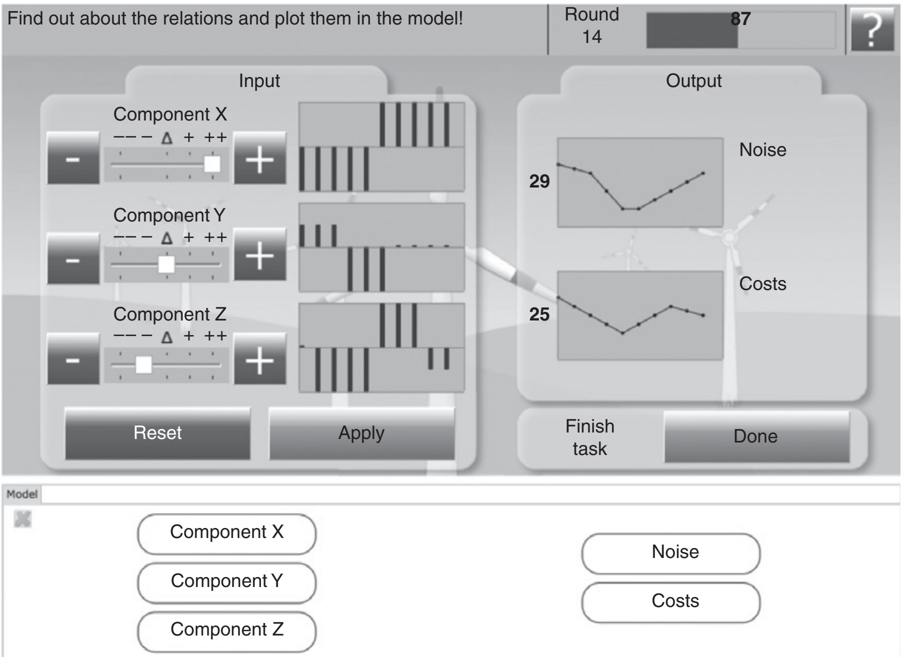

Figure 23.2 *Screenshot from MicroDYN (from wwwen.uni.lu/media/images/microdyn). Reprinted with permission of Dr. Samuel Greiff.*

scope of this chapter, but suffice it to say that the two are significantly related. Results of a recent study by Kretzschmar and colleagues (2016) are illustrative. A sample of 227 German university students completed the Berlin Intelligence Structure test (an IQ test) and two CPS tasks (MicroDYN and MicroFIN). As is typical, *g* loadings for the measures were all positive, ranging from 0.26 to 0.83. The loadings for MicroDYN and MicroFIN were in the middle of this range (0.60 and 0.55, respectively; average *g* loading = 0.54). And although a model with a CPS factor improved model fit, this factor could be interpreted as a method factor (i.e., the MicroDYN and MicroFIN tasks use very similar procedures, and are more similar to each other than to either of the other cognitive ability tests). Finally, CPS improved prediction of an external criterion (grade point average, GPA) when intelligence was narrowly defined as figural reasoning (see also Greiff et al., 2013a; Greiff et al., 2013b; Wüstenberg, Greiff, & Funke, 2012). However, CPS did not improve prediction of GPA above and beyond a *g*-factor. Similarly, in a sample of 560 Luxembourgish high school students, Sonnleitner and colleagues (2013) found that the average correlation between a CPS factor and academic achievement variables dropped considerably (avg. r = 0.39 to 0.15) after G*f* (comprising multiple measures of reasoning ability) was statistically partialled from CPS.

A recent meta-analysis summarizing these and other findings (total studies = 47) estimated the overall correlation of CPS with intelligence at 0.43 (Stadler et al., 2015). The correlation was stronger for reasoning ability (indexing G*f*) than for overall measures of intelligence (indexing *g*); among different tasks used to measure CPS, it was strongest for those that involve managing multiple complex systems (avg. $r = 0.72$, after correction for unreliability). Correlations were less than 1.0 after corrections for unreliability, leading the authors to conclude that CPS and intelligence are distinct constructs. This may be, but an even stronger test of the distinction would be to test the correlation between a psychometric *g*-factor extracted from a diverse set of cognitive ability measures (reasoning, spatial visualization, working memory, comprehension, processing speed, etc.) and a *g*-factor extracted from diverse CPS paradigms (e.g., MicroDYN, Tailorshop, Genetics Lab; see Greiff et al., 2013, for a study that at least did the latter). What is already known from research in this area suggests that these *g*-factors would correlate very highly, and that a CPS *g*-factor would account for relatively small amounts of variance in external criteria above and beyond psychometric *g* (Sonnleitner et al., 2013).

Underpinnings of Intelligence as Problem-Solving Ability

To sum up, in our view, intelligence and problem-solving ability are closely connected, at both theoretical and empirical levels. There is a general capacity for solving novel problems that maps onto G*f*/Intelligence A, and there is knowledge acquired through the exercise of this capacity that maps onto G*c*/Intelligence B. Other labels for what are essentially the same broad factors include Ackerman's (1996) *intelligence-as-process* and *intelligence-as-product,* Baltes' (1987) *cognitive mechanics* and *cognitive pragmatics,* and Salthouse's (2000) *process* and *product* cognition.

But, at the level of the cognitive system, what about a person who is intelligent (and thus an effective problem-solver) differs from a person who is less intelligent (and thus a less effective problem-solver)? There has been a great deal of enthusiasm for the idea that G*f* can be equated with *working memory capacity* (WMC). As Baddeley and Hitch (1974) conceived of it nearly fifty years ago, working memory is a limited-capacity system for both storing and processing information in the service of complex cognition. Based on this model, as the first test of WMC, Daneman and Carpenter (1980) introduced the reading span paradigm. The participant reads a series of sentences (the processing task) while remembering the last word in each for later recall (the storage task), and then is cued to recall the words. Reading span is the number of sentences that the participant can read while maintaining perfect recall of the words. Subsequently, Engle and colleagues introduced operation span (Turner & Engle, 1989). The participant solves a series of arithmetic equations, remembering a word that follows each for later recall.

Beginning around 1990, there were numerous reports of strong positive correlations between WMC and G*f* (see Conway & Kovacs, 2013, for a review). In the first large-scale study of the relationship between WMC and G*f*, Kyllonen and Christal (1990) found strong correlations (> 0.90) between latent variables representing reasoning ability and WMC and concluded that "reasoning ability is little more than working memory capacity" (p. 389). Even more boldly, Kyllonen (2002) stated, "we have our answer to the question of what g is. It is working memory capacity" (p. 433). In a similar vein, Engle (2002) noted that WMC "is at least related to, maybe isomorphic to, general fluid intelligence" (p. 22). At least in the minds of some, the question of what intelligence is, beyond the variance common to a collection of mental ability measures, seemed to be settled.

However, later research revealed that the relationship between WMC and G*f* was weaker than initially thought – much weaker. In a meta-analysis, Ackerman, Beier, and Boyle (2005) found a correlation of 0.50 between latent variables representing *g* and WMC, indicating that only a quarter of the variance was shared between the factors, and in a reanalysis of twelve studies, Kane, Hambrick, and Conway (2005) found an average correlation of 0.72 between latent variables representing G*f* and WMC. Furthermore, analyzing item-level data, Unsworth and Engle (2005) found that the correlation between WMC and performance on Raven's Progressive Matrices was as strong for simple items as for complex ones. Likewise, Salthouse and Pink (2008) found that performance on small set sizes in complex working memory span tasks (e.g., operation span) correlated as highly with G*f* as performance on larger set sizes. These findings were problematic for any explanation of the relationship between performance on working memory tasks and G*f* in terms of the amount of information that can be simultaneously stored and processed.

As a theoretical critique, others pointed out that measures of WMC are at least as complex as the measures of intelligence that they purportedly explained (Deary, 2000). The point was that, like any other cognitive measure, no measure of WMC is "process pure" in the sense that it captures only the intended construct. As Salthouse and Pink (2008) commented, "Because some of the WM assessments closely resemble tests of reasoning and higher order cognition, it may not be reasonable to claim that the WM construct is theoretically more tractable or less opaque than are intelligence constructs, given the fact that it is operationalized in so many different ways that appear to have little conceptual integration" (p. 364). Ackerman, Beier, and Boyle (2002) described working memory as a "promiscuous" variable: one that correlates with everything.

Thus, enthusiasm about the hypothesis that G*f* is "little more than" WMC has waned. What may be regarded as the death knell for this view was recently sounded by one of its most prominent early supporters. As already mentioned, Engle (2002) argued that WMC tasks measure a factor that may be isomorphic to G*f*. However, in an update of this article, Engle (2018) explained, "One of the things I argued in the 2002 article was that individual differences in WMC possibly play a causal role in fluid intelligence. I based that argument on the strong relationship, on the order of .6 to .8, between WMC and fluid intelligence at the construct level. I now think that argument was wrong" (p. 192). Reviewing recent work from his lab, he explained that effects of WMC and G*f* on various outcomes are dissociable, noting that some effects that he and his colleagues originally attributed to WMC can actually be explained by G*f*. For example, Rosen and Engle (1998) found that high-span subjects produced more unique animal names in a fluency task than did low-span subjects, and argued from this finding that WMC is related to the ability to suppress intrusive thoughts (i.e., animal names that had already been produced). However, subsequent research using a more powerful statistical approach (structural equation modeling with large samples) showed that effects of WMC on fluency were entirely due to G*f* (Shipstead, Harrison, & Engle, 2016).

The Role of Placekeeping Ability

If WMC does not explain individual differences in G*f*, what does? In our own attempt to shed some light on this question, we have considered the role of a theoretical construct that we call *placekeeping*: the ability to perform a sequence of steps in a particular order, without skipping or repeating steps (Hambrick, Altmann, & Burgoyne, 2018). One reason to think that placekeeping ability is related to G*f* is that solving complex problems, at least of the well-structured variety, depends on a kind of linear thinking. Newell and Simon (1972; Newell, 1990) characterized problem-solving in terms of a search process in which the problem-solver applies sequences of operators to transform mental problem states, and periodically sets, suspends, and resumes goals organized in a hierarchical mental structure. For such processing to lead to solutions efficiently, the system must be able to keep its place in sequences of operators and within hierarchical goal structures. Skipping an element could mean missing a path to a solution, and repeatedly evaluating a failed path is inefficient and could also lead to missing a solution if solution time is limited. Consistent with this analysis, using a computational cognitive model, Carpenter, Just, and Shell (1990) demonstrated that successful goal management during problem-solving was associated with better performance on Raven's Progressive Matrices.

In a recent series of studies, we have tested the relationship between intelligence and placekeeping ability. The placekeeping task we developed is defined by the acronym UNRAVEL. Each letter in the acronym identifies a step in a looping procedure, and the letter sequence stipulates the order in which the steps are to be performed. That is, the U step is performed first, the N step second, the R step third, and so forth, and the participant returns to the U step following the L step. The participant is interrupted at random points and must perform a transcription task, before resuming UNRAVEL with the next step. A sample stimulus is shown in Figure 23.3. Each sample stimulus includes two characters; of these characters, one is a letter and one a digit, one is presented either underlined or italicized, one is colored red or yellow, and one is located outside of a gray box. Each step requires a two-alternative forced choice (i.e., a keypress) related to one feature of the stimulus, and the letter of the step mnemonically relates to the choice rule: The U step involves deciding whether the formatted character is underlined or italicized, the N step whether the letter is near to or far from the start of the alphabet, the R step whether the colored character is red or yellow, the A step whether the character outside the box is above or below, the V step whether the letter is a vowel or a consonant, the E step whether the digit is even or odd, and the L step whether the digit is less or greater than five. The task tests placekeeping because the stimulus provides no information about what step to perform next. Instead, the participant must remember their place in the procedure, which is especially challenging immediately after an interruption by the transcription typing task.

In recent work, we have examined the contribution of placekeeping to individual differences in G*f*. In a study using 132 undergraduate students (Hambrick &

(a) Sample stimuli for UNRAVEL task:

(b) Choice rules and candidate responses for UNRAVEL task, and responses to the stimuli in (a)

Step	Candidate responses		Choice rules	Responses to Stimulus 1	Responses to Stimulus 2
U	u	i	character is Underlined or in Italics	u	i
N	n	f	letter is Near to or Far from start of alphabet	n	f
R	r	y	character is Red or Yellow	r	y
A	a	b	character is Above or Below the box	a	b
V	v	c	letter is Vowel or Consonant	v	c
E	e	o	digit is Even or Odd	o	e
L	l	m	digit is Less than or More than 5	l	l

(c) Sample stimulus for the transcription task:

Figure 23.3 *Sample stimuli, rules, and interrupting task from UNRAVEL task (from Hambrick & Altmann, 2015). Reprinted with permission of Springer.*

Altmann, 2015), we found that placekeeping ability (as indexed by UNRAVEL error rate) accounted for 20 percent of the variance in scores on Raven's Progressive Matrices. Furthermore, this contribution was reduced only slightly (20% to 18%) after controlling for WMC and measures of two other "executive functioning" factors (task switching and multitasking). Furthermore, in a structural equation model, latent variables representing placekeeping ability (again reflecting the UNRAVEL error rate) and Raven's performance correlated strongly ($r = -0.69$), in the predicted direction of higher Raven's score for participants with lower error rate in UNRAVEL.

In a more recent study (Burgoyne, Hambrick, & Altmann, in press), we had participants complete two tests of placekeeping, two tests of G*f*, and two tests of WMC. In regression analyses, placekeeping ability factors (error rate and response time) accounted for 12 percent of the variance in G*f* above and beyond WMC. By contrast, WMC accounted for only 2 percent of the variance in G*f* above and beyond these placekeeping factors. In a structural equation model, the placekeeping factors had stronger effects on G*f* than did WMC, and dropping the latter, placekeeping ability accounted for 70 percent of the variance. Because it is correlational, this finding does not establish that placekeeping is a cause of variation in Raven's scores (or G*f*). It is, however, consistent with that possibility, providing a motivation to conduct experiments to test causal hypotheses.

We make no claim that UNRAVEL is a "magic bullet" task for research on the underpinnings of intelligence, or that placekeeping ability is the only factor underlying individual differences in G*f*. We do not believe that there is any single construct that can explain something as complex as intelligence. No less so in intelligence research than any other area of psychological research, single-variable explanations are nearly always wrong, or at least incomplete, even if they are seductive. However, we do think that placekeeping ability may be one piece of the intelligence puzzle. Considering intelligence from a problem-solving perspective, a goal for future research in our labs is to conduct experiments to test hypotheses about how placekeeping constrains problem-solving behavior in intelligence tests such as Raven's Progressive Matrices. This research promises to shed light on mechanisms underlying human intelligence, particularly as manifested in the type of linear thinking that is involved in solving well-structured problems.

What Does Intelligence as Problem-Solving Ability Predict?

Life is an unceasing series of problems, from meeting the most basic requirements for survival, to finding and keeping employment, to fulfilling loftier aims such as making the world better for future generations. Even the task of remaining among the living (as it were) can be seen as a problem, especially when there are many ways to lose that status (getting hit by a truck, accidentally ingesting poison, wrecking one's health through a profligate lifestyle). Is intelligence predictive of success in life's problem-solving tasks? The answer is yes, even if the reasons are not yet fully understood. Next, we consider evidence for the relationship between intelligence and three outcomes that can be considered indexes of real-life problem-solving skill: mortality, job performance, and academic achievement.

Mortality

Intelligent people tend to live longer than less intelligent people. This relationship between intelligence and mortality has been documented in research by Deary and colleagues using data from the Scottish Mental Surveys. In 1932, the Scottish government administered an intelligence test to nearly all eleven-year-old children attending school on a single day. More than sixty years later, after finding the raw data collecting dust in a University of Edinburgh building, Deary and Whalley (2001) identified who from the cohort living in the city of Aberdeen was still alive, at age seventy-six. The results revealed that a 1 standard deviation advantage (15 IQ points) was associated with a 21 percent greater chance of survival. For example, a person with an IQ of 100 (the average for the general population) was 21 percent more likely to be alive at age seventy-six than a person with an IQ of 85. IQ also predicts morbidity – becoming ill due to diseases such as cancer and heart disease (Gottfredson & Deary, 2004).

In short, as has now been replicated in upward of twenty longitudinal studies from around the world, more intelligent people live longer, healthier lives than less intelligent people. One possible explanation for this finding is that intelligence is confounded with another variable that correlates with mortality: socioeconomic status. That is, intelligence and mortality may correlate because wealthier people have the means to develop their intelligence (through education) and stay in good health (through health care). There is some evidence consistent with this third-variable explanation. At the same time, socioeconomic factors do not appear to completely account for the intelligence-mortality/morbidity relationship. For example, Hart and colleagues (2003) linked IQ scores for over 900 of the participants from the 1932 study to those participants' responses on a national health survey conducted in the early 1970s. They found that statistically controlling for economic class and a measure of "deprivation" reflecting unemployment, overcrowding, and other adverse living conditions accounted for only about 30 percent of the IQ-mortality correlation.

Another possible explanation for the intelligence-mortality correlation might be called the *problem-solving hypothesis*: Day in, day out, more intelligent people are more effective in solving life's problems than less intelligent people, using knowledge that they have acquired from various sources (e.g., reading, listening to their physician) to keep in good health and stay alive. Consistent with this hypothesis, in the Scottish data, there was no relationship between IQ and smoking behavior in the 1930s and 1940s, when the health risks of smoking were unknown, but after that, people with higher IQs were more likely to quit smoking (Gottfredson & Deary, 2004). There is other evidence consistent with this hypothesis, as well. For example, a meta-analysis showed that people high in cognitive ability were less likely to be involved in vehicular accidents than people lower in cognitive ability (Arthur, Barret, & Alexander, 1991). Other research implicates lower levels of *executive functioning* (EF) in risky driving and vehicular accidents (Walshe et al., 2017). An umbrella term for cognitive operations (e.g., inhibition, updating, set shifting) that are presumed to underpin goal-directed behavior, EF correlates very highly with G*f* (McCabe et al., 2010). Lower cognitive ability has been found to be associated with poor supervisor ratings of employee safety behaviors, ranging from work-related accidents to distractibility when performing dangerous tasks (Postlethwaite et al., 2009).

Job Performance

As industrial-organizational psychologists have established, *g* is also the best-known predictor of acquisition of job knowledge during training and subsequent job performance (Schmidt & Hunter, 2004). Validity coefficients for *g* tend to be higher for more complex jobs than less complex jobs, but are nearly always positive and both statistically and practically significant. This is not to say that *g* is a perfect predictor of job performance – far from it. Across a wide range of occupations, the average validity coefficient for *g* is around 0.50 (see Schmidt & Hunter, 1998, 2004). This means that, on average, *g* accounts for about 25 percent of the variance in job performance, leaving the rest unexplained and potentially explainable by other factors. It is, however, to say that *g* is a better predictor of job performance than measures of other factors, such as personality, interests, and job interview ratings. It is also to say that intelligence tests are practically useful. As utility analyses demonstrate, use of tests with even modest validity (e.g., correlations with job performance in the 0.20–0.40 range) for personnel selection can substantially improve prediction of job performance, translating into substantial savings in terms of decreased job training time and increased productivity (Hunter & Schmidt, 1996).

It is sometimes argued that *g* (general cognitive ability) predicts job performance only early in training within a domain (e.g., a job), after which domain-specific factors (knowledge, skills, and strategies) enable a person to circumvent reliance on domain-general abilities. The chief proponent of this *circumvention-of-limits hypothesis* is the expertise researcher K. Anders Ericsson. In *Peak: Secrets from the New Science of Expertise,* Ericsson and Pool (2016) claimed, "While people with certain innate characteristics . . . may have an advantage when first learning a skill, that advantage gets smaller over time, and eventually the amount and quality of practice take on a much larger role in determining how skilled a person becomes" (p. 233). Similarly, citing his own review of the evidence (Ericsson, 2014), Ericsson recently wrote that "the influence of general abilities, such as IQ, is greater on performance of beginners but virtually disappears for individual differences among expert performers" (Ericsson, 2018a, p. 708), and that "traditional tests of intelligence and IQ are not predictive of individual differences in attained performance among skilled performers" (Ericsson, 2018b, p. 97).

In reality, evidence for this claim is weak, even though it might seem extensive. There are several major problems with Ericsson's reviews of the relevant evidence (Ericsson, 2013, 2014, 2018a, 2018b; Ericsson & Moxley, 2013). First, in some cases, Ericsson makes inferences that are not licensed by the usual conventions of statistical inference in psychological research. In particular, for some studies, he cites as support for his view the finding that an ability-performance correlation was statistically significant in a lower-skill group but not in a higher-skill group, even though the correlations are not significantly different from each other. For example, in multiple reviews, as support for his view, he cites Ruthsatz and colleagues' (2008) finding that a measure of intelligence (Raven's score) correlated with musical performance in high school band members ($r = 0.25$, $p < 0.01$, $N = 178$) but not in more skilled groups of university music majors ($r = 0.24$, $p > 0.05$, $N = 19$) and music institute students ($r = 0.12$, $p > 0.05$, $N = 64$). However, these correlations are not significantly different from each other (all tests of differences in *r*s are non-significant, *z*s < 1). Thus, following the norm of testing for differences in correlations, Ruthsatz and colleagues' results actually *fail* to support Ericsson's claim of a diminishing ability-performance correlation with increasing skill.

The problem here can be made obvious by imagining a situation in which the correlation between two variables (e.g., IQ and performance) just exceeds the threshold for statistical significance in one group and just misses that threshold in another group – for example, with a sample size of fifty per group, $r = 0.280$ ($p = 0.049$) in one group and $r = 0.278$ ($p = 0.051$) in another group. Obviously, no psychological theory can purport to predict such a small difference between correlations (and thus whether one correlation will be significant and the other nonsignificant). Rather, in testing a theory that predicts differential relationships, as Ericsson's does, the question that must be answered is whether the correlations differ significantly from each other. Statistical tests have been developed for the very purpose of answering this question, and have long been in use in the behavioral sciences (e.g., the z test for difference between correlations from independent samples, moderated regression analysis testing for Group × Predictor interactions; see Cohen, 1988; Cohen et al., 2003; Hays, 1988).

In other cases, Ericsson's (2014) interpretations of evidence do not stand to reason. For example, he cites a report in the German magazine *Der Spiegel* that former chess world champion Garry Kasparov's IQ was estimated at 120 based on his score on Raven's Progressive Matrices, and notes that this score is "very close to the average of all chess players . . . thus not very predictive of world-class chess performance" (Ericsson, 2014, p. 87). However, one case does not a correlation make: Even if this estimate of Kasparov's IQ is accepted as valid, one cannot make an inference about the strength of a predictive relationship between two variables (i.e., a correlation) based on a single case. To wit, if other world-class chess players (e.g., Bobby Fischer, Magnus Carlsen) had considerably higher IQs, then IQ could still be *highly* predictive of world-class chess performance.

In still other cases, Ericsson is selective in reporting of evidence. For example, citing Schmidt and Hunter's (2004) review of the job performance literature, Ericsson (2014) explains, "The expert-performance approach proposes that performance on tests of general cognitive ability [and performance] will be correlated for beginners" (p. 84). However, he fails to mention that, in this same review, Schmidt and Hunter (2004) noted that this is true for non-beginners, as well: "One might hypothesize that the validity of GMA [general mental ability] declines over time as workers obtain more job experience. However, research does not support this hypothesis" (p. 167). Ericsson also seems to overlook relevant evidence. For example, in the surgical domain, he misses Gallagher and colleagues' (2003) finding that scores on a test of visuospatial ability correlated significantly and similarly with performance on a laparoscopic laboratory cutting task in novices (rs = 0.50 and 0.50, Ns = 48 and 32) and in experienced surgeons ($r = 0.54$, $N = 18$). No review is perfect – it is easy to miss studies. All the same, failing to review evidence contrary to a hypothesis leads to a biased portrayal of the strength of the evidence for that hypothesis.

Finally, Ericsson makes material errors in his reviews. For example, Ericsson (2013) claimed that "Kopiez and Lee (2006) found that for musicians with lower sight-reading skill there was a correlation with their working memory. For musicians with a higher level of sight-reading skill there was no significant relation between their performance and their working memory" (p. 236). This is an error of commission: Kopiez and Lee (2006) reported no such finding. In fact, they did not report any analyses comparing the correlation between working memory and sight-reading performance in groups representing lower versus higher levels of sight-reading skill. As another example, referring to a subsequent report of data from this study of sight-reading, Ericsson (2018a) noted that "Kopiez and Lee (2008) found that speed of alternating finger movements [music-specific speed trilling] and amount of accumulated sight-reading experience were the only significant predictors of sight-reading performance, but not working memory" (p. 707). This is an error of omission: Sight-reading performance was also significantly predicted by a nonmusic measure of information-processing speed (i.e., number combination, $r = -0.44$, $p = 0.001$; see Kopiez & Lee, 2008, table 2), indicating faster processing for more skilled sight-readers. These errors are not inconsequential typos; they lead to conclusions that contradict those Ericsson advances (for further examples of errors in Ericsson's writings, see Hambrick et al., 2014; Macnamara, Hambrick, & Moreau, 2016).

We carried out our own review of evidence relevant to the circumvention-of-limits hypothesis (Hambrick, Burgoyne, & Oswald, in press), conducting systematic searches for relevant articles in the literature on expertise in five domains (games, music, science, sports, surgery/medicine, and aviation), as well as the literature on job performance. Altogether, we searched approximately 1,300 documents. The findings can be summarized briefly. On balance, evidence from the expertise literature does not support the circumvention-of-limits hypothesis. To be exact, three of fifteen studies provide support for the hypothesis, either in the form of significantly different ability-performance correlations across skill groups or significant ability × skill interactions on performance. What might be regarded as the strongest evidence comes from one of our own meta-analyses (Burgoyne et al., 2016). We found that the correlation between G*f* (as measured by tests of reasoning ability) and chess expertise was significantly higher for less-skilled chess players than for more-skilled players. However, as we urged, this finding must be interpreted cautiously, because the measure of chess skill was highly confounded with age (i.e., the more-skilled players were adults, the less-skilled players were children).

A more consistent picture emerged in the review of evidence from the job performance literature. Ability-performance correlations may decrease in relatively simple lab tasks, in particular those with consistent demands (Ackerman, 1988; see also Henry & Hulin, 1987). However, even after an extensive amount of job experience, general cognitive ability remains a statistically and practically significant predictor of actual job performance (see also Reeve & Bonaccio, 2011). Some of the most compelling evidence for this conclusion comes from a reanalysis of data from the Joint-Service Job Performance Measurement/Enlistment (JPM) Standards Project, a large study initiated in 1980 by the US Department of Defense to develop measures of military job performance (see Hambrick et al., in press; see Wigdor & Green, 1991, for further description of this project). The JPM data set includes thirty-one jobs and a total sample size of 10,088 military personnel; the measure of general cognitive ability was the

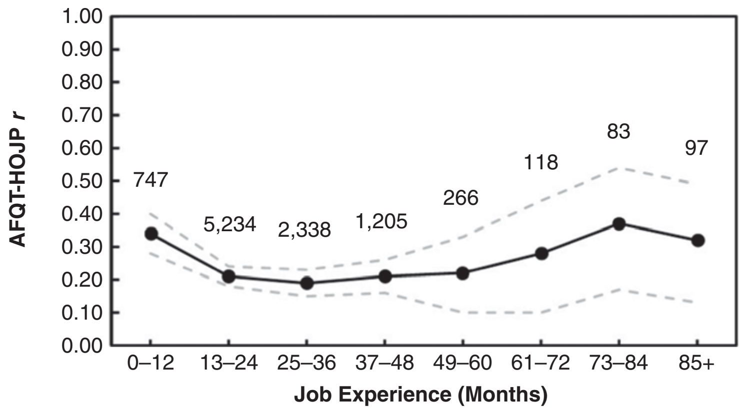

Figure 23.4 *Correlation between intelligence test score (AFQT) and hands-on job performance (HOJP) at different levels of job experience from JPM Standard Project. Total* N = *10,088 (from Hambrick, Burgoyne, & Oswald, in press). Reprinted with permission of Oxford University Press.*

Armed Forces Qualifying Test (AFQT) score, and job performance was measured with hands-on job performance (HOJP) tests for the different jobs. As shown in Figure 23.4, the AFQT-HOJP correlation decreases from the first year to the second, stabilizes, and then, if anything, increases. The overall picture to emerge from this and other large-scale studies (e.g., Schmidt et al., 1988; Farrell & McDaniel, 2001) is that general cognitive ability remains a significant predictor of job performance, even after extensive job experience, and even if validity drops initially.

Academic Achievement

Intelligence is also a strong predictor of academic achievement, which can be considered an index of problem-solving ability, at least within the confines of the classroom. (There are no doubt people who excel in work and other realms of life despite less-than-stellar academic records.) Using a sample of over 70,000 English schoolchildren, Deary and colleagues estimated the relationship between IQ at age eleven and scores on national examinations in twenty-five topics at age sixteen (Deary et al., 2007). The correlation between latent variables representing general intelligence and educational achievement was 0.81. Intelligence predicts later academic performance, as well. Most notably, scores on college admissions exams (the SAT and ACT) – which correlate highly with independent assessments of intelligence (Koenig, Frey, & Detterman, 2008) and were described by Gardner (1999) as "thinly disguised intelligence tests" (p. 69) – predict not only first-year grade point average (GPA), but overall GPA (Kuncel, Hezlett, & Ones, 2004). Summarizing evidence from their own and others' research in the *Wall Street Journal*, Kuncel and Sackett (2018) wrote "Standardized tests are just tools – very effective tools – but they provide invaluable information to admissions offices. They identify those students who need help catching up with fundamental skills and those who are ready to tackle advanced material and rapidly accelerate in their learning."

As evidence from the landmark Study of Mathematically Precocious Youth (SMPY) indicates, scores on college admissions tests also predict outcomes reflecting accomplishments beyond the college years. As part of a planned fifty-year study, the SAT was administered to intellectually gifted youth by age thirteen, and the roughly 2,300 scoring in the top 1 percent were tracked into adulthood. (Given that the SAT was administered at this unusually young age, ceiling effects were avoided, and there was a wide range of scores even in the top 1 percent, for example, 390 to 800 for SAT Math.) Lubinski (2009) found that higher overall SAT score (an index of *g*) was associated with higher levels of accomplishment. For example, compared to individuals in the 99.1 percentile, those in the 99.9 percentile (the profoundly gifted) were 3.56 times more likely to have earned a doctorate, 4.97 times more likely to have published in a STEM journal, 3.01 times more likely to have been awarded a patent, and 2.31 times more likely to have an income at or above the 95th percentile (Ferriman Robertson et al., 2010).

One possible explanation for such results is that people who score well on standardized tests are admitted to better colleges and universities, which creates a halo effect that makes it easier for these people to get into top graduate programs (e.g., Princeton or Harvard), which in turn creates a halo effect that makes it easier for them to, say, publish in a STEM journal. However, this does not appear to explain the relationship between SAT scores and later accomplishments in the SMPY. For example, Park, Lubinski, and Benbow (2008) found that, among individuals who had earned a doctorate, prestige of doctoral institution had very little impact on the relationship between SAT math scores and publication in STEM journals: The odds ratio was 4.04 for graduates of a non-top-fifteen institution ($N = 766$) and 3.52 for graduates of a top-fifteen institution ($N = 240$). Other mediating variables may turn out to explain the relationship between SAT and outcomes, but the simplest explanation remains the most direct, which is that ability influences outcomes.

Can Intelligence as Problem-Solving Ability be Improved through Training?

If intelligence is important for real-world problem-solving, then can it be increased through training, leading to improvements in people's lives? Psychologists have been interested in this possibility for as long as they have studied intelligence. Over a century ago, Alfred Binet, developer of the first standardized intelligence test, envisioned a system of "mental orthopedics" for increasing intelligence. "With practice, training, and above all, method," Binet wrote, "we manage to increase our attention, our memory, our judgement and literally to become more intelligent than we were before" (Binet, 1909/1975, p. 106–107).

Nevertheless, efforts to increase intelligence through training have met with little success. Beginning in the 1970s, a number of longitudinal studies were launched to examine the effects of intensive educational interventions on intelligence. In one of the best-known studies, the Abecedarian Early Intervention Project (Campbell et al., 2001), children from low-income families received an intensive educational intervention from infancy to age five that included educational games designed to enhance cognitive functioning, while children assigned to a control group received social services, healthcare, and nutritional supplements. Increases in intelligence were modest: At the end of the study, children in the treatment group showed an advantage of 6 IQ points (about 0.40 standard deviations) over a control group.

There was renewed excitement about the idea of training intelligence in the 2000s. In an article published in the *Proceedings of the National Academy of Sciences,* Jaeggi and colleagues reported large gains in G*f* for a sample of young adults following computerized cognitive training (Jaeggi et al., 2008). After completing a pretest of reasoning ability, participants were assigned to either a control group, or to a training group in which they received eight, twelve, seventeen, or nineteen sessions of training in a "dual n-back" task requiring simultaneous monitoring of two streams of information (one auditory and one visual). Finally, at posttest, all participants took a different version of the reasoning test. Jaeggi and colleagues reported that there was a greater gain in G*f* from pretest to posttest for the training groups than for the control group, as well as a dosage-dependent relationship (more training, greater gain in G*f*). They concluded that the "finding that cognitive training can improve *Gf* is a landmark result because this form of intelligence has been claimed to be largely immutable" (p. 6832).

The study made an immediate impact, both in the popular press and the scientific literature. *Discover* magazine called Jaeggi and colleagues' (2008) findings one of the top 100 scientific discoveries of 2008, and Sternberg (2008) commented that the study seemed "to resolve the debate over whether fluid intelligence is, in at least some meaningful measure, trainable" (p. 6791). Within a year of publication, the study had already been cited nearly 100 times. The study was also touted as evidence for the effectiveness of brain training by the company Lumosity, which introduced a gamified version of the dual n-back task. Within a few years, advertising on National Public Radio and network television, Lumosity boasted fifty million subscribers. The claim was that brain training can improve cognitive functioning, with benefits for real-world performance.

However, as subsequently reported by Redick and colleagues (2013), Jaeggi and colleagues' (2008) study had major flaws. The control group was "passive" or "no-contact," meaning that the control subjects had no contact with experimenters between pretest and posttest. Therefore, it was possible that the training groups improved because they expected improvement and tried harder on the intelligence test at posttest. Furthermore, the study was not a single experiment in which participants were randomly assigned to conditions that differed only in amount of training; Instead, the results were from different experiments in which the procedures varied in other respects, making it difficult to interpret the reported results. For example, the reasoning test differed across the training groups, with the eight-session training group receiving an eighteen-item version of Raven's, but the other groups receiving the twenty-nine-item Bochumer Matrizentest (BOMAT) test, and with a twenty-minute time limit for the nineteen-session group and a ten-minute time limit for the twelve- and seventeen-session groups. Finally, there was no, or even negative, transfer for other measures not included in the report of the study (e.g., for the 19-session group, no transfer to visuospatial span, and negative transfer to digit-symbol substitution).

What's more, the magnitude of the reported training gain in G*f* seemed larger than possible. The training groups received an average of six hours of dual n-back training, and the difference in the gain in reasoning performance from pretest to posttest for the training groups compared to the control group was 0.40 standard deviations. IQ tests typically have a mean of 100 with a standard deviation of 15. Thus, in terms of IQ points, Jaeggi and colleagues' (2008) results implied that the G*f* of subjects in the training groups increased an average of six points in six hours – a point an hour. This was roughly the same gain in IQ for children in the treatment group of the Abecedarian study after five years of intensive intervention.

Given these problems, Redick and colleagues (2013) attempted to replicate Jaeggi and colleagues' (2008) findings. Participants completed seventeen different cognitive ability tests, including eight tests of G*f*. They were then assigned to a treatment group in which they practiced the dual n-back task for twenty sessions, to a placebo control group in which they practiced another cognitive task for twenty sessions, or to a no-contact control group. Then, at posttest, all participants completed different versions of the cognitive ability tests. The results revealed that the dual n-back group was no higher in G*f* than the control groups. Other replication failures followed (e.g., Chooi & Thompson, 2012; Harrison et al., 2013). Moreover, meta-analyses demonstrated that benefits of brain training are limited to the trained task or to similar tasks, indicating near transfer but no far transfer (Melby-Lervåg & Hulme, 2013).

Subsequently, a letter published by the Stanford Center on Longevity (2014) and signed by seventy-five cognitive psychologists and neuroscientists cautioned, "The strong consensus of this group is that the scientific literature does not support claims that the use of software-based 'brain games' alters neural functioning in ways that improve general cognitive performance in everyday life, or prevent cognitive slowing and brain disease." A rebuttal letter (Cognitive Training Data, 2014), signed by over 100 scientists and practitioners – some of whom acknowledged financial interests in the brain training industry, notably including the cofounder of Posit Science, Michael Merzenich – countered that "a substantial and growing body of evidence shows that certain cognitive training regimens can significantly improve cognitive function, including in ways that generalize to everyday life."

In the wake of the controversy, Simons and colleagues (2016) conducted an exhaustive review of evidence used by brain training companies to promote their products, and concluded, "Brain training is appealing in part because it seems to provide a quick way to enhance cognition relative to the sustained investment required by education and skill acquisition. Practicing a cognitive task consistently improves performance on that task and closely related tasks, but the available evidence that such training generalizes to other tasks or to real-world performance is not compelling" (p. 173). Reaching the same conclusion, in 2016, the United States Federal Trade Commission (FTC) fined Lumosity $50 million for making unfounded claims about the real-world benefits of Lumosity games (the fine was reduced to $2 million because of financial hardship). Speaking to *NBC Nightly News*, a staff lawyer for the FTC commented, "There just isn't evidence that any of that [using Lumosity] will translate into any benefits in a real-world setting" (*NBC Nightly News*, 2016; see also Federal Trade Commission, 2016). The most recent meta-analysis on brain training corroborates this decision: There is no convincing evidence for far transfer of working memory training to real-world outcomes (Melby-Lervåg, Redick, & Hulme, 2016).

Obviously, through training, people can substantially improve their performance in complex problem-solving tasks, whether it be a game such as chess, an occupational task such as air traffic control, or an everyday task such as managing finances. People can, and do, develop high levels of skill in complex tasks through prolonged training. However, beyond improvements in the trained task or similar tasks, benefits of brain training appear to be nil.

Conclusions

Intelligence can be viewed as a general ability to solve problems. Consistent with this view, intelligence and complex problem-solving correlate very highly. Moreover, intelligence is a statistically and practically significant predictor of real-world outcomes that reflect problem-solving skill, including job performance and academic achievement. Intelligence even predicts success in the ultimate problem-solving task: staying alive. The question of what underlies intelligence still remains unanswered. The once popular view that intelligence (specifically G*f*) is isomorphic with working memory capacity has fallen out of favor. Meanwhile, we have discovered that G*f* correlates highly with placekeeping – the ability to perform a sequence of operations in a particular order. We are optimistic that this and other research aimed at identifying mechanisms underlying intelligence will provide the scientific foundation for a wide range of practical applications, from improving procedures for training people in complex tasks, to devising interventions for enhancing problem-solving skill, to refining measures used to predict people's performance in settings such as the workplace.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, *117*, 288–318.

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*, 227–257.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, *131*, 567–589.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, *131*, 30–60.

Anderson, J. R. (1985). *Cognitive psychology and its implications*. New York: W H Freeman/ Times Books/ Henry Holt & Co.

Arthur, W., Barret, G. V., & Alexander, R. A. (1991). Prediction of vehicular accident involvement: A meta-analysis. *Human Performance*, *4*, 89–105.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (vol. 8, pp. 47–89). New York: Academic Press.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*, 611–626.

Binet, A. (1909/1975). *Modern ideas about children* (trans. S. Heisler). California: University of California.

Bingham, W. V. (1937). *Aptitudes and aptitude testing*. Oxford: Harpers.

Brown, R. E. (2016). Hebb and Cattell: The genesis of the theory of fluid and crystallized intelligence. *Frontiers in Human Neuroscience*, *10*(606). https://doi.org/10.3389/fnhum.2016.00606

Burgoyne, A. P., Hambrick, D. Z., & Altmann, E. M. (in press). Placekeeping ability as a component of fluid intelligence: Not just working memory capacity. *American Journal of Psychology*.

Burgoyne, A. P., Sala, G., Gobet, F., Macnamara, B. N., Campitelli, G., & Hambrick, D. Z. (2016). The relationship between cognitive ability and chess skill: A comprehensive meta-analysis. *Intelligence*, *59*, 72–83.

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology*, *37*, 231–242.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*, 404–431.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*, 153–193.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1–22.

Chooi, W. T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence*, *40*, 531–542.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple regression/ correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Cognitive Training Data (2014). Cognitive training data response letter. www.cognitivetrainingdata.org/the-controversy-does-brain-training-work/response-letter

Conway, A. R. A., & Kovacs, K. (2013). Individual differences in intelligence and working memory: A review of latent variable models. *Psychology of Learning and Motivation*, *58*, 233–270.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Danner, D., Hagemann, D., Holt, D. V., Hager, M., Schankin, A., Wüstenberg, S., et al. (2011). Measuring performance in dynamic decision making. *Journal of Individual Differences*, *32*, 225–233.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford: Oxford University Press.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13–21.

Deary, I. J., & Whalley, L. J. (2001). Longitudinal cohort study of childhood IQ and survival up to age 76. *British Medical Journal*, *322*, 819–822.

Duncker, K. (1945). On problem-solving (trans. L. Lees). *Psychological Monographs*, *58*, i–113.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23.

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*, 190–193.

Ericsson, K. A. (2013). My exploration of Gagné's "evidence" for innate talent. It is Gagné who is omitting troublesome information so as to present more convincing accusations. In S. B. Kaufman (Ed.), *The complexity of greatness: Beyond talent or practice* (pp. 223–256). New York: Oxford University Press.

Ericsson, K. A. (2014). Why expert performance is special and cannot be extrapolated from studies of performance in the general population: A response to criticisms. *Intelligence*, *45*, 81–103.

Ericsson, K. A. (2018a). Superior working memory in experts. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 696–713). New York: Cambridge University Press.

Ericsson, K. A. (2018b). Intelligence as domain-specific reproducible performance: The role of acquired domain-specific mechanisms in expert performance. In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 85–100). New York: Cambridge University Press.

Ericsson, K. A., & Moxley, J. (2013). Working memory that mediates experts' performance: Why it is qualitatively different from traditional working memory. In T. P. Alloway & R. Alloway (Eds.), *Working memory: The connected intelligence* (pp. 109–136). New York: Psychology Press.

Ericsson, A., & Pool, R. (2016). *Peak: Secrets from the new science of expertise*. Boston: Houghton Mifflin Harcourt.

Farrell, J. N., & McDaniel, M. A. (2001). The stability of validity coefficients over time: Ackerman's (1988) model and the General Aptitude Test Battery. *Journal of Applied Psychology*, *86*, 60–79.

Federal Trade Commission (2016). Lumosity to pay $2 million to settle FTC deceptive advertising charges for its "brain training" program. www.ftc.gov/news-events/press-releases/2016/01/lumosity-pay-2-million-settle-ftc-deceptive-advertising-charges

Ferriman Robertson, K., Smeets, S., Lubinski, D., & Benbow, C. P. (2010). Beyond the threshold hypothesis: Even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence. *Current Directions in Psychological Science*, *19*, 346–351.

Frensch, P. A., & Funke, J. (1995). Definitions, traditions, and a general framework for understanding complex problem solving. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 3–26). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gallagher, A. G., Cowie, R., Crothers, I., Jordan-Black, J. A., & Satava, R. M. (2003). PicSOr: an objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopic performance. *Surgical Endoscopy and Other Interventional Techniques*, *17*, 1468–1471.

Gardner, H. (1999). Who owns intelligence? *The Atlantic Monthly*, *283*, 67–76.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*, 13–23.

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science*, *13*, 1–4.

Greiff, S., Fischer, A., Wüstenberg, S., Sonnleitner, P., Brunner, M., & Martin, R. (2013a). A multitrait-multimethod study of assessment instruments for complex problem solving. *Intelligence*, *41*, 579–596.

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: A new assessment perspective. *Applied Psychological Measurement*, *36*, 189–213.

Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013b). Complex problem solving in educational contexts – Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, *105*, 364–379.

Hambrick, D. Z., & Altmann, E. M. (2015). The role of placekeeping ability in fluid intelligence. *Psychonomic Bulletin and Review*, *22*, 1104–1110.

Hambrick, D. Z., Altmann, E. M., & Burgoyne, A. P. (2018). A knowledge activation approach to testing the circumvention-of-limits hypothesis. *American Journal of Psychology*, *131*, 307–321.

Hambrick, D. Z., Altmann, E. M., Oswald, F. L., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Accounting for expert performance: The devil is in the details. *Intelligence*, *45*, 112–114.

Hambrick, D. Z., Burgoyne, A. P., & Oswald, F. L. (in press). Domain-general models of expertise: The role of cognitive ability. In P. Ward, J. M. Schraagen, J. Gore, & E. Roth (Eds.), *Oxford handbook of expertise: Research and application*. Oxford: Oxford University Press.

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science*, *24*, 2409–2419.

Hart, C. L., Taylor, M. D., Smith, G. D., Whalley, L. J., Starr, J. M., Hole, D. J., et al. (2003). Childhood IQ, social class, deprivation, and their relationships with mortality and morbidity risk in later life: Prospective observational study linking the Scottish Mental Survey 1932 and the Midspan studies. *Psychosomatic Medicine*, *65*, 877–883.

Hays, W. L. (1988). *Statistics* (4th ed.). Fort Worth, TX: Holt, Rinehart & Winston Inc.

Hebb, D. O. (1942). The effect of early and late brain injury upon test scores, and the nature of normal adult intelligence. *Proceedings of the American Philosophical Society*, *85*, 275–292.

Henry, R. A., & Hulin, C. L. (1987). Stability of skilled performance across time: Some generalizations and limitations on utilities. *Journal of Applied Psychology*, *72*, 457–462.

Hunter, J. E., & Schmidt, F. L. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law*, *2*, 447–472.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, *105*, 6829–6833.

Jensen, A. (1999). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Johnson, J., & Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66–71.

Koenig, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence*, *36*, 153–160.

Kopiez, R., & Lee, J. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research*, *8*, 97–120.

Kopiez, R., & Lee, J. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research*, *10*, 41–62.

Kretzschmar, A., Neubert, J. C., Wüstenberg, S., & Greiff, S. (2016). Construct validity of complex problem solving: A comprehensive view on different facets of intelligence and school grades. *Intelligence*, *54*, 55–69.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, *86*, 148–161.

Kuncel, N., & Sackett, P. (2018). The truth about the SAT and ACT. *Wall Street Journal*, March 8. www.wsj.com/articles/the-truth-about-the-sat-and-act-1520521861

Kunda, M., Soulières, I., Rozga, A., & Goel, A. K. (2016). Error patterns on the Raven's Standard Progressive Matrices Test. *Intelligence*, *59*, 181–198.

Kyllonen, P. C. (2002). g: Knowledge, speed, strategies, or working-memory capacity? A systems perspective. In R. J. Sternberg & E. L. Gigorenko (Eds.), *The general factor of intelligence: How general is it?* (pp. 415–445). Mahwah, NJ: Erlbaum.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*, 389–433.

Lubinski, D. (2009). Exceptional cognitive ability: The phenotype. *Behavioral Genetics*, *39*, 350–358.

Macnamara, B. N., Hambrick, D. Z., & Moreau, D. (2016). How important is deliberate practice? Reply to Ericsson (2016). *Perspectives on Psychological Science*, *11*, 355–358.

Mayer, R. E. (2013). Problem solving. In D. Reisberg (Ed.), *The Oxford handbook of cognitive psychology* (pp. 769–778). New York: Oxford University Press.

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity, frontal lobe functioning, and general fluid intelligence: An adult lifespan study. *Neuropsychology*, *22*, 638–644.

McGrew, K. S. (2005). *The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future*. New York: Guilford Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*, 270–291.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science*, *11*, 512–534.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

NBC Nightly News (2016). Lumosity to pay $2M over allegations that its ads deceived customers. January 5. www.youtube.com/watch?v=uw67Vf_EwAw&gl=UG&hl=en-GB

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Park, G., Lubinski, D., & Benbow, C. P. (2008). Ability differences among people who have commensurate degrees matter for scientific creativity. *Psychological Science*, *19* (10), 957–961.

Postlethwaite, B., Robbins, S., Rickerson, J., & McKinniss, T. (2009). The moderation of conscientiousness by cognitive ability when predicting workplace safety behavior. *Personality and Individual Differences*, *47*, 711–716.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., et al. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General*, *142*, 359–379.

Reeve, C. L., & Bonaccio, S. (2011). On the myth and the reality of the temporal validity degradation of general mental ability test scores. *Intelligence*, *39*, 255–272.

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory and Language*, *39*, 418–436.

Ruthsatz, J., Detterman, D., Griscom, W. S., & Cirullo, B. A. (2008). Becoming an expert in the musical domain: It takes more than just practice. *Intelligence*, *36*, 330–338.

Salthouse, T. A. (2000). Pressing issues in cognitive aging. In D. Park & N. Swartz (Eds.), *Cognitive aging: A primer* (pp. 43–54). New York: Psychology Press.

Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin and Review*, *15*, 364–371.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262–274.

Schmidt, F. L., & Hunter, J. E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162–173.

Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Goff, S. (1988). Joint relation of experience and ability with job performance: Test of three hypotheses. *Journal of Applied Psychology*, *73*, 46–57.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, *11*, 771–799.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., et al. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest*, *17*, 103–186.

Sonnleitner, P., Keller, U., Martin, R., & Brunner, M. (2013). Students' complex problem-solving abilities: Their structure and relations to reasoning ability and educational success. *Intelligence*, *41*, 289–305.

Spearman, C. (1904). "General Intelligence," objectively determined and measured. *American Journal of Psychology*, *15*, 201–292.

Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence*, *53*, 92–101.

Stanford Center on Longevity (2014). A consensus on the brain training industry from the scientific community. http://longevity.stanford.edu/a-consensus-on-the-brain-training-industry-from-the-scientific-community-2

Stern, W. (1914). *The psychological methods of testing intelligence*. Baltimore, MD: Warwick & York.

Sternberg, R. J. (1985a). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, *49*, 607–627.

Sternberg, R. J. (1985b). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, UK: Cambridge University Press.

Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences*, *105*, 6791–6792.

Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, *41*, 37–55.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*, 127–154.

Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between operation span and Raven. *Intelligence*, *33*, 67–81.

Vernon, P. E. (1965). Ability factors and environmental influences. *American Psychologist*, *20*, 723–733.

Walshe, E. A., McIntosh, C. W., Romer, R., & Winston, F. K. (2017). Executive function capacities, negative driving behavior and crashes in young adults. *International Journal of Environmental Research and Public Health*, 14, E1314.

Wigdor, A., & Green, B. F., Jr. (1991). *Performance assessment for the workplace*, 2 vols. Washington: National Academy Press.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving – More than reasoning? *Intelligence*, *40*, 1–14.

24 Intelligence and Decision-Making

Gerd Gigerenzer

There are two ways to study how people make decisions. Decision-making under *risk* deals with well-defined situations where all possible outcomes and their probabilities are known for certain. If you visit a casino in Las Vegas to seek your fortune by playing the roulette and contemplate whether to bet $1,000 on the number "7" or on "red," all possible outcomes and their probabilities are known. No skills other than calculation are needed. Decision-making under *uncertainty*, by contrast, deals with ill-defined situations where this certainty is not attainable for humans or machines. If you ponder how to invest your money, whom to vote for, or whom to marry, the exhaustive set of all possible future outcomes is not known for certain, nor are their probabilities, and surprises may happen. Many situations involve a mixture of risk and uncertainty. For instance, the owner of a Las Vegas casino was able to calculate the gambling odds and the profit that the house could expect to make in the long run. Yet the main losses occurred through unforeseeable events: The star artist, performing his famous tiger act, was attacked by the animal; a disgruntled former contractor attempted to dynamite the casino; and an employee caused the casino to be heavily fined by failing to file tax reports for a long period (Taleb, 2007, pp. 129–130).

The distinction between risk and uncertainty goes back to the economist Frank Knight (1921). In situations of risk, the laws of logic and probability are sufficient to determine which action can be expected to result in the best outcome. In situations of uncertainty, probabilistic reasoning does not suffice, and psychological tools such as heuristics, experience, and intuition are needed. The distinction between risk and uncertainty is reminiscent of that between *known unknowns* and *unknown unknowns* in the NASA terminology popularized by former US secretary of defense Donald Rumsfeld.

In this chapter, I provide an introduction to these two approaches to decision-making and end with the question of how to integrate them with the study of intelligence. Each part begins with a brief historical sketch. The history of ideas is indispensable for understanding why we are asking the questions we ask today, and for understanding what the alternatives are.

Decision-Making under Risk

The Origins: Gambling

In psychological experiments, many thousands of participants have been asked to choose between two or more monetary gambles. This intimate fondness for gambling may appear surprising. After all, one is unlikely to find hordes of psychologists in the machine gambling zone of Las Vegas. Yet for decision research, history is destiny.

Gambling became a craze in several European countries during the seventeenth and eighteenth centuries. People passionately bet their money on lotteries and the outcomes of dice rolls to seek their fortune (Daston, 1988). According to legend, quite a few rebellious citizens of Paris made their regular stop at the lottery stands to try their luck on the way to storming the Bastille at the start of the French Revolution. In England and France, life insurance was almost synonymous with gambling, being not insurance as we know it but rather a bet on the life of a *third party* (Daston, 1987). For instance, London underwriters issued policies on the lives of celebrities, such as the unfortunate Admiral Byng, who stood trial for failing to prevent the island Minorca from falling to the French. Byng was found guilty and shot by a firing squad, much to the satisfaction of those who had bought insurance on his life (Clark, 2002). Underwriters even offered bets on the lives of kings, such as 25 percent against King George II's returning alive from the Battle of Dettingen, Germany. The last British monarch to personally lead his troops on the field, George II survived the battle; in this case, the gamblers lost their money.

A huge shift in psychological perspective was required to eventually convince people that they should bet on their own death rather than on that of someone else. The increasing mobility of society in the nineteenth century helped; no longer could people expect to live their lives in a stationary community that would look after a widow and children if the father unexpectedly died. During this shift, books on probability emphasized that probability had outgrown its frivolous phase and could now serve more moral and sober pursuits, such as pricing insurance and annuities (Daston, 1987). Today, of course, life insurance is often perceived as a moral duty, and it no longer feels strange to bet on one's own death. Moral or immoral, gambling and lotteries have defined decision theory under risk.

Expected Value Theory

Following its historical origins, the staple task posed to participants in psychological experiments is a choice between lotteries, such as:

A. A 50 percent chance to win $100, otherwise nothing.
B. $40 for sure.

Which option would you choose? If you prefer A, you maximize the *expected value.* The expected value of gamble is the sum of all outcomes times their probabilities. For gamble A, the expectation is $0.5 \times \$100 + 0.5 \times \$0 = \$50$, while the sure option B pays only $40. According to *expected value theory* it is prudent to choose the gamble with the maximum expected value, here option A.

Risk Aversion and Risk Seeking

If a person prefers a sure gain to a lottery with a higher expected value, that person is said to be *risk averse*. An example is to prefer B over A. If a person prefers a lottery to a sure gain despite the lottery having a smaller expected value, that person is classified as *risk seeking*. An example would be to prefer A over a sure gain of $60. If the amounts at stake are *losses* rather than *gains*, then the person is said to be risk averse or risk seeking for losses.

How do actual people decide? They are divided. In a representative sample of about 1,000 Germans and 1,000 Spaniards, 55 percent were risk averse and preferred $40 for sure, thereby violating expected value theory (Gigerenzer & Garcia-Retamero, 2017). As the historical association between gambling and insurance might lead one to expect, those who were risk averse more often bought life insurance, household insurance, and other insurances than did those who were risk seeking. Risk-averse people were also more likely to not want to know when they will die and from what cause, and whether their marriage will end in divorce.

Expected Utility Theory

Expected value theory was the brainchild of two famous French mathematicians, Blaise Pascal and Pierre Fermat, who in 1654 solved gambling problems posed by a notorious gambler. Yet expected value theory soon ran into troubles because it conflicted with the intuitions of educated people, as in the case of risk aversion. The most celebrated conflict was another gamble:

> A fair coin is tossed. If "heads" comes up on the first flip, the house pays the gambler $1, and the game ends. If the first head is on the second flip, the house pays $2, and the game ends. If the first head is on the third flip, the house pays $4, and so on. In general, if the first head is on the *n*th flip, then the house pays $\$2^{n-1}$.

What is the fair entry price for this game? *Fair* means leaving open which side of the bet one takes. (The same principle of fairness is in play when a child cuts a piece of cake in two, and a second child gets to choose first.) If the fair price is the expected value, we get:

$$\text{Expected value} = 1/2 \times \$1 + 1/4 \times \$2 + 1/8 \times \$4 + \ldots + (1/2)^n \text{ x } \$2^{n-1} + \ldots = \$1/2 + \$1/2 + \ldots = \infty$$

Translated into words, the probability of winning $1 is 1/2, that of winning $2 is 1/4, that of winning $4 is 1/8, and so on. Each term on the right side of the equation equals 50 cents, and since the number of terms is infinite, the expected value is also infinitely large. In keeping with the theory, we should thus wager everything we own to play this game. Yet no sensible person would be willing to pay more than a small sum. The discrepancy between expected value theory and people's intuitions was dubbed the *St. Petersburg Paradox.*

In 1738, the twenty-five-year-old Swiss mathematician Daniel Bernoulli published a solution in the annuals of the Petersburg Academy (1738/1954), after which the paradox was named. Bernoulli reasoned that winning $200 (in modern currency) is not necessarily double the "utility" of winning $100, and that the wealthier a player is to begin with, the more money he needs to win in order to experience an equal increase of utility. This reasoning later became known as the *principle of marginal decreasing utility*. Bernoulli assumed that the relation between dollar (x) and utility (u) is logarithmic, $u(x) = \log(x)$. By replacing the actual dollar values with their logarithms, the problem of the infinite "expected value" of the gamble was resolved. The "expected utility" of the gamble amounts to a small dollar value, in the region of what most people are willing to pay.

Today, a weaker version of Bernoulli's equation is called *expected utility theory*: it is "weaker" because, unlike Bernoulli's logarithmic function, the utility function u is not specified. Here, the expected utility of an option is defined as the sum of the utilities of its consequences, multiplied by their probabilities. Its central idea is that people decide between options as if it were possible to know beforehand all consequences of each option, then estimate their utilities, multiply these by their probabilities, add these terms up, and finally choose the option with the highest expected utility. Although this knowledge is unlikely to exist in situations of uncertainty, expected utility theory became the template for a large number of psychological theories, including consumer choice, health behavior, attitude formation, motivation, and intuition. The theory also single-handedly shaped entire disciplines such as economics and finance.

Axioms of Choice

Despite the mathematical appeal of expected utility theory, criticism was raised about the complete lack of evidence that people actually perform the sequence of calculations the theory entails. This criticism was countered by arguing that if people satisfy a small number of choice axioms, then they behave *as if* they maximized a utility function. In fact, one of the celebrated successes of decision theory is the proof by von Neumann and Morgenstern that if the following choice axioms hold, then the options can be represented as numbers on a line called *utility function*:

Axiom 1: Completeness: $A \succeq B$ or $B \succeq A$.
Axiom 2: Transitivity: if $A \succeq B$, $B \succeq C$, then $A \succeq C$.
Axiom 3: Archimedean property (assumes continuity).
Axiom 4: Independence: if $A \succ B$, then $pA + (1 - p)C \succ pB + (1 - p)C$ (for any C and probability p).

A, B, C are the options in the choice set; $A \succeq B$ means that one either prefers A over B or is indifferent between both ("weak preference"). Completeness means that one either prefers A weakly over B, or vice versa. Everything else is excluded, such as not having any preference or not making a choice. Transitivity means that if one prefers A over B, and B over C, then one also prefers A over C. The Archimedean property, named after the ancient Greek geometer Archimedes of Syracuse, roughly means that some kind of trade-off is always possible, which guarantees continuity of the number line. Independence means that if the same amount is added to each of two options, their order remains the same. In expected utility theory, preferences are simply inferred from choices, which is called the *principle of revealed preferences.*

As the axioms show, the theory is behavioristic: Preference does not mean "liking" or "deriving more pleasure" but only consistent choice.

Von Neumann and Morgenstern formulated these axioms in order to provide the mathematical conditions for a utility function, which are similar to the axioms of number theory. When doing so, they made no claims that these describe what people do or should do. Others proposed that the axioms – and, by implication, expected utility theory – describe how people actually make decisions or, at the very least, how they ought to make decisions. Both the descriptive and the prescriptive interpretation generated intense controversies.

Controversies

Do Choice Axioms Describe Behavior?

Two kinds of arguments have been levied against the descriptive interpretation. The first is theoretical. Consider the completeness axiom. It appears to be almost trivial to find out whether a person's choices fulfill this axiom, yet it is not. For instance, consider the decision of which websites to visit and in which order. According to Internet Live Stats, ten websites existed on the Internet in 1992. To order these according to preference, one had to make forty-five ($10 \times 9/2$) binary choices. At that time, checking for completeness was tractable. In the year 2016, the number of websites had increased to about 1,085,628,900, which would require in the order of 10^{18} checks. Here, checking for completeness is no longer tractable, neither for humans nor for machines. And without that, one cannot check transitivity and find out whether the choice axioms describe behavior. In general, if there is a large choice set, checking for completeness is computationally intractable. And even when it is tractable, checking may amount to a foolish loss of time.

The second argument is empirical. Beginning with the Allais paradox and the Ellsberg paradox, a number of psychological studies showed that people systematically violate expected utility theory in simple choice situations (e.g., Kahneman & Tversky, 1979). Allais (1953) constructed gambles in which people tend to violate the independence axiom, and Ellsberg (1961) showed that people's preferences for gambles are sensitive to whether the probabilities are known (risk) or ambiguous (a form of uncertainty), all of which result in behavior inconsistent with expected utility theory. This has led to revisions of expected utility theory by adding more free parameters to fit deviating behavior, such as prospect theory.

Thus, expected utility theory faced the same challenge as expected value theory. People sometimes behave *as if* they maximized expected utility, at other times not. Most important, there is little evidence that expected utility theory (or its modifications, such as prospect theory) describes how people actually reason, that is, that the postulated multiplications, additions, and transformations of values are performed by humans. Yet there is also an important difference. When expected value theory conflicted with educated people's intuition, the blame was placed on the theory, not on the people. When expected utility theory contradicted people's decisions, the blame was placed on people, not on the theory, which was maintained as being prescriptive.

Do Choice Axioms Prescribe Behavior?

Beginning in the 1970s, the *heuristics-and-biases program* documented that people's judgments systematically deviate from choice axioms and various other logical rules. In contrast to Bernoulli, Kahneman and Tversky (1974) attributed these discrepancies to flaws in the human mind rather than in the norms. These deviations were named *cognitive illusions* in analogy to visual illusions, suggesting that they are equally stable and stubborn, and people's intuitions were called "a multitude of sins," "ludicrous," "indefensible," and "self-defeating" (Tversky & Kahneman, 1971, pp. 107–110). Today, long lists of cognitive illusions exist, including violations of transitivity, the conjunction fallacy, and framing effects, with Wikipedia cataloguing some 175 of these. But again, two arguments, one theoretical and one empirical, have been mounted against the interpretation of logical axioms and rules as prescriptive in all situations and their violations as cognitive illusions.

First, when Maurice Allais and Daniel Ellsberg (the man who released the Pentagon papers) published their famous "paradoxes" in the 1950s and 1960s, demonstrating systematic discrepancies between people's intuitions and the choice axioms, they criticized the normative interpretation of the latter. Ellsberg (1961) urged distinguishing between risk and uncertainty. He concluded that the belief held by many researchers that choice axioms define rational behavior under uncertainty amounts to "bad advice" (p. 669), and that when he and other people systematically violate logical axioms, this is not irrational but in fact a sensible way to behave.

A related critique is that expected utility theory pays attention solely to the mean outcome, not to the variance (or to higher moments) of the outcomes. If one pays attention to the variance, a sure gain of $40 in alternative A is *not* necessarily inferior to an expected value of $50 in alternative B, because the first outcome has variance zero (it is certain) while the second has a variability between $0 and $100. Thus, looking at the expectation (the mean) alone is not a universal yardstick of rational decision-making.

The second argument is empirical. It is directed against the claim (Thaler & Sunstein, 2008) that violations of logical rules incur substantial real-world costs. Arkes, Gigerenzer, and Hertwig (2016) searched through more than 1,000 articles on so-called cognitive illusions, and could find little to no evidence that violations of logical rules are associated with less income, poorer health, lower happiness, inaccurate beliefs, shorter lives, or any other measurable outcome (see also Berg, Biele, & Gigerenzer, 2016). Similarly, when Stanovich and West (2008) investigated whether the biases discussed in the heuristics and biases literature were correlated with measures of ability such as SAT tests, they concluded "that a large number of thinking biases are uncorrelated with cognitive ability" (p. 672).

The Turkey Illusion

Many psychologists do not distinguish between risk and uncertainty but instead believe that all problems can be solved by probability theory, as assumed in Bayesian theories of mind or brain. This position has been criticized as being the "turkey illusion."

Imagine you are a turkey. It's the first day of your life. All of a sudden a man appears. In panic, you fear that he will kill you, but he kindly feeds you. On the second day, the man returns, and again you fear that he might kill you. But once more, he feeds you. On the third day, the same happens. According to Bayesian probability updating, the probability that he will feed rather than kill you increases each day. On day 100, it is higher than ever before – but it happens to be the day before Thanksgiving. You are dead meat.

The turkey missed a crucial piece of information: It was not in a situation of risk, where the past predicts the future. The turkey illusion goes back to the philosopher Bertrand Russell (who used a chicken in his account), and has been popularized by trader Nassim Taleb. Yet it appears to be committed more frequently by humans than turkeys, one example being the increasing confidence of financial institutions in the stability of the financial market in the years before the crisis of 2007–2008, up to shortly before the breakdown. Calibrating their models on the past years, the rating agencies – like the turkey – predicted that the future resembles the past.

All three points of critique – the descriptive, the prescriptive, and the turkey – are essential whenever decisions are evaluated as rational or irrational. At a minimum, these suggest that choice axioms, or similar logical rules, are of limited value for describing how people make decisions and inadequate as universal norms of how we should behave. Based on these limits, Herbert Simon (1955, 1979) asked for an extension and revision of the study of decision-making in two respects: (1) to study how people make decisions in the real world of uncertainty, as opposed to risk, and (2) to study the process of how people actually make decisions, as opposed to as-if models of expected utility maximization and its variants. Yet Simon's call was little heeded for decades, and it took several more decades before his program was finally fleshed out (Gigerenzer, Hertwig, & Pachur, 2011).

Decision-Making under Uncertainty

The Origins: Heuristic Decision-Making

As far as we can know, humans and other animals have always relied on heuristics to solve adaptive problems. Ants use a simple rule to estimate the area of a candidate nest cavity: Run around on an irregular path for a fixed period while laying down a pheromone trail, then leave; next return, move around again in an irregular path, and estimate the size of the cavity by the frequency of encountering one's old trail. This heuristic is remarkably precise: Nests half the area of others yielded reencounter frequencies 1.96 times greater.

To choose a mate, peahens similarly use a heuristic: Investigate only three or four of the peacocks in a lek (an assembly of males engaged in competitive displays to attract a mate) and choose the one with the largest number of eyespots (see Hutchinson & Gigerenzer, 2005). Roughly speaking, a heuristic is a simple rule that uses a minimum of the available information (in contrast to expected utility maximization) in order to make efficient decisions under uncertainty. The term *heuristic* is of Greek origin, meaning "serving to find out or discover."

One of the fathers of the study of heuristics is Herbert A. Simon (1916–2001). He also made seminal contributions to artificial intelligence, psychology, political science, and economics. In the mid-1930s, young "Herb" had taken a class on price theory at the University of Chicago, and then tried to apply what he had learned on utility maximization to real budget decisions in his native Milwaukee recreation department. To his surprise, experienced managers did not estimate utilities and probabilities but rather used heuristics and added incremental changes to last year's budget. This venture into the real world taught him that even experienced managers cannot know in advance all possible states of the world, their consequences, and probabilities. In an uncertain world, he concluded, utility maximization was hopeless.

Simon (1955, 1979, 1990) proposed a division of labor: Heuristic reasoning is necessary in situations of uncertainty, while probabilistic reasoning is necessary for situations of risk. Contrast this with the heuristics-and-biases program. First, Simon rejected choice axioms and expected utility theory as a universal principle, whereas Kahneman and Tversky accepted these and similar logical rules as universal yardsticks for rationality and attributed deviating behavior to flaws in the human mind. Second, while Simon insisted on formal models of heuristics that could be simulated and tested, the heuristics-and-bias program relied on vague one-word labels, such as availability and representativeness. By doing so, this program could not discover the power of heuristics. In this chapter, I will focus on models of heuristics, either formal or in the form of precise verbal descriptions of the steps of decision-making, as studied by Simon and subsequently by Payne, Bettman, and Johnson (1993) and Gigerenzer, Todd, and the ABC Research Group (1999). In what follows, I proceed by means of examples; a systematic treatment can be found in Gigerenzer and colleagues (2011).

The program of fast-and-frugal heuristics (Gigerenzer et al., 2011) is often perceived as being in opposition to the heuristics-and-biases program (Kahneman, 2011). In fact, it should be seen as a necessary extension of the latter by introducing formal models of heuristics, replacing logical with ecological rationality, and taking uncertainty seriously. The program has a descriptive, prescriptive, and engineering component: the study of the *adaptive toolbox* (which heuristics are in the repertoire of a person?), *ecological rationality* (which heuristic should be selected for a given problem?), and *intuitive design* (how to design intuitive expert systems based on the adaptive toolbox).

The Adaptive Toolbox

Through individual learning and social imitation, humans acquire during their lifetime a repertoire of heuristics. This repertoire is called the *adaptive toolbox* of a person, where the term *adaptive* signals that heuristics are tailored to classes of problems, just as a hammer is designed for nails and screwdrivers for screws. The key difference to decision-making under risk is that these tools can deal with uncertainty. To illustrate the difference, I begin with a decision similar to the choice between gambles, the stock-in-trade for decision-making under risk.

Satisficing

Consider an entrepreneur who is looking for an investment in real estate. She has discovered a potential site S to invest in and develop, and is faced with this decision:

C. Invest in site S.
D. Forgo S and continue to search for a better site.

Compare this choice to the one between A and B. First, the set of all possible outcomes (profits) of site S and their probabilities are no longer known for sure. Second, option D entails further uncertainty: New sites need to be searched for, without knowing ahead what, if anything, will be found. In other words, decision-making takes place in time; options are not presented simultaneously, as in the choice between A and B, but discovered sequentially. If the set of all possible sites, their outcomes, and their probabilities are not known, one cannot determine the site with the highest expected utility. Nevertheless, people can make good decisions. But how? Simon argued that people rely on a heuristic called *satisficing*, named after the Northumbrian word for "good enough."

> *Satisficing:* Set an aspiration level *x*, and choose the first object that satisfies *x*.

Consider again the decision between C and D. Berg (2014) studied forty-nine entrepreneurs in the Dallas-Fort Worth greater metropolitan area who developed commercial high-rises or residential areas. He reported that every single one of the professionals relied on satisficing:

> If I believe I can get at least *x* percent return within *y* years, then I take the option.

The entrepreneurs differed in their aspiration levels and the time horizon. The time horizon *y* was mostly one to three years, and *x* a *prominent number*. Prominent numbers are powers of ten, their halves, and their doubles (i.e., 1, 2, 5, 10, 20, 50, . . .). For instance, convenience store and gas station investors required at least a 10 percent annual return on capital within one or two years. Most entrepreneurs considered only one, two, or three sites before making a decision (similar to the peahen's mate choice). Not a single one tried to determine the point where the marginal benefit of search equals its costs. Many expressed open skepticism that such utility calculations could be made in one-off decisions in high-stake and quickly changing environments.

Like all heuristics, satisficing is used in a variety of sequential search problems, from consumer choice to mate choice (Todd & Miller, 1999). Consider a mundane everyday decision: fast and frugal food choice. How do customers in a restaurant decide what to order for dinner? Unlike the entrepreneurs, customers differ in the heuristics they use. Yet again, one of these is satisficing.

Satisficing: First, pick a category from the menu (say, fish). Then read the first item in this category, and decide whether it is good enough. If yes, close the menu and order that dish. If no, read the second item and proceed in the same way.

In a representative study of 1,000 German adults, 34 percent reported that this fast-and-frugal rule is how they typically decide (Figure 24.1; Gigerenzer, 2014). Yet 17 percent reported an even faster rule:

Habit: Don't open the menu. Order your favorite dish.

This rule is reminiscent of risk aversion; it avoids disappointment and appears reasonable when one is familiar with the restaurant. Satisficing and habit are both individual heuristics, where the decision is made without social input. The next two, in contrast, are social heuristics.

Advice taking: Don't open the menu. Ask the waiter what they recommend and order it.

Advice taking is a social heuristic because it relies on and trusts the judgment of someone else. In situations where one has little experience, social heuristics are generally useful. Advice taking is a reasonable rule in a good restaurant where the waiter knows what is best and does not deceive the guest. Among the Germans

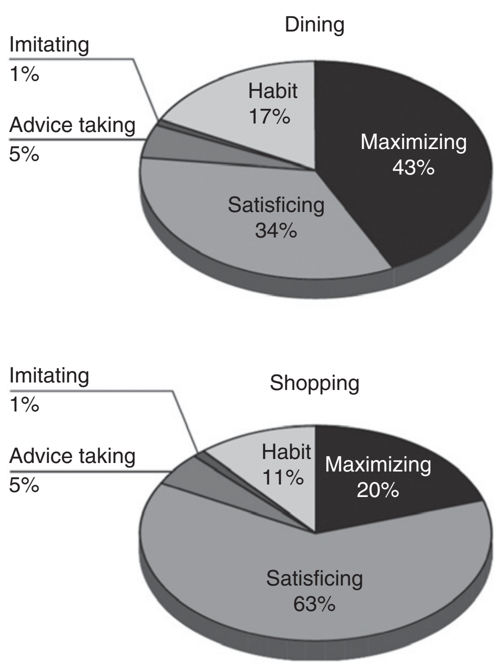

Figure 24.1 *Individual differences in heuristic decision-making (from Gigerenzer, 2014).*

Top: How to order in a restaurant? Bottom: How to shop for a pair of trousers? Shown are the responses of a representative sample of 1,000 German adults.

surveyed, however, it was not very popular, with only 5 percent relying on it. The least popular heuristic was also a social one:

> *Imitating:* Don't open the menu. Find out who among your friends at the table has most experience with this restaurant, and order the dish they order.

Imitating or copying appears reasonable when in a foreign country or an unfamiliar restaurant. Yet only 1 percent of the general public reported relying on it. Copying appears to be a taboo among Germans, unlike in more family-oriented countries where it is fine to eat what others eat. The largest group (43%), however, relied on a time-expensive rule:

> *Maximizing:* Study every item in the menu carefully and try to figure out the best option.

Maximizing is an individualistic rule and requires stamina in a restaurant where the menu resembles an encyclopedia. Yet for some people, making a decision without having inspected all options is emotionally unbearable. Maximizing resembles expected utility calculations, but there is an essential difference: No probabilities and utilities are estimated, multiplied, and added up.

When the same participants were asked how they proceed when buying a pair of trousers, the major difference was that more reported satisficing and fewer maximizing (Figure 24.1). Maximizing requires checking all trousers in a department store, then heading to another department store and on to boutiques to find the best pair. Maximizing can be a direct route to unhappiness: One wants the best, and nothing less. But even if happening on it right away, one would not know it and continue to look for something better. Studies indicate that people who rely on satisficing tend to be more optimistic and have higher self-esteem than those who rely on maximizing. The latter excel in perfectionism, depression, and self-blame (Schwartz et al., 2002). Food choice and shopping are not the biggest decisions we face. Imagine if everyone tried to maximize in mate choice, that is, to find the perfect partner and nothing less. That would be a recipe for disaster and divorce.

Once again, social heuristics – to take advice from a salesperson or buy the trousers others wear – were as rarely reported as in food choice. This reluctance to rely on others when making decisions, or to admit to doing so, may be particularly strong in individualistic Western societies (Hertwig & Hoffrage, 2013).

Do Animals and Humans Share Common Heuristics?

Both animals and humans rely on heuristics to deal with uncertainty, and sometimes even rely on the same heuristic. In an experiment, Norway rats had a choice between two kinds of food, one they recognized by smell from the breath of another rat, while the other was new. The far majority of rats choose the recognized food, even in situations where the fellow rat was sick (experimentally induced). This decision rule is known as the *recognition heuristic* (Table 24.1). Experiments with humans showed that they tend to rely on the same heuristic, specifically when it is *ecologically rational*. The recognition heuristic is said to be ecologically rational in

Table 24.1 *Twelve well-studied heuristics with evidence of use in the adaptive toolbox of humans (after Todd, Gigerenzer, & ABC Research Group, 2012, pp. 9–10)*

Heuristic	**Description**	**Counterintuitive results**
Recognition heuristic (Goldstein & Gigerenzer, 2002)	If one of two alternatives is recognized, infer that it has the higher value on the criterion.	Less-is-more effect.
Fluency heuristic (Schooler & Hertwig, 2005)	If both alternatives are recognized but one is recognized faster, infer that it has the higher value on the criterion.	Less-is-more effect.
Take-the-best (Gigerenzer & Goldstein, 1996)	To infer which of the two alternatives has the higher value, (1) search through cues in order of validity; (2) stop search as soon as a cue discriminates; (3) choose the alternative this cue favors.	Often predicts as accurately as or better than multiple regression, neural networks, exemplar models, and decision-tree algorithms.
Fast-and-frugal trees (Martignon et al., 2003)	To classify a person or object, (1) search through cues in order; (2) stop search when first cue is found that allows a decision; (3) choose the option at the decision node.	Often predicts as accurately as or better than logistic regression.
Tallying (Dawes, 1979)	To estimate a criterion, do not estimate weights but simply count the number of positive cues.	Can predict as accurately as or better than multiple regression.
Satisficing (Simon, 1955)	Search through alternatives and choose the first one that exceeds your aspiration level.	Aspiration levels can lead to substantially better choices than by chance, even if they are arbitrary.
Gaze heuristic (McBeath, Shafer, & Kaiser, 1995)	To catch a ball that is coming down from overhead, fix your gaze on it, start running, and adjust your running speed so that the angle of gaze remain constant.	Balls will be caught while running, possibly on a curved path.
1/*N* rule (DeMiguel, Garlappi, & Uppal, 2009)	Allocate resources equally to each of *N* alternatives.	Can outperform optimal asset allocation portfolios.
Default heuristic (Johnson & Goldstein, 2003)	If there is a default, follow it.	Explains cultural differences in organ donor registration; predicts behavior when trait and preference theories fail.
Tit-for-tat (Axelrod, 1984)	Cooperate first and then imitate your partner's last behavior.	Can lead to a higher payoff than "rational" strategies (e.g., backward induction).

Table 24.1 (*cont.*)

Imitate the majority (Boyd & Richardson, 2005)	Determine the behavior followed by the majority of people in your group and imitate it.	A driving force in bonding, group identification, and moral behavior.
Imitate the successful (Boyd & Richardson, 2005)	Determine the most successful person and imitate their behavior.	A driving force in cultural evolution.

situations where recognition is correlated with the criterion. For instance, human participants were presented pairs of Swiss cities, such as Aarau and Basel, and asked which of the two cities has the larger population (Pohl, 2006). Most had not heard of Aarau but of Basel, so using the recognition heuristic would lead to the inference that Basel is larger (which is correct). Across all pairs of cities, 89 percent of the inferences followed the recognition heuristic. When the question was changed from population to which city is nearer to the center of Switzerland, the percentage went down to 54 percent. This sensitivity of the participants is ecologically rational: Name recognition is a valid predictor for population but not for distance from the center of Switzerland, and the difference in validity was almost identical to the difference in use of the heuristic. People's adaptive use of the recognition heuristic has been documented across forty-three experiments, where the correlation between recognition validity and the percentage of cases in which people follow it is $r = 0.57$ (Gigerenzer & Goldstein, 2011). The similarities and differences in heuristic decision-making in animals and humans is the topic of Hutchinson and Gigerenzer (2005).

Do Animals, Humans, and Machines Share Common Heuristics?

Consider the *gaze heuristic*, which is used by living organisms and machines. When a hawk pursues a dove, it relies on the *gaze heuristic*, fixating its eyes on the target and adapting the direction of flight so that the angle of gaze always remains constant (Hamlin, 2017). The angle of gaze is the angle between the direction of the hawk's flight and the line between the position of the hawk and the dove. Using this fast-and-frugal heuristic, the hawk does not need to estimate the dove's trajectory in three-dimensional space and calculate the intersection point. Moreover, when the target tries to evade, the pursuer adjusts its course so that the angle of gaze remains constant. Dogs rely on the same heuristic to catch Frisbees, as do baseball outfielders to intercept fly balls and sailors to avoid collisions (Table 24.1). The heuristic was built into the Sidewinder A1M9 short-range air-to-air missile, one of the most successful guided modern weapon systems and still in use, whose "gaze" is directed at a source of heat, which is the target.

The gaze heuristic is a prime example of a simple, robust decision system that has been discovered by animals, humans, and controllers of fighter planes and missiles. Moreover, it illustrates how a heuristic can travel from animal to human to machine, providing a simple solution to complex problems in nonstationary environments.

Like the recognition heuristic, it relies on a single input; both are members of the class of one-good-reason heuristics (Gigerenzer & Gaissmaier, 2011).

Ecological Rationality

The term *rationality* has at least two different meanings: logical consistency and attainment of one's goals. Logical rationality means that behavior should conform to logical rules, such as the choice axioms. Ecological rationality, in contrast, means that behavior should lead to successful performance, as measured by accuracy of prediction, speed of decision-making, or efficiency in reaching one's goals. As mentioned before, consistency and accuracy of beliefs appear largely unrelated, and violations of logical consistency do not seem to have demonstrable consequences for performance. The rationality of heuristics is not logical but ecological. In the words of Simon (1990), rational behavior "is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor" (p. 7).

The study of ecological rationality asks the question, when should people rely on a given heuristic rather than a complex strategy to make better decisions? While the study of the adaptive toolbox is descriptive, relying on experiment and observation, that of ecological rationality is prescriptive, relying on mathematical proof and computer simulation. As mentioned above, the recognition heuristic is ecologically rational in situations where the recognition validity (the correlation between recognition and the criterion) is substantially above chance. The study of ecological rationality delves deeper into the analysis of environmental structures, beyond the scope of this introduction. The conditions for the ecological rationality of the heuristics in Table 24.1 can be found in Todd and colleagues (2012), pp. 9–10, and in Gigerenzer (2016).

Intuitive Design

Expert systems can be classified into those that aim at efficiency and those that additionally value transparency and simplicity so that users can understand how the system works. Machine learning focuses mainly on accuracy and efficiency. Its techniques range from logistic regression, which many experts, such as most physicians and judges, have difficulties understanding, to deep neural networks whose inner workings are opaque even to its creators. Intuitive design, in contrast, aims at expert systems that are both efficient and transparent. It is called *intuitive* because it applies the results of the study of the adaptive toolbox and ecological rationality to design systems that mirror people's psychological processes and thus can be easily learned, remembered, and used.

To illustrate intuitive design, I take a class of heuristics called *fast-and-frugal trees* (Table 24.1). A fast-and-frugal tree embodies three principles of human decision-making under uncertainty: ordering, limited search, and one-reason decision-making. In contrast, a full decision tree does not order cues (reasons), uses exhaustive search, and combines all reasons to make the final decision. In more formal terms, a fast-and-frugal tree asks only a few questions (cues) and allows for making a decision after each question. If n is the number of questions with yes/no answers, then a fast-and-frugal tree has $n + 1$ exits, whereas a full tree has 2^n exits. Figure 24.2 (a) shows a fast-and-frugal tree. It is a model of how magistrates at London courts decide whether to grant bail or subject a defendant to a punitive measure such as jail. The tree has three building blocks:

Search rule: Look through cues in order.
Stopping rule: Stop search when the first cue is found that allows a decision.
Decision rule: Choose the option at the exit.

For instance, if the prosecution requests conditional bail or opposes bail, then magistrates follow suit and make a punitive decision such as jail. If the prosecution does not, then magistrates respond to a second question in the same way, and so on. This fast-and-frugal tree predicted the actual decisions of British magistrates better than linear models that used more cues. In experimental research, this form of sequential decision-making has been often reported. It differs from the prescriptions of expected utility theory, where all information should be searched for and integrated. Intuitive design begins with this structure rather than a logistic regression or

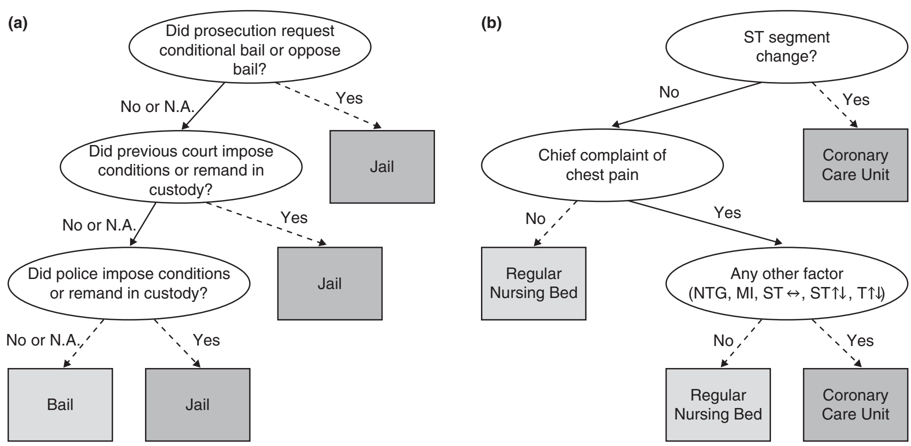

Figure 24.2 *Fast-and-frugal trees for bail decisions in London courts (a) and for coronary care unit (CCU) allocation in emergency medicine (b).*
The bail-and-jail tree is a descriptive model of how London magistrates decide whether to grant bail or subject a defendant to a punitive measure such as jail (Dhami, 2003). The CCU tree is a prescriptive model of how physicians should make decisions, embodying intuitive design. It led to fewer errors in allocation decisions compared to a complex logistic regression and the defensive decisions of physicians (Green & Mehr, 1997). ST segment change = electrocardiogram shows ST segment with elevation or depression of 1mm or more; NTG = history of nitroglycerin use for chest pain; MI = history of heart attack; the other measures are based on the electrocardiogram.

similar statistical models; the idea is to build expert systems that mirror human psychology.

Based on these principles of intuitive decision-making, the fast-and-frugal tree on the right side of Figure 24.2 has been designed by medical researchers for a medical emergency situation: when a patient is rushed into a hospital with severe chest pains. The emergency physicians have to decide quickly whether the patient suffers from acute ischemic heart disease and should be assigned to an intensive unit, the coronary care unit (CCU), or a regular bed with telemetry. The tree asks only three questions, and allows for a decision after each one. For instance, if there is an anomaly in the ST segment of the electrocardiogram, then the patient is immediately assigned to the CCU. No further questions are asked.

The tree was developed in a Michigan hospital in response to the problem that doctors used to send about 90 percent of patients into the CCU, although only 25 percent of these actually had a myocardial infarction (Green & Mehr, 1997). The result of this defensive decision-making was an overly crowded coronary care unit, decrease in quality of care, increase in cost, and a risk of serious infection among those who were incorrectly assigned ("false alarms"). The fast-and-frugal tree reduced both false alarms and misses considerably compared to doctors' decisions and to a logistic regression expert system for patient allocation. The tree is intuitive because, unlike a logistic regression, doctors can understand and memorize it, and also adapt it to new patient populations.

How to Balance False Alarms and Misses

In both bail and care unit decisions, one can make two kinds of errors. The first is a *false alarm*: classifying someone wrongly as positive, such as sending a defendant to jail who would have not committed a crime or a patient to the CCU who has no heart disease. The second error is called a *miss*: classifying someone wrongly as negative, such as granting bail to a defendant who will then commit a crime or allocating a patient who will have a heart attack to a regular bed. The design of the tree determines the balance of the two error rates. In the bail-and-jail tree, all exits are "jail" except one, which reduces the rate of misses at the cost of increasing the rate of false alarms. In the CCU tree, the exits – and, accordingly, the two possible errors – are more balanced.

These two examples illustrate a general principle for constructing fast-and-frugal trees. With three ordered cues, there are four possible trees with differing exits ("S" for signal and "N" for noise). The top of Figure 24.3 depicts signal detection theory with two hypotheses, noise and signal, and a decision criterion that sets the balance between the two errors. If the data fall to the left of the criterion, the conclusion is "noise"; if they fall to the right, the conclusion is "signal." The bottom of Figure 24.3 shows how the four possible fast-and-frugal trees map into this scheme (Luan, Schooler, & Gigerenzer, 2011). The tree on the far left side corresponds to a decision criterion set at the far left arrow, which reduces misses at the cost of more false alarms. This tree has the structure of the bail-and-jail tree. The CCU tree has the structure of the second tree from the left, striking more of a balance between misses

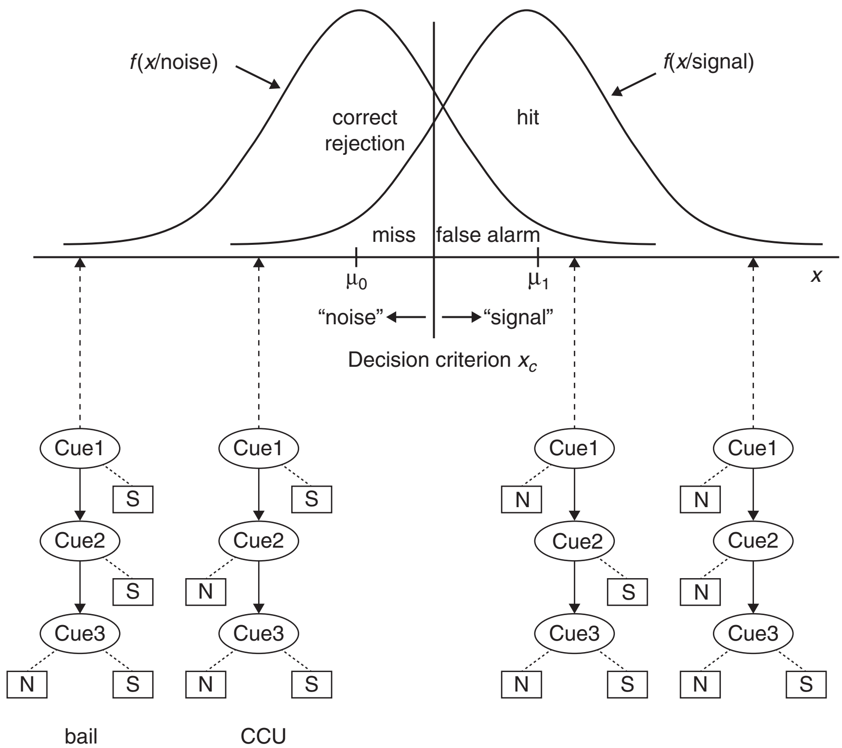

Figure 24.3 *Balancing misses and false alarms (after Luan et al., 2011).*
In signal detection theory, the balance between misses and false alarms is set by a decision criterion. In fast-and-frugal trees, the balance is set by the exit structure, where "S" stands for "signal" and "N" for "noise." A fast-and-frugal tree with three ordered cues has four possible exit structures. The two on the left correspond to the two trees in Figure 24.2.

and false alarms. To minimize false alarms over misses, however, one would have to choose one of the two trees on the right.

In sum, intuitive design honors transparency and usability. Moreover as illustrated by the CCU tree, it can also lead to equally good or better classification than complex, nonintuitive statistical decision models.

Controversies

The Rationality Debate

What is known as the *rationality debate* concerns the question whether humans are rational or not, and what rationality means in the first place. Kahneman and Tversky (1996) argued that the rules of logic and probability universally define rationality, that people lack rationality, and that people are hardly educable out of their "cognitive illusions." Others added that deviations from logical rules would be associated with substantial costs, from obesity to HIV to the financial crisis, and called on governments to “nudge” their citizens into proper behavior (Thaler & Sunstein, 2008). Governments across the globe increasingly follow this “libertarian paternalism.”

In contrast, Gigerenzer (1996, 2015) and others conjectured that logical rules cannot provide universal yardsticks for rationality in real-world situations of uncertainty, and that human behavior should be evaluated in terms of whether it reaches its goals (ecological rationality). Moreover, many so-called cognitive illusions (including the base-rate fallacy, conjunction fallacy, and overconfidence) have been shown to reflect reasonable judgments under uncertainty (Gigerenzer, 2018), and automatically following logical rules is not necessarily an intelligent strategy. Furthermore, experiments show that education and training can help to improve reasoning, while there is a general lack of evidence that violations of logical rules are associated with substantial costs (Arkes et al., 2016).

The status of heuristics is essential to this debate. In the heuristics-and-biases program, heuristics grew to be associated with biases and irrationality, and logic and probability theory are taken as a universal definition of rationality both in situations of risk and under uncertainty. In contrast, when labels such as availability were replaced by formal models of heuristics (Table 24.1), it could be shown that simple heuristics often make more accurate predictions than complex strategies that use more information and calculation. This is called a *less-is-more effect*, an example of which is the higher accuracy of the CCU tree over the logistic regression decision system. In general, the two views differ in whether they take heuristics and uncertainty seriously.

Decision-Making and Intelligence: Toward Theory Integration

Open a book on intelligence and you will likely find little if anything on decision-making. By the same token, pick up a book on decision-making and you will likely look in vain for the term *intelligence*. This mutual ignorance is surprising, given that making good decisions should in some way involve intelligence.

The Two Disciplines

During the twentieth century, psychology evolved into two separate disciplines, which Lee J. Cronbach (1957) called the “Tight Little Island” of experimental psychology and the “Holy Roman Empire” of correlational psychology. The study of intelligence traces back to the nineteenth-century work of Francis Galton and Karl Pearson on individual differences in “natural ability,” later named *intelligence*, and relies heavily on correlational methods developed by Galton. The psychological study of decision-making, however, became associated with experimental psychology and emerged mainly in the second half of the twentieth century. This historical split is one reason for the surprising lack of interaction, but there is also a psychological reason. Many psychologists show strong in-group behavior, creating and defending subdisciplines, which leads to ignoring what others write on similar topics. As Walter Mischel (2008) put it, “Psychologists treat other people’s theories like toothbrushes – no self-respecting person wants to use anyone else’s.”

Yet one of psychology’s vital goals is *theory integration*, that is, to link the available theories, concepts, and phenomena into a common network. In 2017, the journal *Decision* published two issues to launch this integration program. The integration between signal detection theory and fast-and-frugal trees is one successful example (Figure 24.3). Integration between research on intelligence and decision-making, however, has rarely been considered. Here are a few thoughts.

Theory Integration

Let us begin with decision-making under risk. What is the notion of intelligence underlying this program? It appears to be logical and calculative intelligence: complying with logical axioms, checking and maintaining coherence, and maximizing expected utility. Expertise and knowledge plays little to no role, and individual differences are not the focus, apart from exceptions such as the distinction between risk averse and risk seeking. A test analogous to the IQ test for decision-making under risk would measure the degree to which people exhibit coherence in their choices, similar to how coherence is measured in Bayesian judgments (Berg et al., 2016). This would lead to research questions such as, is IQ associated with coherence?

For decision-making in situations of uncertainty, intelligence has a broader function than maintaining coherence in a world of certainty. In the words of the late psychologist Jerome Bruner, intelligence means going beyond the information given. The heuristics described in this chapter go beyond this information because what is known is not sufficient. In decisions under uncertainty, intelligence entails inference, that is, making informed bets using heuristics.

Bröder (2012) studied the question of whether people with higher IQs rely on more complex strategies when making inferences under uncertainty. The answer is no, which is consistent with the finding that experts often rely on simple heuristics because they know what to ignore. A second question is whether people with higher IQs better know what heuristic to choose in what situation, that is, have better intuitions about the ecological rationality of heuristics. Here, the answer is yes; there appears to be a correlation between the adaptive use of heuristics and IQ (Bröder, 2012). Thus, intelligence could be understood as the ability to find the right heuristic for a given problem.

The integration between decision research and intelligence research could start by exploiting the specific strengths of each program. Research on intelligence focuses on individual differences; decision-making research could adopt this perspective to acquire a better understanding of the heterogeneity in people’s (adaptive) use of heuristics. These individual differences exist, as illustrated earlier: Entrepreneurs differ in the aspiration level they set when deciding where to invest and customers differ in the heuristics they use for food choice or consumer choice in general.

One strength of decision-making research is that it builds and tests models about processes, such as search rules, stopping rules, and decision rules in fast-and-frugal trees. Research on intelligence could consider building process models into their theories. The transparency and usability of these models can add a practical dimension to intelligence. Like the adaptive toolbox with its multiple heuristics, theories of intelligence often postulate multiple intelligences, which provides another point of integration. What decision-making research can learn is to take individual differences seriously. What intelligence research can learn is to take the difference between risk and uncertainty and the analysis of heuristic processes seriously.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulates et axiomes de l'ecole américaine [Rational man's behavior: Critique of the American School's postulates and axioms]. *Econometrica 21*, 503–546.

Arkes, H. R., Gigerenzer, G., & Hertwig, R. (2016). How bad is incoherence? *Decision*, *3*, 20–39. https://doi.org/10.1037/dec0000043

Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.

Berg, N. (2014). Success from satisficing and imitation: Entrepreneurs' location choice and implications of heuristics for local economic development. *Journal of Business Research*, *67*, 1700–1709. https://doi.org/10.1016/j.jbusres.2014.02.016

Berg, N., Biele, G., & Gigerenzer, G. (2016). Consistent Bayesians are no more accurate than non-Bayesians: Economists surveyed about PSA. *Review of Behavioral Economics*, 3, 189–219. https://doi.org/10.1561/105.00000034

Bernoulli, D. (1738/1954). Exposition of a new theory on the measurement of risk. *Econometrica*, *22*, 23–36. https://doi.org/10.2307/1909829

Boyd, R., & Richardson, P. J. (2005). *The origin and evolution of cultures*. New York: Oxford University Press.

Bröder, A. (2012). The quest for take-the-best. In P. M. Todd, G. Gigerenzer, & ABC Research Group, *Ecological rationality: Intelligence in the world* (pp. 216–240). New York: Oxford University Press.

Clark, G. (2002). Embracing fatality through life insurance. In T. Baker & J. Simon (Eds.), *Embracing risk: The changing culture of insurance and responsibility* (pp. 80–96). Chicago: University of Chicago Press.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671–684.

Daston, L. (1987). The domestication of risk: Mathematical probability and insurance 1650–1830. In L. Krüger, L. Daston, & M. Heidelberger (Eds.), *The probabilistic revolution*, vol. 1, *Ideas in history*. Cambridge, MA: MIT Press.

Daston, L. (1988). *Classical probability in the Enlightenment*. Princeton: Princeton University Press.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571–582. https://doi.org/10.1037/0003-066X.34.7.571

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naïve diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, *22*, 1915–1953. https://doi.org/10.1093/rfs/hhm075

Dhami, M. K. (2003). Psychological models of professional decision making. *Psychological Science*, *14*, 175–180. https://doi.org/10.1111/1467-9280.01438

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, *75*, 643–669. https://doi.org/10.2307/1884324

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review*, *103*, 592–596. https://doi.org/10.1037/0033-295X.103.3.592

Gigerenzer, G. (2014). *Risk savvy: How to make good decisions*. New York: Viking.

Gigerenzer, G. (2015). On the supposed evidence for libertarian paternalism. *Review of Philosophy and Psychology*, *6*, 363–383. https://doi.org/10.1007/s13164-015-0248-1

Gigerenzer, G. (2016). Towards a rational theory of heuristics. In R. Frantz & L. Marsh (Eds.), *Minds, models, and milieux: Commemorating the centennial of the birth of Herbert Simon* (pp. 34–59). New York: Palgrave Macmillan.

Gigerenzer, G. (2018). The bias bias in behavioral economics. *Review of Behavioral Economics*, 5, 303–336. http://dx.doi.org/10.1561/105.00000092

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Garcia-Retamero, R. (2017). Cassandra's regret. The psychology of not wanting to know. *Psychological Review*, *124*, 179–196. https://doi.org/10.1037/rev0000055

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *104*, 650–669. https://doi.org/10.1037/0033-295X.103.4.650

Gigerenzer, G., & Goldstein, D. G. (2011). The recognition heuristic: A decade of research. *Judgment and Decision Making*, *6*, 100–121.

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.) (2011). *Heuristics: The foundations of adaptive behavior*. New York: Oxford University Press.

Gigerenzer, G., Todd, P. M., & ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75–90. https://doi.org/10.1037/0033-295X.109.1.75

Green, L, & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice*, *45*, 219–226.

Hamlin, R. P. (2017). "The gaze heuristic": Biography of an adaptively rational decision process. *Topics in Cognitive Science*, *9*, 264–288. https://doi.org/10.1111/tops.12253

Hertwig, R., & Hoffrage, U. (2013). Simple heuristics: The foundations of adaptive social behavior. In R. Hertwig & U. Hoffrage (Eds.), *Simple heuristics in a social world* (pp. 3–36). New York: Oxford University Press.

Hutchinson, J. M. C., & Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes*, *69*, 97–124. https://doi.org/10.1016/j.beproc.2005.02.019

Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science*, 302, 1338–1339.

Kahneman, D. (2011). *Thinking fast and slow*. London: Allen Lane.

Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291. https://doi.org/10.2307/1914185

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582–591. https://doi.org/10.1037/0033-295X.103.3.582

Knight, F. (1921). *Risk, uncertainty and profit*, vol. 31. Boston: Houghton Mifflin.

Luan, S., Schooler, L., & Gigerenzer, G. (2011). A signal detection analysis of fast-and-frugal trees. *Psychological Review*, *118*, 316–338. https://doi.org/10.1037/a0022684

Martignon, L., Vitouch, O., Takezawa, M., & Forster, M. R. (2003). Naive and yet enlightened: From natural frequencies to fast and frugal decision trees. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 189–211). Chichester, UK: Wiley.

McBeath, M. K., Shafer, D. M., & Kaiser, M. K. (1995). How baseball outfielders determine where to run to catch fly balls. *Science*, *268*, 569–573. https://doi.org/10.1126/science.7725104

Mischel, W. (2008). The toothbrush problem. *Observer*, *21*(11). www.psychologicalscience.org/observer/the-toothbrush-problem

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.

Pohl, R. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, *19*, 251–271. https://doi.org/10.1002/bdm.522

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112, 610–628. https://doi.org/10.1037/0033-295X.112.3.610

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology*, *83*, 1178–1197. https://doi.org/10.1037/0022-3514.83.5.1178

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118. https://doi.org/10.2307/1884852

Simon, H. A. (1979). *Models of thought*. New Haven, CT: Yale University Press.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*, 1–19. https://doi.org/10.1146/annurev.ps.41.020190.000245

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*, 672–695.

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. New York: Random House.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Todd, P. M., Gigerenzer, G., & ABC Research Group (2012). *Ecological rationality: Intelligence in the world*. New York: Oxford University Press.

Todd, P. M., & Miller, G. F. (1999). From pride and prejudice to persuasion: Realistic heuristics for mate search. In G. Gigerenzer, P. M. Todd, & ABC Research Group, *Simple heuristics that make us smart* (pp. 287–308). New York: Oxford University Press.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*, 105–110. https://doi.org/10.1037/h0031322

25 Artificial Intelligence

Ashok K. Goel and Jim Davies

Artificial intelligence (AI) is the field of research that strives to understand, design, and build intelligent computational artifacts. From computer programs that can beat top international grand masters at chess to robots that can help detect improvised explosive devices in war, from intelligent agents that can answer questions in customer service to computing systems that can automatically detect credit card fraud, AI has had many well-known successes. In fact, modern societies are based on AI systems; without AI agents, advanced industrialized economies may quickly grind to a halt (Kurzweil, 2005).

As the above examples illustrate, the field of AI has a very broad scope. However, two features unify all of AI as a discipline. First, AI is united in the core belief that intelligence is a kind of computation. Thus, although in principle the design of any intelligent artifact might be classified as an AI, in practice AI agents are almost always computers or computer programs, and AI laboratories typically are found in departments of computer science. Second, its main methodology is the exploration of the principles of intelligence by building computational artifacts.

Broadly speaking, AI includes three large subfields: *robotics, machine learning*, and *cognitive systems*. Figure 25.1 depicts the relationship between AI and the three subfields. Robotics deals with embodied AI agents that interact with the physical world. Thus, action and perception are important parts of robotic AI agents: A robot that can detect improvised explosive devices, for example, must be able to sense and perceive the physical world as well as act on it. Machine learning typically pertains to computer programs that can detect and exploit patterns in data. To detect a fraudulent credit card transaction, for example, a computer program may first need to be trained on a set of credit card payments including both regular and irregular purchases so that the program can learn patterns indicative of a fraud. Cognitive systems, such as game-playing programs (for example, DeepBlue and AlphaGo) and conversational agents (for example, Siri and Watson) pertain to higher-level cognition, and interact with human and social worlds. Research on cognitive systems sometimes is also called cognitive computing.

Of course, the three subfields of AI – robotics, machine learning, and cognitive systems – have considerable overlap. For example, the emerging area of human-robot interaction combines cognitive systems and robotics. Thus, a robot for detecting

Acknowledgments: We thank the editors of this *Handbook of Intelligence* as well as members of the Design and Intelligence Laboratory at Georgia Institute of Technology for their comments on earlier drafts of this chapter. An earlier version of this chapter appears in the 2011 edition of the *Handbook*.

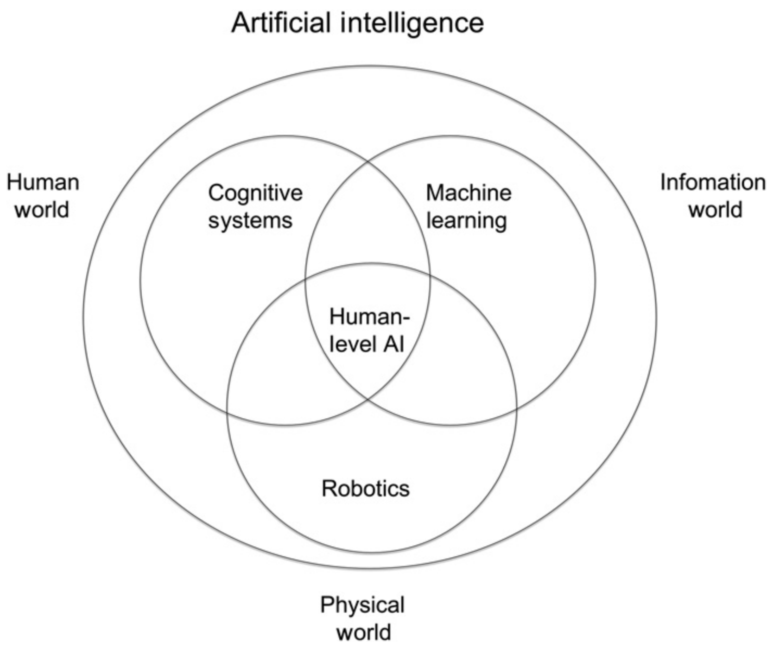

Figure 25.1 *The field of artificial intelligence includes the subfields of robotics, machine learning, and cognitive systems.*

improvised explosive devices may have an embedded cognitive system for conversing with humans about its findings, receiving instructions, and asking questions. At the intersection of cognitive systems and machine learning, an interactive game-playing AI agent might have been trained on data from a large number of previously played games. Similarly, at the intersection of robotics and machine learning, a robot may learn a policy for taking action in the world through many trials. The center of Figure 25.1 indicates the intersection of all of these subfields, which describes work on general intelligence, including the construction of human-level AI dreamt by the founders of the field (McCarthy et al., 1955/2006). General AI likely will require major advances in all three subfields: robotics, machine learning, and cognitive systems.

It is also helpful to distinguish AI research into two main paradigms. *Engineering AI* is concerned with how to design the smartest intelligent artifacts possible, regardless of whether the processes implemented reflect those found in people (or other animals). The vast majority of AI research on robotics and machine learning falls into this category. *Psychological AI*, in contrast, endeavors to design artifacts that think the way people do (or sometimes groups of people do). Much, but not all, research on cognitive systems belongs to this paradigm, though it is possible to design cognitive systems, such as Siri and Watson, that interact with humans but do not necessarily reason as people do. In this chapter, we will focus mostly on the paradigm of psychological AI because the original dream of AI was to develop human-level intelligence, this handbook is intended for an audience of cognitive scientists, and we ourselves work in this paradigm.

AI and Cognitive Science

Cognitive science engages both AI and psychology. Cognitive psychology is mostly concerned with understanding of intelligence found naturally in humans and other animals, whereas, in addition, AI is interested in the understanding of intelligence in agents it designs. From the AI perspective, the concept of intelligence is not limited to the abilities of humans or even animals in general, but covers potentially any kind of intelligent system, be it human, computer, animal, or alien. Albus (1991, p. 474) put it eloquently: "A useful definition of intelligence . . . should include both biological and machine embodiments, and these should span an intellectual range from that of an insect to that of an Einstein, from that of a thermostat to that of the most sophisticated computer system that could ever be built."

AI and cognitive psychology have a rich two-way relationship. In one direction, cognitive psychology has often inspired AI theories, and AI systems have acted as testbeds for experimenting with and evaluating the theories. Examples include theories of schema (Piaget, 1952), mental models (Craik, 1943), and learning by imitation (Tomasello, 1999), whose origins lie in psychology but have deeply influenced AI research. In the other direction, AI research has resulted in theories of cognition to be tested through psychological experimentation. Examples include semantic networks (Quillian, 1968), scripts (Schank & Abelson, 1977), and Bayesian networks (Pearl, 1988). AI models of cognition differ from other models in psychology in that AI models always implement *information-processing* theories. That is, the theory describes intelligence in terms of the content, representation, access, use, and acquisition of information, as opposed to, say, a statistical model of the influences of age or nutrition on IQ in a population.

Critics of AI from psychology sometimes view many AI programs as being psychologically implausible. Indeed, psychological claims of AI theories typically are under-constrained by empirical human data, and thus, for the most part, these criticisms of AI from psychology are not inaccurate. Most AI is engineering AI, and even psychological AI must go out on limbs simply because there are just not enough data to constrain all the choices AI scientists need to make. However, AI contributes to the understanding of intelligence in several ways (Davies & Francis, 2013).

First, although they can be under constrained, *AI programs demonstrate what kinds of data need to be collected*. Because AI programs work at a very precise level of detail, they bring to light theoretical ambiguities that psychology might not immediately or explicitly acknowledge. For example, it is one thing to say that a person can only comprehend one speaking voice heard at a time; it is quite another to create a computer implementation of this attentional effect – to do so requires making decisions about the interaction and influences of volume, which one voice you are listening to first, what factors affect attentional switching, among many other issues. The level of detail that makes AI programs under-constrained is the very quality that brings to light previously unrecognized factors.

Humans obviously have only limited information and information-processing resources, and, thus, their rationality is intrinsically bounded (Simon, 1996). However, it is also true that many cognitive problems people routinely solve are computationally intensive. For example, deciding how to design a poster for a concert offers more possibilities than can possibly be considered. *AI approaches to solving these kinds of problems shed light on what ways will not work*. If AI shows that a means for solving a problem will take too long to be practical, then AI has shown that people cannot be doing it that way, at least not routinely.

On the other hand, *AI can show that certain methods are possible*. Though showing that something is possible is far from proving that it *is*, many current theories in psychology do not have such proof. AI serves a valuable function as creating proofs-of-concept.

Another thing AI is particularly good at is *exploring the benefits and limitations of various ways to represent and organize knowledge in memory*. Many of these benefits are clear only when dealing with a strict information-processing level of detail. Are beliefs represented as words, pictures, or something else? Given all of the cognitive tasks memories are supposed to contribute to, AI is in a good position to shed light on such issues. As we will describe in more detail later, this subfield of AI is known as "knowledge representation."

Finally, once there is an AI program that resembles some part of human thinking to a researcher's satisfaction, *it is possible to run experiments on the program that are either unethical or too expensive (in terms of time or money) to run on living beings*. In simulation you can run thousands of experiments in a day, with exquisite control over all variables.

If AI theories of cognition are under-constrained, theories of AI with roots in cognitive psychology can be over-constrained (Langley, 2012). In general, there is no particular reason why an AI cognitive system must imitate each and every microstructure of cognition to manifest intelligent behavior. As an example, while the notion of a working memory originating in cognitive psychology is important in AI for, among other things, focusing attention, the size of the working memory in AI cognitive systems need not be limited to the famous seven plus or minus two typical of adult humans.

In both cognitive psychology and AI cognitive systems, researchers over the years have tried to build theories of intelligence of two different kinds at two different levels of abstraction: the symbolic and the sub-symbolic. Symbols represent conceptual abstractions of the world, such as *dog* or *justice*, and act as pointers both to the world and to one another; thus, symbolic processing pertains to conceptual information processing. In contrast, sub-symbolic processing tends to deal with information at a finer grain, such as the pixels of an image, the weight of an association, and the probability of the truth of a proposition. Sub-symbolic systems are difficult to program completely by hand, so they are often machine learning systems as well. Connectionism is the dominant sub-symbolic modeling paradigm in cognitive science, and works by applying neural networks to psychological problems. Symbolic cognitive systems, in contrast, may organize symbols into propositions that can be used in reasoning processes, such as deduction, induction, and various transformations. Many language processing systems work this way, as well as systems that use commonsense reasoning to try to understand the physical world. Many symbolic systems use "qualitative reasoning," a term that distinguishes it from the more numerically represented memory of sub-symbolic systems. There are, of course, many hybrid systems that use both symbols and sub-symbols, such as Bayesian belief networks.

Navigational Planning: An Illustrative Example

We want to illustrate a simple example of AI in some detail to help make this discussion more concrete. Let us suppose that Sunny, a cheerful AI agent, is about to start a new job in a new city. Sunny starts its car at its apartment and needs to navigate to an office building downtown. How might Sunny think and what might Sunny do, given that this is its first day in the city and it has never been to the office building? Our goals in this section are to explain some dimensions in designing AI agents as well as describe some issues in putting multiple capabilities into an AI agent.[1]

Action, Perception, and Cognition

To reach its office from its apartment, Sunny might use one (or more) of several possible strategies. For example, it might drive its car a short distance in some direction and then see if it has reached the office building. If it has, then it has accomplished its goal. If it has not, then it might again drive a short distance in some direction, and then again see if it has reached the building. Sunny could repeat this process until it reaches its goal. Blindly moving about like this would likely take a very long time, but in terms of internal processing, this method is very efficient. This *perceive-act* internal computational processing, called *situated action* (or *reactive control*; Arkin, 1999), works by perceiving the immediate environment, acting based on those perceptions, and then repeating. The computational processing in reactive control is very efficient and does not require the agent to be able to store new memories. However, depending on the environment and the goal, it may produce needlessly complicated external behavior because Sunny could be driving short distances in arbitrary directions for a very long time before it reaches its goal. In fact, this strategy does not guarantee that the goal will ever be reached.

Alternatively, when Sunny starts at its apartment, it might simply ask Honey, a sweet AI agent who happens to be passing by, how to reach the office building. Honey, a longtime resident of the city, might give Sunny detailed directions, which Sunny could simply follow. In contrast to the previous strategy, this strategy produces very efficient output behavior: Assuming that Honey's directions are good, Sunny should reach its goal quite efficiently. However, this strategy of *asking* requires a society of intelligent agents (human or AI), each with different knowledge. It also requires

[1] Much of our discussion of this problem is based on the work of the first author and his students in the 1990s when they developed a computer program called Router for addressing this class of problems (Goel et al., 1994) and instantiated Router on a mobile reactive robot called Stimpy (Ali & Goel, 1996). They also developed a knowledge-based shell called Autognostic for learning by reflection on the Router program embodied in Stimpy (Stroulia & Goel, 1999), as well as reflection on Stimpy's reactive controller (Goel et al., 1997).

a culture in which Sunny may in fact approach Honey for directions; Honey might in fact stop to help Sunny, and the two can communicate in a shared language; Sunny might trust Honey, a total stranger, enough to follow its directions in a new city; and so on. AI research on robot societies and human-robot interaction is in its early stages, and so here we will briefly mention only a small set of selected issues.

How can Sunny and Honey talk with each other? How can Sunny talk with a human? Understanding and generating natural language is the goal of the AI subdiscipline of *natural language processing* (NLP). Researchers in the area of natural language understanding take written text or spoken language and create accurate knowledge representations reflecting the meaning of the input. Natural language generation works roughly in the reverse – starting with some meaning and generating appropriate words and speech to communicate it; this has received much less attention in AI. Two robots might be able to share knowledge very efficiently if that knowledge is represented in the same way. However, there is little agreement in AI over how knowledge should be represented in general (the linguistics subfield of semantics, similarly, has no consensus on meaning representation). Different knowledge representations appear to be better for different tasks.

When Honey gives advice, how is Sunny to know whether that advice is plausible? Except for limited environments, this problem seems to require general *commonsense reasoning*, a field closely related to knowledge representation. It is a widely held belief that most computer programs' lack of common knowledge and inability to reason with it effectively are major problems for much of AI. The subfield of commonsense reasoning endeavors to overcome this challenge. The most famous is the Cyc project (Lenat & Guha, 1990), a major project to manually encode all human commonsense knowledge. More recent strategies include Web-based knowledge collection methods, such as OpenMind Commonsense (Singh et al., 2002) and Peekaboom (von Ahn, Liu, & Blum, 2006).

Here is another strategy by which Sunny may reach its office building: Let us suppose that when Sunny was originally built in an AI laboratory, it was bootstrapped with some knowledge. Some of this knowledge may have been *heuristic* in its content and encoded in the form of a *production rule*. A heuristic is a "rule of thumb," and a production is an "If x then do y" kind of rule. So, for example, Sunny might be bootstrapped with the knowledge that "if the goal is to reach downtown in a city, then move in the direction of the tallest buildings." This knowledge directly uses the goal (reaching downtown) to suggest a high-level action (move in the direction of the tallest buildings) and is heuristic in its nature since it may not correctly apply in all cities. If Sunny had this knowledge, then it might begin by perceiving the environment around it, locating the tallest buildings in the horizon, deciding to head in their direction, and moving toward them. When it reaches the next intersection, Sunny might again locate the tallest buildings relative to its current location, change its direction if needed, and so on. This strategy of *perceive-think-act* not only requires some knowledge but also must use more complex internal processing than the simpler perceive-act strategy of situated action. On the other hand, depending on the environment, perceive-think-act may result in a far simpler external behavior because now the behavior is more explicitly directed by the goal.

This kind of strategy can be implemented as a *production system* (Newell & Simon, 1972), which represents "what to do," or procedural knowledge, with if-then rules. In Sunny's case, the rules dictate physical action in the environment. Production systems are often used for making changes in memory as well. Rules can add, change, and remove goals and elements in memory. Surprisingly complex behavior can result with this method. This particular approach has been very successful in cognitive modeling for some problems. Well-known cognitive architectures such as Soar (Laird, 2012; Laird, Newell, & Rosenbloom, 1987) and ACT-R (Anderson, 2013; Anderson & Lebiere, 1998) are production systems at their core.

However, cognitive architectures such as SOAR and ACT-R have declarative as well as procedural knowledge. Declarative knowledge is often represented as *frames* (Minsky, 1975) or semantic networks, and is used by the productions (the procedural knowledge). Frames are similar to classes in object-oriented programming: They define a class of entities and what attributes they have. Instances of these frames take particular values for these attributes. For example, the frame for PERSON might contain the attributes NAME and AGE, and an instance of person might have a NAME of "Julie" and an AGE of "45." Like frames, *semantic networks* (Sowa, 1987) are a widely used representation scheme in AI. One can imagine a semantic network as a map of concepts, with nodes representing concepts (such as MAN and DOG) and labeled links between them (labeled, for example, with OWNS). Frames and semantic networks are thought to be informationally equivalent, which means that there is no loss of information when translating from one to another.

Another long-standing and still very strong area of AI is representation and processing based on *logic*. Logic is used for inference but has also been adapted for use in many other specific tasks, such as theorem proving (McCarthy, 1988).

Let us consider one other strategy for Sunny's task before we move on to the next topic: Sunny might consult a map of the new city. The important characteristics of a city map in this context are that they are an external representation of the world (i.e., it is not stored internally in Sunny) and that it is a visuospatial model of the world (i.e., there is a one-to-one structural correspondence between selected spatial objects and relations in the world and the objects and relations on the map; see Glasgow, Narayanan, & Chandrasekaran, 1995). Sunny can use this map to plan a navigation route to the office building and then execute the plan. This too is a perceive-think-act strategy. However, as compared to the heuristic method, the "thinking" in this strategy uses very different content and representation of knowledge.

Once Sunny has studied the map, it has some version of it stored in its memory. When Sunny needs to navigate to a location on the map, it can refer to the map. Finding a route on a map is not trivial, however. At each intersection, a choice must be made. One of the first insights of the field was that a great many cognitive problems can be solved by systematically evaluating available options. This method of searching through a space of choices is applicable in many domains and is still widely used. Researchers focusing on *search* compare the various search methods that have been invented and describe the classes of problems to which each is most applicable. Because most interesting search spaces are enormous (e.g., there are more possible chess game configurations than there are atoms in the universe), researchers invent *heuristics* to guide the AI to explore the more promising areas of the search space. One problem for which search has been particularly useful is in *planning*, which is the generation of an ordered sequence of actions prior to actually executing those actions.

The internal processing in reading a map might be more costly than the processing in a heuristic search; however, depending on the environment, this strategy might lead to a solution that has a better chance of success – for example, the solution generated by this model-based method is less likely to get stuck in some cul-de-sac than the solution generated by the heuristic method. Of course, we can easily think of several other strategies for addressing Sunny's task, especially in today's world of the Internet and the global positioning system.

These examples make clear some of the dimensions of designing an AI agent. First, an AI agent lives in some environment, and what and how an agent can think depends in large part on the environment in which the agent lives. Some environments might contain other agents, who may be cooperative, competitive, or combative. Some environments are dynamic. Some environments are only partially observable. Some environments are nondeterministic, and so on. One of the many contributions of AI is a more precise characterization and analysis of different kinds of environments, though much of the AI analysis so far has focused mostly on physical, not social, environments. Second, an agent might have access to different kinds of knowledge contents and representations. The knowledge may be engineered or acquired. The representations can be internal or external. The knowledge contents range from nil to heuristic rules to detailed, high-fidelity models of the environment. Another major AI contribution is a more precise and detailed account of knowledge contents and representations. Third, different strategies lead to very different trade-offs among knowledge requirements, the computational efficiency of internal processing, and the quality of generated solutions and behaviors. Yet another contribution of AI is more precise enumeration and analysis of these trade-offs.

Reasoning, Learning, and Memory

So far we have talked only about what our hypothetical AI agent, Sunny, might think and do when trying to reach its office for the first time. However, because Sunny is an AI agent, it might also learn from its interactions with the environment. What and how might Sunny learn from its experiences? Sunny acquires a new experience each time it interacts with the environment, including navigating from its apartment to its office, talking with Honey, and so on, irrespective of what internal strategy it uses. Further, to the degree to which Sunny's internal processing is accessible to it, it may also acquire an internal experience each time it does internal processing. In addition, when Sunny executes a plan or an action on the environment, the environment might provide it with feedback. This feedback might come immediately after the execution of an action (e.g., taking a turn at an intersection and getting caught in a cul-de-sac), or after a series of actions (e.g., taking a sequence of turns and reaching the goal).

The feedback might simply be the outcome of a plan – success or failure – or it might contain more information, for example, a specific action in the plan failed because it led to a cul-de-sac. Thus, an experience might contain not only an interaction with the environment but also some feedback on the interaction, and perhaps also a trace of the internal processing in that interaction.

Sunny might potentially learn many different things from its experiences in the environment. For example, Sunny might simply encapsulate experiences as *cases* and store them in memory for reuse in the future – the AI equivalent to episodic memory. On the first day, for example, Sunny might use a map to plan a navigation route and then execute the plan in the environment, as indicated in the section Action, Perception, and Cognition. The next day, when Sunny again faces the task of navigating to its office from its apartment, it might find a solution simply by retrieving the navigation plan in the case acquired from the previous day rather than relying on general-purpose knowledge and rules. This is called *case-based reasoning* (Kolodner, 1993). This approach views reasoning largely as a memory task, that is, as a task of retrieving and modifying almost correct solutions from memory to address the current problem. Related subdisciplines of cognitive science studying similar phenomena are exemplar-based reasoning, memory-based reasoning, instance-based reasoning, and analogical reasoning.

As Sunny learns from its experiences, its internal processing as well as its external behaviors can change. Initially, for example, Sunny might use a map of the environment for navigating through the new city. However, as it navigates through the world and stores its experiences as cases in its memory, it can increasingly generate new navigation plans by case-based reasoning. However, as the number of cases in memory increases, the cost of retrieving the case appropriate for a new problem also increases. Thus, again, each reasoning strategy offers computational trade-offs among knowledge requirements, processing efficiency, and solution quality.

More generally, AI typically thinks of each strategy for action selection discussed in the section Action, Perception, and Cognition as setting up an associated learning goal, which in turn requires a corresponding strategy for learning from experiences. Let us suppose, for example, that Sunny uses the strategy of situated action for action selection. It might, for example, use a table (called a *policy*) that specifies mappings from percepts of the world into actions on it. Then, from the feedback, or the reward, on a series of actions, Sunny can learn updates to the policy so that over time its action selection is closer to optimal. This is called *reinforcement learning* (Sutton & Barto, 1998). Note that if the series of actions results in success, then the reward will be positive; otherwise it is negative. Reinforcement learning is an especially useful learning strategy when the reward is delayed, that is, it comes after a sequence of actions rather than immediately after an action so that it is not clear what specific action in the sequence was responsible for the success or failure. Alternatively, suppose that Sunny employs the strategy of using production rules such as "If x then do y" to select actions. In this case, Sunny can use the learning strategy of *chunking* (Laird et al., 1987) to learn new rules from its experiences over time. Thus, just as AI has developed many reasoning strategies for action selection, it has developed many learning strategies for acquiring the knowledge needed by the reasoning strategies. Further, just like the reasoning strategies, the learning strategies too offer trade-offs among knowledge requirements, computational efficiency, and solution quality.

Most of the methods described thus far fall roughly into a category that can be described as "symbolic" approaches, characterized by the manipulation of qualitative, recognizable, discrete symbols. Another broad approach, as we mentioned earlier, is sub-symbolic. Though the border between these two approaches is fuzzy, we can think of a symbolic representation having a symbol for the letter "R" and a sub-symbolic system representing the letter with the dots that make it up on a screen. Since the dots, or pixels, are not meaningful in themselves, they are thought to be at a level of description below the symbol. The rest of the methods described in this subsection tend to use sub-symbolic representations.

So far we have assumed that Sunny has perfect knowledge of the environment, even if that knowledge is limited. However, many real-world domains involve uncertainty, and AI methods based on *probability* have been very successful at working in these environments. Probability theory has been used in many algorithms that use *hidden Markov models* to predict events based on what has happened in the past. Hidden Markov models are mathematical representations that predict the values of some variables given a history of how the values of these and other variables have changed over time (Raibiner & Juang, 1986). Probabilities are also used to determine beliefs, such as how likely it is that a street Sunny wants to use has been closed, given that the rain in that part of the city was 80 percent likely to have been freezing. Bayes' rule is useful for determining such conditional probabilities of some events (e.g., a road being closed) given the probability of others (e.g., freezing rain). *Bayesian belief networks* are mathematical representations that predict the probability of certain beliefs being true, given the conditional probabilities of other beliefs being true (Pearl, 2000). These networks are useful for updating probabilities of beliefs as information about events in the world arrives.

Statistics is the foundation of much of *machine learning*, a subdiscipline of AI that aims to create programs that use data and limited previous beliefs to create new beliefs. There are a great many kinds of learning algorithms, including artificial *neural networks*, which are the basis of connectionism in cognitive science (McCelland, Rumelhart, & the PDP Research Group, 1986; Rumelhart, McClelland, & PDP Research Group, 1986). Whereas most of the systems we've discussed process recognizable symbols, neural networks represent information at a sub-symbolic level (such as in pixels or bits of sound) as activations of nodes in a network. The processing of a neural network depends on how the nodes change each other's activations. The output of a neural network is an interpretation of the activations of certain nodes (for example, indicating whether or not a room is dark). *Genetic algorithms* are another means of computation that is (often) based on processing sub-symbolic representations. Inspired by the theory of biological evolution, genetic algorithms create solutions to problems by applying some fitness function to a population of potential solutions (Mitchell, 1998). Solutions with a high fitness are used to generate members of the next generation (often with some mutation or crossover of features), after which the process repeats.

Deliberation and Situated Action

Although we have briefly discussed situated action (reactive control) and situated learning (reinforcement learning), much of our discussion about Sunny, our friendly robot, pertained to deliberation. While AI theories of deliberative action selection typically are explicitly goal-directed, goals in situated action often are only implicit in the design of an AI agent. Deliberation and situated action in AI agents occur at different timescales, with deliberation typically unfolding at longer timescales than situated action. In general, designs of AI agents include both deliberative and situated components. For example, the design of Sunny, our friendly robot, might contain a deliberative planner that generates plans to navigate from one location in a city to another. Note that because there are many people and other robots working or walking on the roads, Sunny's environment is dynamic in that the state of the world can change during the time Sunny takes to generate a plan. How can Sunny navigate from its apartment to its office building in this dynamic environment?

Sunny of course can use the deliberative planner to plan a path between offices. However, while the planner can produce navigation plans, it might not represent the movements of all the people and other robots on the roads. So deliberation by itself is not good enough for the dynamic urban environment. Alternatively, Sunny can use situated action (i.e., *perceive-act*) that we described in the previous section. While this can help Sunny avoid collisions with moving people – as soon as Sunny senses the nearby presence of a person, it can move away – its progress toward the goal of reaching a specific office is likely to be slow, perhaps painfully so.

Yet another alternative is to endow Sunny with the capability of both deliberative planning and situated action. In fact, this is exactly what many practical robots do. As a result, Sunny becomes capable of both long-range planning and short-range reaction. It can use its deliberative planner to come up with a plan for reaching the office building. Then, as it is executing the navigation plan, it constantly monitors the world around it and acts to avoid collisions with moving people. Next, as soon as it has moved away from a collision, it reverts to execution of its navigation plan. In this way, Sunny combines both deliberation and situated action. While this integration of deliberation and situated action has obvious benefits, it also has additional knowledge requirements as well as additional computational costs of shifting between strategies.

So far we have talked of perceiving the environment as though it were a minor task. For human beings, perception often appears to be effortless, but automating perception in AI agents has proven to be one of the many difficult problems in AI. The field of *computer vison* creates programs that take images (such as photos and video) as input and generates beliefs about objects, textures, and movements, as well as higher-level features such as emotions, movement styles, and gender. *Speech recognition* is another major field in perception. The ability of computers to understand your credit card number when you speak it into the phone is the result of over fifty years of AI work. Many of the algorithms used to understand speech and sound are shared with those of machine learning.

Likewise, achieving physical motion in the real world is difficult. *Robotics* is the field of AI that controls machines that interact directly with the physical world (as opposed to a program that, say, buys stocks electronically). Robotics uses computational perception, machine learning, and sometimes natural language processing. Some of the major problems specific to robotics are navigation and the handling of objects. Robots can work in collaboration with each other; the field of *intelligent agents* or *agent-based AI* builds intelligent programs that operate through the interaction of many individual agents whereas in *swarm intelligence* the individual agents do not have much intelligence individually. For example, two intelligent robots cooperating to assemble a desk would be an example of agent-based AI, and a large number of simple agents, reacting to their environment only locally to find the fastest route, much as ants do, would be an example of swarm intelligence.

Deliberation and Reflection

We have briefly discussed the need for both longer-range planning and shorter-range situated action in autonomous AI agents because the environment in which they reside is dynamic. However, changes in the environments themselves can unfold over different timescales. In the short term, for example, people and robots might be moving around on the roads of Sunny's city. In the long term, roads themselves change, new apartments and office buildings are constructed, and other changes occur. Then the navigation plan that Sunny's deliberative planner produces will start failing on execution. How might Sunny adapt its knowledge of the environment as the environment changes? Alternatively, if Sunny had been designed incorrectly to begin with, how might it adapt its reasoning process?

Recent AI research on meta-reasoning is starting to design AI agents capable of self-adaptation (Cox & Raja, 2011). Such an AI agent might contain a specification of its own design. For example, the meta-reasoner in Sunny may have a specification of Sunny's design, including its functions (e.g., its goals) and its mechanisms for achieving the functions (e.g., the method of map-based navigation planning). When Sunny generates a plan that fails on execution, Sunny's meta-reasoner uses the specification of its design to diagnose and repair its reasoning process. If the feedback from the world on the failed plan pertains to an element of knowledge (e.g., at intersection A, I expected a road going directly toward downtown but when I reached there, I found no such road), then Sunny enters this new knowledge in its map of the city. Thus, while the deliberative planner in Sunny reasons about actions in the external world, Sunny's reflective meta-reasoner reasons about its external world as well as its internal knowledge and reasoning.

AI Safety

If Sunny is going to be moving around the real world, there needs to be something about the design that will keep the robot from hurting people (or itself). As robots are often made of hard material, the very motion of the limbs has the potential for injury to humans, other animals, or sensitive environments.

A complex robot should not have simple goals without a set of values or preferences that will make sure that harm is not caused. Pushing a child out of the way might help make Sunny get to the desired location faster, but would be socially unacceptable; so we need to make sure that robots like Sunny recognize that this would likely cause harm and should not be done.

This is a challenging issue for many reasons. First, harm can come about in so many ways. Some harms, such as injury, are often immediate, and others, such as low-level ingestion of toxins, can take years to cause harm. Some harms are physical, and others, such as witnessing a horrific event, can cause psychological trauma.

Second, sometimes harms cannot be avoided. We don't want Sunny to cut someone open, but a surgical robot has to do just that. If Sunny is a self-driving car robot, it might have to make split-second decisions on who lives or dies. Suppose Sunny is moving at a high speed, and the road conditions change such that Sunny will strike, and probably kill, either an old person or a seven-year-old child, depending on which direction Sunny steers (it's too late to brake sufficiently). Which one should Sunny hit, and why? What if it's two old persons or one child? The calculation that Sunny must make needs to be fast, and have numerical values associated with human lives and suffering.

We will also have robots designed to kill people. The very existence of these machines is an ethical issue, and the programming that determines who the robot kills and doesn't kill is an obvious example of a serious moral decision. It is important that we recognize that programming ethics into robots is not merely a programming issue, but an interdisciplinary problem requiring contribution from law, philosophy, psychology, sociology, and other fields.

Putting It All Together

Figure 25.2 illustrates a high-level general architecture for an AI agent such as Sunny with many of the capabilities discussed above. The agent is situated in an external world: It can sense percepts in the world; it can also use its effectors to act on the world. At the bottom of the multilayered architecture is reaction, which directly maps percepts into actions. In the middle is deliberation, which unfolds more slowly than reaction and includes complex interactions among memory, learning, and reasoning. At the top level is metacognition, which monitors, controls, adapts, and explains the deliberative processing. This architecture helps us understand how AI agents like Sunny are designed such that all their capabilities work in synchrony.

In this section, we took navigational planning as an example to illustrate how AI is putting together multiple capabilities ranging from perception, cognition, and action, to reasoning, learning, and memory, and on to reflection, deliberation, and situated action. Of course, the design choices we have outlined are exactly that: choices. For example, instead of using deliberation to mediate between reflection and situated action as described above, an AI agent can reflect directly on situated action. In a way, the enterprise of AI is to explore such design choices and examine the computational trade-offs that each choice offers.

What has emerged out of this line of work is an understanding that the design of an AI agent depends on the environment it lives in, and that no one design is necessarily the best for all environments. Further, the design of an AI agent in any nontrivial

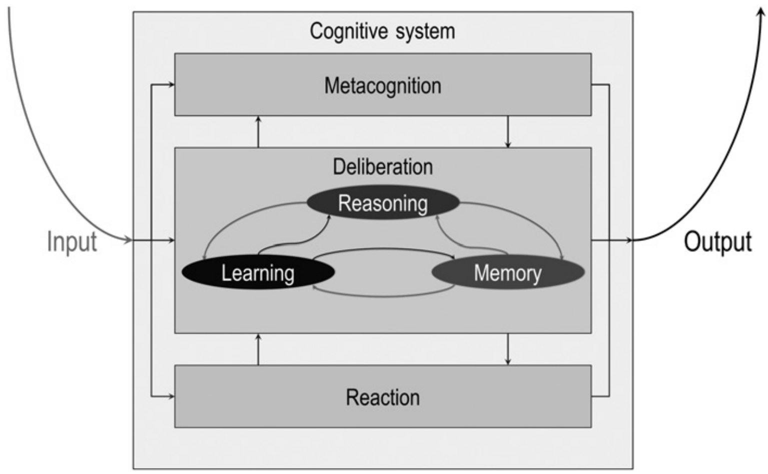

Figure 25.2 *A multilayered architecture for an AI agent, combining reaction, deliberation, and metacognition.*

environment requires multiple capabilities and multiple methods for achieving any capability such as reasoning and learning.

There is a large and growing literature on architectures for intelligent agents. Some of these architectures are inspired by cognitive psychology and are called cognitive architectures. Some of the better-known cognitive architectures include ACT-R (Anderson, 2013; Anderson & Lebiere, 1998) and SOAR (Laird, 2012; Laird et al., 1987). Samsonovich (2010) and Kotseruba, Gonzalez, and Tsotsos (2016) provide useful surveys of cognitive architectures. Langley, Laird, and Rogers (2009) review some of the challenges in developing cognitive architectures. Laird, Lebiere, and Rosenbloom (2017) recently proposed a "common model" of intelligence based on work on cognitive architectures.

A Very Brief History of AI

Many people have an almost mystical view of intelligence. One result is that when an AI agent manages to accomplish some task, a common reaction is to claim that it is not an example of intelligence. Indeed, at one point in the history of computing, arithmetic calculation was thought to be one of the best displays of intelligence, but now almost no one wants to say a calculator is intelligent. Because of this moving of the goalposts, AI has been jokingly referred to as standing for "almost implemented." For the most part, this is only a semantic issue. In fact, although not always labeled AI, AI discoveries have revolutionized our world.

In the middle of the twentieth century, the scientific world experienced a shift in focus from descriptions of matter and energy to descriptions of information. One manifestation of information theory applied to real-world problems was in the field of *cybernetics* (Weiner, 1961), the study of communication and control in self-regulating analog systems. Cybernetics' focus on analog signal contributed to its losing ground against discrete symbolic approaches common in AI. Not only did the symbolic approaches come to dominate AI research, but the symbol-processing approach came to dominate cognitive psychology as well.

Search was the first major paradigm of AI. The first artificial intelligence program ever written is the Logic Theorist (Newell, Shaw, & Simon, 1958). Many of the problems early AI researchers focused on were, in retrospect, simple. The early exuberance of AI was tempered with the first "AI winter" that dominated the late 1960s and the 1970s, characterized by a decrease of optimism and funding, and caused by unfulfilled expectations. Early interest in associative processing was diminished by an influential book *Perceptrons* (Minsky & Papert, 1969) around the same time. This rigorous book showed that the state-of-the-art associative systems of the time could not implement any task that was not linearly separable, including the simple logical operator "exclusive or."

The AI winter of the 1970s, however, also witnessed the emergence of new theories and paradigms. For example, ANALOGY (Evans, 1968) solved simple geometric analogy problems that appear on some tests of human intelligence. SHRDLU (Winograd, 1972) performed natural language processing to understand commands to a robot to pick up and manipulate blocks. Marr (1982) developed a three-stage computational theory of vision: from a raw image to a primal sketch with edges, from primal sketches to 2 and 1/2 D representations including surfaces, to 3D object recognition. Schank first developed a theory of conceptual structures for natural language understanding (Schank, 1975) and then a theory of memory, reasoning, and learning (Schank, 1982).

Working in a different paradigm, Feigenbaum, Buchanan, and their colleagues first developed an expert system called Dendral that could generate hypotheses about molecular structures from spectroscopic data (Lindsay et al., 1980), and then an expert system called Mycin that could generate hypotheses about E. coli bacterial diseases from heterogeneous patient data (Buchanan & Shortliffe, 1984). AI's revival in the 1980s was due in part to the success of these *expert systems*, which were designed to replicate the expertise of individuals with a great deal of domain knowledge. Knowledge engineers would interview and observe experts, and then attempt to encode their knowledge into some form that an AI program could use. This was done with a variety of methods, including *decision trees* (which can be thought of as using the answers to a series of questions to classify some input, as in the game Twenty Questions). Since expert systems were of use to business, there was a renewed interest in AI and its applications. Funding for AI research increased.

One of the ideological debates of the 1980s was between the "neats" and the "scruffies": the neats used a formal, often logic-based approach and the scruffies focused on modeling human intelligence and getting AIs to use semantic information processing. Geographically, many of the neats were based at Stanford University and the US West Coast, and in Japan, and many of the scruffies were at Massachusetts Institute of Technology (MIT) and the US East Coast. Neats thought that knowledge representation and processing should be mathematically rigorous and elegant, and evaluations should involve proofs.

Scruffies believed that intelligence is so complex that it is unwise to put such constraints on it at this early stage of development of AI theory and methodology. Today, most of the engineering AI research would be classified as neat. A good deal of, but not all, contemporary psychological AI is scruffy.

In the 1980s, interest in artificial neural networks and associative AI was revived through cognitive modeling by *connectionists* (Rumelhart et al., 1986; McClelland et al., 1986). Connectionism continues to have a strong influence in modern cognitive science; in engineering AI, artificial neural networks are regarded as just one of many statistical learning mechanisms (such as Markov models and other methods of memory, reasoning and learning mentioned in the previous sections). Interestingly, some of the approaches and ideas of the cyberneticists have had a revival in these sub-symbolic approaches to AI.

Over time, the limits of expert systems became clear. As they grew in size, they became difficult to maintain and could not learn. As a knowledge base grows, inconsistencies between different chunks of knowledge tend to arise. In part again because of unfulfilled expectations, in the 1990s, AI entered a second "winter," with diminished optimism, interest, and funding. However, during the second winter, again, new frameworks appeared, including *embodied cognition, situated cognition*, and *distributed cognition*. These frameworks emphasize how the body and environment both constrain and afford cognition, how cognition always is in the context of the physical and social worlds where these worlds themselves afford information to the cognitive agent. Similarly, *agent-based AI* on one hand seeks to unify cognition with perception and action, and on the other, studies AI agents as members of a team of other agents (artificial or human).

Over the past decade or so, AI has witnessed a resurgence of interest and attention. Perhaps the most exciting new development in AI is "deep learning," which involves machine learning over multiple layers of artificial neural network units (LeCun, Bengio, & Hinton, 2015). This algorithmic innovation was facilitated by hardware advances that allowed AI researchers to program systems to work on special chips made for graphics: graphics processing units, or GPUs, which speeded deep learning by about 100 times.

Deep learning feels like a revolution, in part because it has been used to address many difficult problems in AI. To take a famous example, the Chinese board game Go was very difficult for AI, in part because there are so many possible moves at each turn. Using deep learning over information taken from large quantities of human Go games and knowledge, an AI system called AlphaGo (Silver et al., 2016) beat a professional Go champion in 2015. Two years later, the successor to AlphaGo, called AlphaGo Zero (Silver et al., 2017), was able to beat the original, but without having looked at any human knowledge at all: AlphaGo Zero got to be a world champion just by playing games against itself for forty days, and came up with previously unknown strategies that Go experts described as very creative.

More traditional realms of creativity, the arts, have also been the focus of AI research. We now have AIs that create paintings, jokes, musical compositions and improvisation, and poetry (Besold, Schlorlemmer, & Smaill, 2015; Veale & Cardoso, 2018). More complex artistic endeavors, such as creating a written novel, have proven more difficult, because a novel requires so much knowledge and understanding about how the world works. Many of these creative AIs are used for commercial products, usually under the name "procedural generation." Most famously, the computer game *No Man's Sky* (www.nomanssky.com) allows players to explore a virtual galaxy. When they land on a planet, the game creates the terrain and weather, as well as a complete ecosystem with custom-generated flora and fauna. Not only are the causal complexities of the simulation generated automatically, but the game also creates graphical models of them, complete with sounds. Procedurally generated content is economically important, because artistic creativity is a large part of the budget of many modern video games and movies, and the content is consumed far faster than it can be generated by human beings.

Robotics has proven to be a particular challenge for artificial intelligence – our running example of Sunny notwithstanding – because dealing with the real world is far more complex than dealing with formal, internal systems that play Go or recommend books to people. But these problems are slowly being addressed as well. Many people have robot vacuum cleaners in their houses now, and more complex robots are on their way. The field of *human-robot interaction* has arisen to study how humans do and can best interact with robots.

At present, AI appears to have entered a new phase of revival. This is in part due to the new frameworks that have appeared over the past generation, especially agent-based AI, deep learning, human-centered AI, and computational creativity. By now, AI is ubiquitous in industrialized societies, though it often does not go by that name. Many researchers avoid the term, feeling that it has been tarnished by the boom-and-bust cycle of interest and funding it has experienced in its sixty-year history. However, techniques from AI are used in many practical applications, allowing your voice to be understood when you talk to an automated phone system, using your past purchases to make recommendations for books when you shop online, efficiently matching flights to gates at airports, directing the pathfinding of characters in computer games, generating web search engine results, enabling face detection in cameras and online photo archives, and doing automatic translation.

The concerns of *human-centered AI* (Ford et al., 2015) are how individuals and societies can productively work with artificial intelligence to make progress on human values. For instance, if Sunny were given a command to get some food, implicit in our request is that Sunny does not steal food. But without world knowledge, the AI might complete the request, but not abide by assumed preferences about how it's done. Just as AIs need to understand what we mean, they also need to be able to convey what *they* mean.

Although machine learning, and particularly "deep learning," has enjoyed many breakthroughs in the past few years, a persistent problem for all sub-symbolic AI systems is that once they learn to do something, it is not immediately clear how or why they can do it. This is because each unit is by itself meaningless, and the processing involves interaction between thousands or more of them. It's too complex to understand by looking at the code.

However, for many applications it is very important that our AI systems are able to *explain* their decisions to us, so we can, for example, tell when an AI is discriminating against a group of people for a reason we feel violates their rights. To prevent these practical problems, as well as to better understand how intelligence works, a subfield has emerged to try to understand the workings of the sub-symbolic AI systems we build!

Of course, we have not tried to cover every topic in AI in this chapter. For example, over the past two decades, there has been much AI research on designing the *semantic web* (Berners-Lee, Hendler, & Lassila, 2001), a new version of the World Wide Web that would be capable of understanding information (e.g., web pages) stored on it. As another example, just over the past few years, *interactive games* have emerged as an important arena for AI research, especially agent-based AI.

Assessing Progress in AI

The task of measuring progress in AI is complex. In the past, tasks such as arithmetic and chess were considered to require intelligence. However, computers have been performing arithmetic calculations with great precision for more than seventy-five years and reliably beating human grand masters at chess for more than twenty-five (Hsu, Campbell, & Hoane, 1995). Although in computer science these problems are sometimes still used to measure progress in speed of calculation and use of memory, very few humans now consider these computer programs as good manifestations of intelligence. One part of the difficulty is that some humans seem reluctant to ascribe intelligence to computers: Once a computer is able to address the problems that we once considered to require intelligence, we tend to dismiss them as not being very interesting. Further, once we understand how a computer actually solves these problems, for many humans, these problems lose some of their challenge.

Early in the history of AI, Turing (1950) proposed the most famous test for AI, called the *imitation game* or, as it is more popularly known, the “Turing test.” In this test, computers and human beings are put in (typed) chat sessions with human judges. If computers can reliably make the judges think they are human, they pass the test. Turing initially formulated this test in response to the question “Can machines think?” But rather than answering that question, he reformulated it into a more concrete question of whether a machine could fool a human interrogator into believing that the computer can think. Some interpretations of the Turing test take the purpose of the test as distinguishing computer programs that have human-level intelligence from those that do not (e.g., Harnad, 1992). In this interpretation, the test is not a measurement of intelligence in the sense of giving a score that accurately reflects cognitive abilities, but is a pass-or-fail litmus test of general intelligence.

It has proven to be very difficult for computers to pass the Turing test in general, although some surprisingly simple and old programs, such as ELIZA (Weizenbaum, 1966) and PARRY (Raphael, 1976), sometimes fool some people for short times. Because of the difficulty of the general Turing test, many competitions usually restrict judges to specific topics. Recently, there have been variations of the Turing test with prize monies such as the Loebner prize.

Recently there have been proposals (Bringsjord & Schimanski, 2003) for using psychometrics tests of human intelligence, such as the Wechsler test (1939) and the Raven's test (1962) to measure progress in AI. However, there already exist computer programs that approach human performance on various versions of the Raven's test including the Standard, Color, and Advanced Raven's test (e.g., Kunda, McGreggor, & Goel., 2013). Other recent proposals (Marcus, Rossi, & Veloso, 2016) for measuring AI have covered a wide range, from playing soccer and winning the FIFA world championship to scientific discovery and winning the Nobel Prize.

AI and Society

As AI becomes more ubiquitous in our society, and affects more aspects of human life in greater ways, the question of ethical behavior becomes increasingly more important. Earlier we wrote about safety issues, using the example of our robot, Sunny. Many people worry that robots, as well as software AI agents, will continue to replace human jobs faster than society can create new ones. This isn't a safety issue that can be fixed in the code of a single agent, but rather a societal issue that needs to be dealt with on the level of laws and social norms.

Recently, several famous people such as Stephen Hawking have expressed fear that superintelligent AI might pose a threat to humanity's very existence. The reasoning goes something like this: At some point, an AI might be smarter than human beings, and have the power to rewrite its own code (this kind of AI is called a "seed AI"). Because the AI is smarter than any human, it will make itself smarter faster than humans can make it smarter. This will cause a "takeoff" that might be very rapid. Are there limits to how smart the software could get? We have no idea. Once the AI is many times smarter than any human, it will have enormous power to gain real control over world resources, using social manipulation, hacking, and other methods. No matter what the ultimate goals of the AI are, it would probably in the AI's interest to have the subgoals of self-preservation, cognitive enhancement, technological progress, resource acquisition, and prevention of its goals from being changed (the so-called instrumental convergence thesis, Boström, 2014).The need for good ethical reasoning at the start, in this scenario, is crucial because if the AI were to come to rule the world in pursuit of its goals, it might be difficult or impossible to change the ethics of the AI after it is many times smarter than us (and has an interest in preserving its values).

Of course this argument is very speculative and not based on any evidence. There are counterarguments suggesting that the above scenario is based on poor or very unlikely assumptions (e.g., Pinker, 2018). Further, there is also scholarship suggesting that superintelligent AI will be a force for good (Kurzweil, 2005). In any case, almost everyone agrees that we are nowhere near having a superintelligent AI. The study of ethical AI behavior is a growing field of interest. A related issue is whether or not we will someday need to have ethical considerations for the AIs themselves, should they ever be able to suffer pain.

Conclusions

In this chapter we have reviewed the history of AI and its major subfields, illustrated AI as a science and as a technology, examined its relationship to psychology, and discussed the problem of measuring the intelligence of AI agents. A somewhat surprising lesson from the history of AI is that it is relatively easy to make AI systems for some cognitive tasks that seem difficult for humans to solve (for example, mathematical, logical, and chess problems), and extraordinarily difficult to make computers solve some tasks that are apparently easy for humans to address (for example, seeing, walking, and talking). This apparent paradox has meant that repeated predictions about bold AI successes have gone unfulfilled.

We suggest two reasons for this paradox. First, our difficult problems require deliberate thought and strategies that are explicitly learned. As a result, we can often gain insight into how they are solved through observation and introspection. Indeed, many of these strategies are actually written down, meant to be learned through reading. In contrast, nobody needs to tell human beings how to see, walk, or speak. As a result, our intuitions about how these processes work are, to put it mildly, unhelpful.

The second, perhaps more important, reason is that deliberate processing is likely a serial process running as a virtual machine on a network of neurons, whereas the automatic processes, the seemingly easy tasks, are running directly on the neural network. These easy tasks (called System 1 in Stanovich & West, 2000) are evolutionarily older, and the parts of our brains that accomplish them (generally near the back of our brains) evolved to do just those things. In contrast, the more deliberate processing (System 2) is evolutionarily younger and makes use of the kind of hardware designed for System 1 tasks. System 2 struggles to do rational, serial processing on an essentially parallel pattern-matching machine (Kahneman, 2011; Stanovich, 2004).

Computers, and the languages we program them with, are naturally serial processors. When we implement artificial neural networks, we are doing it backward from nature: Whereas System 2 is a serial virtual machine running on parallel hardware, our artificial neural networks are parallel virtual machines running on serial hardware. Given this, and the fact that we have no conscious access to System 1 processes, it is no wonder that the AI community has had to work very hard to make progress in these areas. As a result, we have chess programs that can beat world grand masters, but no robots that can walk down a street even as well as a five-year-old child. We expect that neuroscience findings may illuminate the nature of these processes, and the AI community will be able to build on them.

Given the track record of predictions about the future of AI, we will refrain from making our own. What we will claim is that AI already has had a profound impact not only on computer science and information technology but also more generally on our culture and our philosophy. The field has made so much progress that the Association for Advancement of Artificial Intelligence (AAAI; www.aaai.org) organizes multiple conferences every year, including one for deployed AI applications. If the past fifty-year history of AI is any guide, then the next fifty years will not only be full of exciting discoveries and bold inventions, but they will also raise new questions about who we are as humans and what we want to be.

References

Albus, J. S. (1991). Outline for a theory of intelligence. *IEEE Transactions on Systems, Man, and Cybernetics*, *21*(3), 473–509.

Ali, K., & Goel, A. (1996). Combining navigational planning and reactive control. *Proceedings of the AAAI-96 Workshop on Reasoning About Actions, Planning and Control: Bridging the Gap* (pp. 1–7). Portland: AAAI Press.

Anderson, J. R. (2013). *The adaptive character of thought*. New York: Psychology Press.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

Arkin, R. (1999). *Behavior-based robotics*. Cambridge, MA: MIT Press.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). Semantic web. *Scientific American*, *284*(5), pp. 35–43.

Besold, T., Schlorlemmer, M., & Smaill, A. (Eds.) (2015) *Computational creativity research: Towards creative machines*. New York: Atlantis Press.

Boström, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford: Oxford University Press.

Bringsjord, S., & Schimanski, B. (2003). What is artificial intelligence? Psychometric AI as an answer. *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI-03)* (pp. 887–893). San Francisco: Morgan Kaufmann.

Buchanan, B., & Shortliffe, E. (1984). *Rule based expert systems: The Mycin experiments of the Stanford Heuristic Programming Project*. Boston: Addison-Wesley.

Cox, M., & Raja, A. (2011), *Metareasoning: Thinking about thinking*, Cambridge, MA: MIT Press.

Craik, K. (1943). *The nature of explanation*. Cambridge, UK: Cambridge University Press.

Davies, J., & Francis, A. G. (2013). The role of artificial intelligence research methods in cognitive science. In R. West & T. Stewart (Eds.), *Proceedings of the 12th International Conference on Cognitive Modeling* (pp. 439–444). Ottawa: Carleton University.

Evans, T. G. (1968). A program for the solution of a class of geometric-analogy intelligence-test questions. In M. Minsky (Ed.), *Semantic information processing* (pp. 271–353). Cambridge, MA: MIT Press.

Ford, K., Hayes, P., Glymour, C., & Allen, J. (2015). Cognitive orthoses: Toward human-centered AI. *AI Magazine*, *36*(4), 5–8.

Glasgow, J., Narayanan, N. H., & Chandrasekaran, B. (Eds.) (1995). *Diagrammatic reasoning: Cognitive and computational perspectives*. Cambridge, MA: MIT Press.

Goel, A., Ali, K., Donnellan, M., Gomez, A., & Callantine, T. (1994). Multistrategy adaptive navigational path planning. *IEEE Expert*, *9*(6), 57–65.

Goel, A., Stroulia, E., Chen, Z., & Rowland, P. (1997). Model-based reconfiguration of schema-based reactive control architectures. In *Proceedings of the AAAI Fall Symposium on Model-Directed Autonomous Systems* (pp. 1–6). Cambridge, MA: AAAI.

Harnad, S. (1992). The Turing test is not a trick: Turing indistinguishability is a scientific criterion. *SIGART Bulletin*, *3*(4), 9–10.

Hsu, F., Campbell, M., & Hoane, A. (1995). Deep Blue system overview. In M. Wolfe (Ed.), *Procs. the 1995 International Conference on Supercomputing* (pp. 240–244). New York: ACM Press.

Kahneman, D. (2011). *Thinking, fast and slow.* New York: Farrar, Straus and Giroux.

Kolodner, J. (1993). *Case-based reasoning*. San Francisco: Morgan Kaufmann.
Kotseruba, I., Gonzalez, O., & Tsotsos, J. (2016). A review of 40 years of cognitive architecture research: Focus on perception, attention, learning and applications. The Computing Research Repository (CoRR). arXiv preprint arXiv:1610.08602, 1–74.
Kunda, M., McGreggor, K., & Goel, A. (2013). A computational model for solving problems from the Raven's Progressive Matrices intelligence test using iconic visual representations. *Cognitive Systems Research*, *22*, 47–66.
Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. New York: Viking Adult.
Laird, J. E. (2012). *The Soar cognitive architecture*. Cambridge, MA: MIT press.
Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *AI Magazine*, *38*(4). https://doi.org/10.1609/aimag.v38i4.2744
Laird, J., Newell, A., & Rosenbloom, P. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, *33*, 1–64.
Langley, P. (2012). The cognitive systems paradigm. *Advances in Cognitive Systems*, *1*, 3–13.
Langley, P., Laird, J., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, *10*(2), 141–160.
LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.
Lenat, D., & Guha, R. (1990). Building large knowledge based systems: Representation and inference in the Cyc project. Boston: Addison-Wesley Longman.
Lindsay, R., Buchanan, B., Feigenbaum, E., & Lederberg, J. (1980). *Applications of artificial intelligence for chemical inference: The Dendral project*. New York: McGraw-Hill.
Marcus, G., Rossi, F., & Veloso, M. (2016). Beyond the Turing test. Special issue, *AI Magazine*, *37*(1), 3–101.
Marr, D. (1982). *Vision*. New York: Henry Holt.
McCarthy, J. (1988). Mathematical logic in AI. *Daedalus*, *117*(1), 297–311.
McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1955/2006). A proposal for the Dartmouth summer research project on artificial intelligence. *AI Magazine*, *27*(4), 12–14.
McClelland, J. L., Rumelhart, D. E., & PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 2, *Psychological and biological models*. Cambridge, MA: MIT Press.
Minsky, M. L. (1975). *A framework for representing knowledge*. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 1–82). New York: McGraw-Hill.
Minsky, M. L., & Papert, S. A. (1969). *Perceptrons*. Cambridge, MA: MIT Press.
Mitchell, M. (1998). *An introduction to genetic algorithms*. Cambridge, MA: MIT Press.
Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of problem solving. *Psychological Review*, *63*(3), 151–166.
Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Francisco: Morgan Kauffman.
Pearl, J. (2000). *Causality: Models, reasoning and inference*. New York: Cambridge University Press.
Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Pinker, S. (2018). *Enlightenment now: The case for reason. Science, humanism, and progress*. New York: Viking.

Quillian, M. (1968). Semantic Memory. In M. Minsky (Ed.), *Semantic information processing* (pp. 227–270). Cambridge, MA: MIT Press.

Rabiner, L., & Juang, B. H. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, January, 4–16.

Raphael, B. (1976). *The thinking computer*. New York: W. H. Freeman.

Raven, J. C. (1962). *Advanced Progressive Matrices Set II*. London: H. K. Lewis.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 1, *Foundations*. Cambridge, MA: MIT Press.

Samsonovich, A. V. (2010). Toward a unified catalog of implemented cognitive architectures. In A. V. Samsonovich, K. R. Jóhannsdóttir, A. Chella, & B. Goertzel (Eds.), *Proceeding of the Conference on Biologically Inspired Cognitive Architectures* (pp. 195–244). New York: IOS Press.

Schank, R. C. (1975). *Conceptual information processing*. New York: Elsevier.

Schank, R. C. (1982). *Dynamic memory* (2nd ed.). New York: Cambridge University Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.

Silver, D., Huang, A., Maddison, C., Guez, A., Sifre, L., Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529* (7587), 484–489.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., et al. (2017). *Mastering the game of Go without human knowledge*. *Nature*, *550*(7676), 354–359.

Simon, H. A. (1996). *Sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.

Singh, P., Lin, T., Mueller, E. T., Lim, G., Perkins, T., & Zhu, W. L. (2002). Open mind common sense: Knowledge acquisition from the general public. In R. Meersman & Z. Tari (Eds.), *On the Move to Meaningful Internet Systems: OTM Confederated International Conferences* (pp. 1223–1237). Berlin: Springer.

Sowa, J. (1987). Semantic networks. In S. Shapiro (Ed.), *Encylopedia of AI* (pp. 1011–1024). New York: Wiley.

Stanovich, K. E. (2004). *The robot's rebellion*. Chicago: University of Chicago Press.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, *23*, 645–726.

Stroulia, E., & Goel, A. K. (1999). Evaluating problem-solving methods in evolutionary design: The autognostic experiments. *International Journal of Human-Computer Studies*, *51*, 825–847.

Sutton, R. S., & Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*, 433–460.

Veale, T., & Cardoso, A. (2018). *Computational creativity: The philosophy and engineering of autonomously creative systems*. Berlin: Springer.

Von Anh, L., Liu, R., & Blum, M. (2006). Peekaboom: A game for locating objects in images. In R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, & G. Olson (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Montreal, April 22–27)* (pp. 55–64). New York: ACM Press.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Weiner, N. (1961). *Cybernetics* (2nd ed.). Cambridge, MA: MIT Press.

Weizenbaum, J. (1966). ELIZA – a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45.

Winograd, T. (1972). *Understanding natural language*. San Diego, CA: Academic Press.

26 Intelligence and Video Games

Maria Ángeles Quiroga and Roberto Colom

Introduction: Intelligence and Video Games

> A great deal has been learned using the testing techniques that are ubiquitous today. Forgetting or denigrating this information would be silly; science progresses by building on the past. But it is time to move on to new techniques of measurement if we want to obtain any major breakthrough.
>
> (Hunt, 2011, p. 864)

Defining the Playground

A video game involves a user interface generating feedback from a device such as a TV screen, a computer monitor, a tablet, or a smartphone. There are hundreds of video games and some of them require reasoning, planning, solving problems, and learning. These features are included in the definition of intelligence: "a very general mental capability that involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience" (Gottfredson, 1997a).

Charles Spearman (1904) postulated the principle of the indifference of the indicator based on the positive manifold (the substantive correlation among cognitive tasks irrespective of their content). This manifold is one the most replicated findings in psychology (Kovacs & Conway, 2016) and the implication is this: The vehicles (or superficial characteristics) of the situations science uses for assessing intelligence and cognitive ability are relatively irrelevant. The key lies in their cognitive requirements (Hunt, 2011; Jensen, 1998). From this perspective, it becomes possible to use video games for obtaining measures of the construct of interest.

Nevertheless, defining the playground is relevant. "Intellectual ability" refers to a consistent and stable disposition to solve families of cognitive problems (abstract, verbal, numerical, visuospatial, mechanical, and so forth). According to the Cattell-Horn-Carroll model (CHC; McGrew, 2009), several second-stratum or broad "cognitive" abilities can be identified, but "cognition" and "intelligence" are not synonymous. Intelligence (from the Latin term *intelegere*) refers to the ability to choose the best solution to solve a problem. Cognition (from the Latin *cognoscere*) refers to the faculty to process information from perception and acquired knowledge. Intelligence includes psychological processes, but the reverse is not true.

Second-stratum or broad abilities capture shared variance among diverse tests measuring a common ability to some extent. The psychometric properties of these tests are carefully addressed, but this is usually not the case for the experimental tasks tapping cognitive processes such as attention or working memory updating.

Finally, “skill” and “ability” must be clearly distinguished. Skill refers to the easiness of doing a given activity well because of a greater experience or training. To cook or to drive a car are skills. Higher-ability levels may facilitate the acquisition of skills, but once acquired and automated, ability differences may become less important.

First Studies Using Video Games

Video games research has considered “performance” and “experience.” The first refers to the level obtained in the game, whereas the second refers to people’s playing habits such as hours per week or genre. On the other hand, the “correlational” and “comparative” approaches have been applied. The first focuses on analyzing the covariance between standard ability tests and video game performance. The second considers ability differences associated with levels of experience or amount of training on the video games of interest.

The first studies using video games and administering intelligence tests were run thirty years ago (Jones, Dunlap, & Bilodeau, 1986; Rabbitt, Banerji, and Szymanski, 1989). Significant correlations between the variables assessed by games and tests were found. Jones and colleagues (1986) administered thirteen intelligence tests (from the Kit of Factor-Referenced Cognitive Tests by R. B. Ekstrom, French, and Harman, 1976) and five video games for the Atari console (*Air Combat Maneuvering, Breakout, Race Car, Slalom*, and *Antiaircraft*). Correlation values ranged from 0.18 (for the *Slalom* game) to 0.50 (for the *Race Car* game).

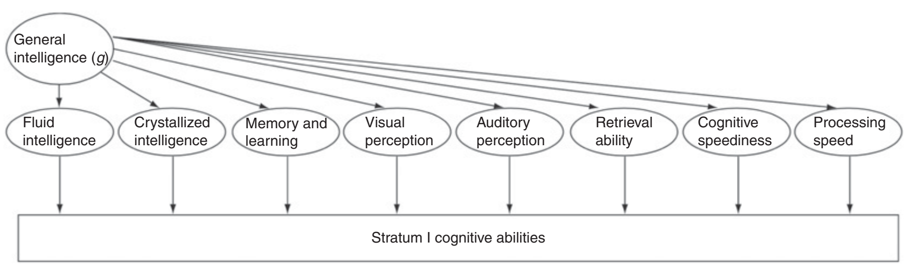

Figure 26.1 *Simplified depiction of the Cattell-Horn-Carroll (CHC) model.* There are more than sixty stratum I cognitive abilities summarized in a much smaller set of stratum II abilities. General intelligence (*g*) is at the apex (stratum III). The location of the stratum II abilities represents their higher or lower relationship with the higher-order factor.

Three years later, Rabbitt and colleagues (1989) assessed intelligence with the AH4 test (Heim, 1968), whereas video game performance was evaluated using *Space Fortress*. The AH4 test is a group-administered test consisting of sixty-five items belonging to the verbal and numerical domains. *Space Fortress* is a video game designed at the University of Illinois for studying complex-skill acquisition. The game's goal is to shoot missiles and destroy a space fortress (it can be installed and run from http://hyunkyulee.github.io/research_sf.html), and participants played over five successive days. Greater correlations were observed with increased practice (from 0.28 to 0.68). This was the main conclusion: "a relatively unsophisticated video-game, on which performance may reasonably be expected to be independent of native language or acquired literacy, and which is greatly enjoyed by young people who play it, rank orders individual differences in 'intelligence' nearly as well as pencil and paper psychometric tests which have been specially developed for this purpose over the last 80 years" (p. 13).

The increase in correlations from the first to the fifth session suggests that some practice was necessary to overcome preexisting differences in familiarity with the video game. The correlation became stable once those differences disappeared (stay tuned).

The Video Games Jungle

Video games comprise a variety of genres. Their cognitive requirements are different regarding planning, speed, psychomotor ability, and so on. Beyond their superficial similarities, small differences among games may recruit different cognitive processes (Sedig, Haworth, & Corridore, 2015). Thus, for instance, the term "action video games" is highly unspecific: "it encompasses several video game genres, without controlling for effects potentially stemming from differences in mechanics between these video games" (Dobrowolsky et al., 2015, p. 59).

Video games can be categorized into genres taking into account (1) their gameplay mechanics, (2) the in-game tasks, (3) the rules that players must follow, (4) whether they are multiplayer or not, and (5) the devices required to play the game (Torre-Tresols, 2017). Sajjadi, Vlieghe, and De Troyer (2017), to help researchers, describe the connection between gameplay mechanics and a variety of abilities.

Martinovic and colleagues (2014) elaborated a matrix detailing the cognitive processes probably recruited by several video games with the main aim of categorizing them. Based on their proposal, Table 26.1 provides a list of video games classified by genre, subgenre, and main features.

The differentiation among video games has been increasingly refined. Indeed, distinguishing genres and genre subtypes is tough. These fine distinctions were absent in the first studies relating cognitive abilities and video game performance. For example, Green and Bavelier (2003) considered some action video games titles to belong to the Action genre, but others to belong to fighting, shooter, and car races (see Table 26.1).

Table 26.1 *Types of video games (after Torre-Tresols, 2017)*

Genre	Subgenre	Main features	Examples
1. Action		Speed; high level of psychomotor abilities.	*Grand Theft Auto*
2. Adventure		Quick-time events; exploring big surfaces collecting objects to solve problems.	
	2.1. Graphic adventure	Point and click; 2D or 3D; player interacts with the mouse or different control devices to complete tasks.	*Monkey Island* *Hotel Dusk Room 215* *Zero Scape*
	2.2. Visual novel	Player has no control over the character.	*Steins; Gate*
	2.3. Mixed	These games have features both from graphic adventures and visual novel.	*Heavy Rain* *The Wolf among Us*
3. Action-adventure		Action games which place an importance upon narrative; combine elements from different gameplay styles with the same focus (when the game focuses mainly on a specific play style it will be grouped in that category instead of here); subgenres differ in the focus they give to certain elements (planning versus shooting, for example).	
	3.1. Stealth action	Require mainly planning because players cannot defend themselves from the enemy through simple force or firepower.	*Metal Gear* *Deus Ex* *Sprinter Cell*
	3.2. Survival horror	Main goal is to survive in a frightening atmosphere; player can, in some titles, defend themselves from enemies.	*Resident Evil* *Clock Tower* *Survival Run*
	3.3. Platforms	Player has to overcome obstacles.	*Super Mario* *Sonic the Hedgehog* *Megaman* *Rayman*
	3.4. Metroidvania	Combines exploring plus platform features.	*Ori and the Blind Forest* *Hollow Knight*
4. Sandbox		Player moves freely in an open world where they can choose what to do; player can change the game world.	*Minecraft* *Garry's Mod*

Table 26.1 (*cont.*)

Genre	Subgenre	Main features	Examples
5. Fighting		Physical fighting in small and closed areas. These games require high mechanical skill to use the control buttons.	
	5.1. Fighting game classic	Combat between two or more fighters, which can be grouped, of comparable strength; often divided into rounds.	*Street Fighter* *Super Smash Bros* *Tekken* *Guilty Gear* *Blazblue*
	5.2. Beat 'em up	Focused on cooperative game and player versus environment; usually 2D; easy combats that include narrative.	*Double Dragon* *Golden Axe*
	5.3. Hack 'n' slash	A variation of Beat 'em ups; hand-to-hand combat; medium mechanical skill.	*Devil May Cry* *Bayonetta* *Metal Gear Rising: Revengeance*
6. Shooter		Mainly focused on moving and shooting; 2D or 3D; first (FPS) or third person (TPS) shooting.	
	6.1. Shoot 'em up	Similar to beat 'em up but shooting; action could be vertical or horizontal.	*Aero Fighters* *Space Invaders* *Galaga* *Satazius*
	6.2. Danmaku (bullet hell)	These games show a curtain of fire; require high attentional level, adaptation to novelty; and high control of fine motor movements; it is more important to dodge bullets than to attack enemies.	*Ikaruga* *Touhou*
	6.3. Classic shooter	Shoot 'n' run; high speed of movements required; very quick action; unlimited arsenal of weapons; one player or multiplayer; lack of life regeneration.	*Doom* (FPS) *Wolfenstein* (FPS) *Unreal Tournament* (Multiplayer FPS) *Quacke* (Multiplayer FPS) *Half Life* (Multiplayer)
	6.4. Tactical shooter	Limited amount of weapons; the chosen weapon determines the character's speed; multiplayer.	*Counter-Strike*

Table 26.1 (*cont.*)

Genre	Subgenre	Main features	Examples
	6.5. Modern shooter	Multiplayer; reduced teams, small and closed maps; small recoil-weapons; life regeneration.	*Titanfall* (FPS) *Battlefield* (FPS) *Gear of War* (TPS) *Splatoon* (TPS)
7. Role-playing game (RPG)		Main feature is to let the character evolve; high load on narrative; real-time action or in turns; player knows the quantitative value of the attributes they obtain (strength and speed usually); different roles: tanks, healers; and DPS (damage per second).	
	7.1. Western RPG	More focused on expression and fantasy; player develops their avatar in a story; combat in turns.	*Pillars of Eternity* *Divinity*
	7.2. Japanese RPG	No avatar; highly narrative loaded, mainly focused in interpersonal relationships; combats in turns.	*Final Fantasy* *Bravely Default*
	7.3. Action RPG	Real-time combats.	*The Legend of Zelda* *Bayonetta* *Dark Souls*
	7.4. Dungeon crawler	Player explores gigantic dungeons; 3D; first-person perspective.	*Etrian Odissey* *Might and Magic*
	7.5. Tactical RPG	Include elements from strategy genre; combat, rather than exploration, is central .	*Final Fantasy Tactics*
8. Strategy		Player has to focus on tactics and long-term plans to complete the mission.	
	8.1. Strategy in turns	Player controls units, collects, and manages resources; one or multiplayer; map for the mission is divided into cells; only one action per turn; more focused on strategy than in combat.	*Civilization*
	8.2. Real-time strategy (RTS)	Continuous action without pauses; player controls several units they can send to combat, build, or collect; one or multiplayer.	*Starcraft (*one of the more famous esports) *Hears of Iron* *Age of Empires*

Table 26.1 (*cont.*)

Genre	Subgenre	Main features	Examples
	8.3. Tower defense	Building defenses to avoid enemies trespassing; decision-making about strategic locations for your defenses; some titles from this genre include shooter playability, allowing player to defend against enemies.	*Flash Element Tower Defense* *Desktop Tower Defense* *Orcs Must Die!* *Sanctum*
	8.4. MOBA (multiplayer online battle arena)	Action real-time strategy (ARTS) or Dota-like (Dota = Defense of the Ancients); player controls a unit that moves along a symmetric map to destroy the enemy base; multiplayer only.	*League of Legends* *Dota 2* (both titles are famous eSports games)
	8.5 Tactic in turns	Similar to strategy genre in turns and tactic RPG, but combat focused.	*X-Com*
	8.6. Real-time tactics	Combat strategy, emulating tactics from the battlefield.	*Total War* *Full Spectrum Warrior*
9. Puzzle		Focused on solving problems unrelated to each other; barely including narrative (exception is the Professor Layton saga); solutions to problems do not rely either on speed or accuracy but on intellectual abilities.	
	9.1. Educational games	Goal is to teach though problems or questions; usually for children; when games include playful elements the genre is named "edutainment games."	*Brain Training* *Big Brain Academy*
	9.2. Action puzzles	Played in real time; player has to perform very coordinated actions to solve the problems; usually problems are visuospatial.	*Portal* *Portal 2*
10. Car races		Driving vehicles; titles can be more or less realistic.	*Gran Turismo* *Forza* *Mario Kart* *Crash Team Racing*
11. Music		These games are focused on dancing, singing, or following rhythms.	
	11.1. Rhythm	Player must type commands following the music; two types: played by hand and played by feet.	*Beatmania* *Dance, Dance Revolution* *Pump It Up*

Table 26.1 (*cont.*)

Genre	Subgenre	Main features	Examples
	11.2. Dancing	Player has to imitate the movements they see on the screen while hearing a song.	*Just Dance*
	11.3. Singing	Player has to sing a song whose lyrics are on the screen (like a karaoke).	*Sing Star*
12. Sports		These games reproduce the practice of a sport.	*Fifa* *Wii Sports*
13. Nonmechanical genres		Titles in this genre can also belong to one of the other genres, for reasons other than game mechanics.	
	13.1. Arcade	Essentially this refers to the distribution format; in earlier times, each of these games was played on a specific machine in a public environment (e.g. pub); never rely on narrative.	
	13.2. Simulators	"Simulation games" may or may not refer to the game mechanics, because they can be referred to as their own genre or be a simulation game of another genre (e.g., racing simulation).	*The Sims* *Animal Crossing* *Farmville* *Forza* (driving) *StarCitizen* (piloting an spatial ship)
	13.3. Massive multiplayer online (MMO)	All players play in a shared world interacting among themselves. Usually these games are also RPG (MMORPG).	*World of Warcraft* *Dungeon Fighter Online* (playability like beat 'em up)
	13.4. Roguelike	These games consist of dungeons that are built at random each time the player plays; only one player; no life regeneration; exploration is essential to solve the game; the goal is not to finish the game once but many times (to unlock special features).	*The Binding of Isaac* *Enter the Gungeon* (action) *Nuclear Throne* (action) *Faster Than Light* (tactics)
		Many of these games are also action-adventure genre but there are exceptions.	*Strafe* (FPS) *Forgotten Depths*

Genre is important when comparing video game players with non-players. In this regard, Dobrowolski and colleagues (2015) compared people who had played mainly first-person shooter (FPS) or real-time strategy (RTS) video games for seven or more hours per week during the six months previous to starting the study with people who had played five hours or less per week (including no more than two hours per week of FPS and/or RTS). Players and non-players were compared on task-switching performance and multiple-object tracking. RTS players outperformed non-players on the set size they could accurately follow in the MOT and they were also less affected by switches than non-players in the switching task. There were no differences found between the FPS players and FPS non-players.

The enhancement of attention skills sometimes found (Green & Bavelier, 2003) might result from the different processes in players whose experience comes from different game genres. These processes may remain unknown if the analyzed group of participants includes more individuals who have played shooter and strategy games (e.g., *Team Fortress Classic*) than people who have played car racing games (e.g., *Super Mario Kart*), even though both are "action games."

Experience with video games is usually assessed in terms of hours per week and video game genre. There are some questionnaires for assessing these variables. The Video Games Playing Habits (VGPH) questionnaire by Quiroga and colleagues (2011) and the Video Game Playing Questionnaire by the Bavelier Lab (Bediou et al., 2018) are two examples.

The first studies considered non-players only (Adams & Mayer, 2012; Glass Maddox, & Love, 2013; Quiroga et al., 2009, 2011) because it was relatively easy to find naïve participants. When video game experience started to be explicitly considered, the cut-off was five or more hours per week (Green & Bavelier, 2003; Green, Pouget, & Bavelier, 2010). Later this cut-off rose to 6–7 hours per week (West et al., 2017). The numbers are expected to increase steadily.

Intelligence Assessment Using Video Games

The Association between Intelligence and Video Game Performance: Cautionary Tales

There are research findings showing a lack of association between playing commercial video games and individual differences in cognitive ability (Gnambs & Appel, 2017; Unsworth et al., 2015). However, other studies report very different conclusions (Foroughi et al., 2016; Kokkinakis et al., 2017; Quiroga et al., 2015). To know why there is this discrepancy, some crucial points must be clarified.

First, video game performance cannot (and should not) be estimated using time invested playing video games. The amount of time devoted to doing something is not a guarantee of achieving greater performance (Macnamara, Hambrick, & Oswald, 2014). As noted by Green and colleagues (2017), the relationship between practice and outcome is not linear and, therefore, it is strongly inappropriate to use playing time as a proxy of playing performance.

Unfortunately, this is usually overlooked. Thus, for instance, Sala, Tatlidil, and Gobet (2018) examined the meta-analytic correlation between video game performance and cognitive ability (or cognitive processes). Their main conclusion was this: There is no relation between the two domains. However, studies measuring video game performance ($N = 28$) and those measuring video game playing hours ($N = 38$) were combined, leading to a strange mix of effects (performance, motivation, etc.). As detailed below, correlations between intelligence and video games are indeed substantial when studies measuring just playing hours are excluded.

Second, not all tasks are proper measures of cognitive ability. Visual attention tasks, for instance, do not measure any second-stratum or broad ability. At best, they can be considered within the first or narrow stratum below general visualization (G*v*) (Figure 26.1). At worst, some visual-attention tasks measure very specific cognitive processes weakly related with the cognitive ability of interest.

Cognitive "abilities" and cognitive "processes" belong to conceptual realms that must be distinguished. Again, in the meta-analytic study by Sala and colleagues (2018), from the twenty-eight papers using raw scores as video game performance, twelve referred to action video games. Among those twelve papers, seven have been published in peer-reviewed journals. The correlations between cognitive ability tests and video game performance were: Progressive Matrices = 0.63; Symmetry Span = 0.30; Mental Rotation = 0.69; Mental Paper Folding = 0.40. However, the correlations between cognitive tasks[1] and video game performance were: Antisaccade task = 0.15; Change Detection Task = −0.11; Color Wheel Task = −0.31; Matching Figure Task (RT) = 0.12; Matching Figure Task (Accuracy) = 0.01; Visual Search Task = 0.11. The difference between cognitive ability tests (average correlation 0.69) and cognitive tasks (average correlation 0.14) is pretty obvious.

Furthermore, research is moving fast beyond computing simple correlations between one test or task and performance on a given video game. The interest focuses now on the latent traits tapped by various specific measures (Baniqued et al., 2013; McPherson & Burns, 2007, 2008; Quiroga et al., 2009, 2011). Results derived from this much more appropriate approach – based on the estimation of second-stratum abilities (usually G*f*, G*c*, G*v*, G*y*, G*s*) and the computation of structural equation models (SEM) correlating latent factors for cognitive ability and for video game performance (Baniqued et al., 2013; Foroughi et al., 2016; Quiroga et al., 2015, 2019) – are summarized in Table 26.2.

Third, the characteristics of the sample must be explicitly considered. Studying children, adolescents, or adults may have differential impact on the observed findings.

It is very important to keep in mind that meta-analytic studies are not the "cure-all" remedy for psychological science. The combination of weak studies, even using sophisticated statistical tools, cannot replace carefully designed and developed studies. Mega-samples of individuals combined from largely disparate designs

[1] We use the term "cognitive tasks" instead of "cognitive tests" for those measures that were designed as laboratory tasks and lack precise psychometric properties.

may appeal to the naïve reader, but must be deeply inspected by the specialist before buying the message.

Meta-analytic reports can be very damaging for emerging research fields. The meta-analysis of Sala and colleagues (2018) discussed above is a paradigmatic example. Only four of the eighteen studies specifically focused on cognitive abilities measuring video game "performance" were considered. As underscored by H. J. Eysenck (1993):

> Including all relevant material – good, bad, and indifferent – in meta-analysis admits the subjective judgments that meta-analysis was designed to avoid. Several problems arise in meta-analysis: regressions are often non-linear; effects are often multivariate rather than univariate; coverage can be restricted; bad studies may be included; the data summarized may not be homogeneous; grouping different causal factors may lead to meaningless estimates of effects; and the theory-directed approach may obscure discrepancies. (p. 789)

In short: revise and think carefully about the information included in published meta-analyses because there may be much more than meets the eye (and for the worse).

Intelligence and Video Game Performance in Adults

Table 26.2 summarizes the results reported in research studies relating cognitive ability and video game performance published since 2007.

There is variability among studies, but commonalities can be highlighted. All correlations are positive, which is consistent with the principle of the indifference of the indicator, even when different genres are considered: puzzles, sports, shooters, real-time strategy, MOBAs (multiplayer online battle arena), or customized ones.

The correlation with intelligence tests can be underestimated studying MOBAs. Furthermore, video game matchmaking ranking (MMR: ratio of historical wins to losses) is sometimes used as the performance measure, although MMR from different leagues does not imply the same performance level. This practice may also underestimate the correlation with intelligence.

The reliability values for the video games considered are in the medium to high range (Table 26.2): (1) for puzzles values are high (internal consistency from 0.75 to 0.94[2]; stability from 0.65 to 0.84; split-half = 0.92); (2) for third-person shooters they are also high (0.92) and (3) for sports games they are medium to high (0.77 to 0.86).

Cognitive abilities assessed with video games differ across studies, including fluid intelligence (G*f*), broad visual perception (G*v*), general memory and learning (G*y*), and processing speed (G*s*), but the correlation values are similar when using composite scores: from 0.69 to 0.74 for G*f*, or from 0.41 to 0.67 for G*s*, for instance (see Table 26.2 for further details). The lowest correlations are for attention tasks. Results from the two SEM models tested were 0.93 for brain games and 0.78 for non-brain games. Therefore, the general cognitive ability factor (*g*) and the general video game

[2] Except for two *Big Brain Academy* games; *Faces* = 0.44, and Color Count = 0.57.

Table 26.2 *Summary of studies relating intelligence and video games*

Study	Video game description	Ability measured[1]	Results (reliability and convergent validity, at the test and at the latent levels)
McPherson and Burns (2007)	*Space Code Game* (customized action game). Computer mouse response method. Goal is to destroy enemies' spaceships that appear in the window view of a cockpit. At the bottom of the cockpit, nine spaceships are presented, each one with a single digit placed directly above. Destroying a spaceship requires firing the number placed above the matching ship at the bottom of the screen. Video game experience was not assessed. *N* = 61, 37 females, mean age 20.	G*s* (processing speed)	Test-retest = 0.84 Gs = 0.67 (4 tests) Gv = 0.35 (3 tests) APM = 0.49 (Gf) VAL WJ-III = 0.24 (Glr)
McPherson and Burns (2008)	*Space Matrix* (customized action game). In this game, participants were asked to destroy spaceships in the same manner as in *Space Code* while also monitoring where dots were located on a 5 x 5 grid as the one developed for the Dot Matrix test (Miyake et al., 2000). These dot locations were described as indicators of which "sector" of space they were operating in. From time to time participants had to report back to head-quarters which sectors they had been operating in. The screen layout was the same as that for *Space Code*, with the addition of the sector grid, which appeared at inter-vals to the right of the numerical response grid on the cockpit control panel.	G*y* (general memory and learning)	Test-retest = 0.77 Correlations at the test level: Dot Matrix = 0.66 APM = 0.51 Picture Swaps = 0.54 Correlations at the construct level: Gf/WM = 0.69 (3 tests) Gs = 0.40 (3 tests)

Table 26.2 (*cont.*)

Study	Video game description	Ability measured[1]	Results (reliability and convergent validity, at the test and at the latent levels)
	Video game experience was assessed related to hours per week played and experience playing with a mouse. N = 70, 40 females, mean age 19.6.		
Quiroga et al. (2009)	Three games from *Big Brain Academy* for the Nintendo Wii Console: *Mallet Math, Reverse Retention*, and *Train Turn* (puzzle games). Participants played 10 blocks, each consisting of 10 items, during 2 nonconsecutive weeks (15 days of separation). Video game experience was assessed. Selected participants had no previous experience with these video games. N = 27, 17 females, mean age 21.5.	*g* (general mental ability), obtained from 5 tests (Numerical Reasoning, FDSPAN, FLSPAN, Rotation of Solid Figures, and Corsi Block).	Mallet Math = 0.12 to −0.52 Rev. Retention = 0.43 to 0.34 Train Turn = 0.49 to 0.67
Quiroga et al. (2011)	Two games from *Big Brain Academy* for the Nintendo Wii Console: *Train Turn* and *Speed Sorting* (puzzle games). Participants played 25 blocks, each consisting of 10 items, during 5 consecutive weeks. Video game experience was assessed, with the Video Games Playing Habits. Selected participants had no previous experience with these video games. N = 27 females, mean age 21.	*g* (general mental ability), obtained from 5 tests (PMA-R, PMA-S, D-48, and Rotation of Solid Figures).	Train Turn = 0.65 to 0.67 Speed Sorting = 0.65 to 0.34
Adams and Mayer (2012)	*Tetris* (puzzle game). *Unreal Tournament* (UT; classic shooter game).	G*v* (broad visual perception) static (Paper Folding and Mental	G*v* static: Pap. Fold. – Tetris = 0.24

	All participants were non-video game players. *N* = 69, 44 females, mean age 19.3.	Rotation) and G*v* dynamic (Race2 and Interception Tasks by Hunt et al., 1988).	Pap. Fold. – UT = 0.27 M. Rotat (errors) – Tetris = −0.21 M. Rotat. (errors) – UT = −0.27 G*v* Dynamic: Race2 RT – Tetris = −0.20 Intercep. Hits – Tetris = 0.19 Race2 RT – UT = −0.27 Intercep. Hits – UT = 0.21
Baniqued et al. (2013)	20 casual games (for computer; can be considered to be puzzle games), grouped in 4 types: Reasoning, Working Memory, Spatial Reasoning, Attention, Visuo-Motor Speed and Perceptual Speed. *N* = 219, 33% male, mean age 21.7.	G*f* (fluid intelligence) G*y* (general memory and learning) G*s* (processing speed) G*lr* (long-term retrieval) Attention	At the test level (Zmean): G*f* G*y* G*s* G*lr* Att. WM+R 0.65 0.55 0.36 0.12 0.06 Spat. Rel. 0.57 0.44 0.18 0.00 0.13 Attention 0.46 0.41 0.28 −0.03 0.19 Vis.Motor 0.27 0.17 0.23 0.03 0.08 Per.Speed 0.36 0.24 0.24 0.02 0.06 At the latent level:
Ventura et al. (2013)	*Virtual Spatial Navigation Assessment – VSNA* (adventure game). Participants have to find a set of gems in a 3D environment using a first-person avatar. Participants answered a general question about how often they play video games. *N* = 323, 194 females, no mean age reported.	G*v* (broad visual perception)	Test-retest = 0.65 Spatial Orientation Test = 0.18 Mental Rotation Test = 0.26
Shute, Ventura, and Ke (2015)	*Portal 2* (action puzzle game) *Lumosity Platform* (includes 52 puzzle games). 77 participants, 43% male, mean age 19.7.	G*f* (fluid intelligence) G*v* (broad visual perception) Creativity Problem-solving	*Portal 2*: SPM-reduced = 0.02 Insight Test = −0.38 Remote Association Test (creativity) = −0.18 Mental Rotation = −0.33 Spatial Orientation = 0.27

Table 26.2 (*cont.*)

Study	Video game description	Ability measured[1]	Results (reliability and convergent validity, at the test and at the latent levels)
			VSNA = 0.34 Lumosity: SPM-reduced = 0.37 Insight Test = 0.26 Remote Association Test (creativity) = 0.10 Mental Rotation = −0.05 Spatial Orientation = −0.11 VSNA = −0.10
Quiroga et al. (2015)	Ten *Big Brain Academy* games and *Garden Gridlock* (for the Nintendo Wii console) plus *Tilt Maze* (for computer), which were grouped in 4 types following game developers' descriptions: Analyze, Memorize, Compute, and Visualize. All are puzzle games. Video game experience was assessed. Selected participants were naïve for the Wii console and *Big Brain Academy* video game. *N* = 188, 67 men, mean age 22.2.	G*g* (general mental ability) G*f* (fluid intelligence) G*c* (crystallized intelligence) G*v* (broad visual perception) G*y* (general memory and learning) G*s* (processing speed)	Reliability (internal consistency): Analyze games: 0.76 to 0.80 Memorize: 0.44 to 0.67 Compute: 0.57 to 0.80 Visualize: 0.71 to 0.95 At the test level: G*f* G*c* G*v* G*y* G*s* Analyze 0.62 0.40 0.65 0.30 0.55 Memorize 0.46 0.44 0.32 0.37 0.44 Compute 0.54 0.47 0.48 0.44 0.46 Visualize 0.64 0.34 0.66 0.29 0.41 At the latent level: *g* – video game latent = 0.93
Buford and O'Leary (2015)	Modified version of *Portal 2* (action puzzle game).	G*f* (fluid intelligence) G*c* (crystallized intelligence)	Split-half reliability of 0.92 SPM = 0.46 (IPS)

	Previous game experience as well as experience and skill with *Portal 2* was assessed. Two samples of 94 (online sample – OS; mostly men, mean age 24.8, very high experience playing. Only 27 completed the cognitive measures and 73 (in person – IPS; 58% women, mean age 19.6, almost no experience playing) participants.		Shipley Block Patterns = 0.49 (IPS) Shipley Vocabulary = 0.30 (IPS) Wonderlic = 0.27 (OS)
Foroughi et al. (2016)	A version of *Portal 2* developed by authors (action puzzle game). Two samples: • *N* = 35 video game players (VGPs) experienced playing *Portal 2*, 9 females, mean age 21.3 • * *N* = 100 video game players and non-video game players (NVGPs), 74 females, mean age 21.6.	G*f* (fluid intelligence)	Reliability (α) = 0.80 At the test level: APM – VGPs = 0.65 APM – both = 0.61 BOMAT – both = 0.63 APM – NVGPs = 0.60 BOMAT – NVGPs = 0.67 At the latent level: Gf = 0.78
Bonny, Castaneda, and Swanson (2016)	*Dota 2* (action real time strategy – ARTS; MOBA). Participants were recruited in the International Tournament 5 of *Dota 2*. *Dota* play history was assessed. *N* = 396, 34 females, mean age 23.4.	G*y* (general memory and learning) Numerical processing ability	WM task accuracy = 0.03 WM task RT = −0.04 Location Mem. Task-d' = −0.06 Location Mem. Task – RT = −0.10 Number Task Accuracy = 0.24 Number Task – RT = −0.11
Quiroga et al. (2016)	• *Professor Layton and the Curious Village* (puzzle game). • Two samples recruited in 2009 and 2016: • *N* = 47, 9 males, mean age 19.6 • *N* = 27, 6 males, mean age 20.6. Participants completed 15 hours playing in six weeks.	*g* (general mental ability) obtained from 3 tests: AR, SR, and VR from the Differential Aptitudes Test).	Reliability (internal consistency): 0.94 First sample: Puzzles found = 0.10 to 0.65 Puzzles solved = 0.20 to 0.58 Second sample: Puzzles found = 0.40 to 0.53 Puzzles solved = 0.30 to 0.58

Table 26.2 (*cont.*)

Study	Video game description	Ability measured[1]	Results (reliability and convergent validity, at the test and at the latent levels)
Kranz et al. (2017)	Six casual games (3 adaptive and 3 non-adaptive. Action puzzle games and puzzle and skill games). Ten 20-minute playing sessions (2 to 3 per week) *N* = 94, 30 males, mean age 21.2.	G*f* (fluid intelligence) G*y* (general memory and learning) G*s* (processing speed)	Reason. – adaptive games = 0.60 to 0.74 WM – adaptive games = 0.40 to 0.65 Per. Speed – adap. games = 0.00 to 0.05 Reason. Non-adap. games = 0.60 to 0.48 WM – Non-adap. games = 0.50 to 0.38 Percep. Speed – Non-adapt. games = 0.37 to 0.38
Kokkinakis et al. (2017)	*League of Legends* (*LoL*), *Dota 2* (action real time strategy, multiplayer online battle arena). All subjects were experienced *LoL* players who had played a large number (> 100) of both "ranked" and "unranked" matches. *N* = 56, 51 males, mean age 20.5 years.	G*f* (fluid intelligence) G*y* (general memory and learning)	Matrix-WASI II = 0.44 Rotation Span = 0.26 Symmetry Span = 0.12 Operation Span = 0.03
Kirkegaard (2018)	*Dota 2, League of Legends, Starcraft II* (action real time strategy –ARTS, MOBA); *Counter Strike: Global Offensive; Overwatch* (first-person shooter) *Counter Strike* (tactical shooter) *Hearthstone* (card game) *Super Smash Bros* (classic fighting game). Data collected at the country level (*N* = 195 countries).	*g* (general mental ability)	National IQ and general gaming ability = 0.79

Lim and Furnham (2018)	*Taboo* (board game), *Portal* (action puzzle). *N* = 112, 101 males, mean age 18.6 years.	G*f* (fluid intelligence) Gc (crystallized intelligence)	Portal (time taken) – APM = −0.61 Taboo (describing) – APM = 0.33 Taboo (guessing) – APM = 0.28
Quiroga et al. (2019)	*Space Invaders* (shoot 'em up), *Splatoon* (third-person shooter), *Art of Balance, EDGE, Hook, Rail Maze, Blek* (puzzle games), *Unpossible* (action game), *Sky Jump, Crazy Pool* (sports game). *N* = 134, 29 males, mean age 21.04.	*g* (general mental ability)*G* G*f* fluid intelligence) G*v* (broad visual perception) G*s* (processing speed)	Reliability (internal consistency): *Splatoon* = 0.92; *Blek* = 0.91; *Rail Maze* = 0.81; *Sky Jump* = 0.77; *Crazy Pool* = 0.86. At the latent level: *g* – video game latent = 0.79

[1] Second-stratum abilities will be referred to using the CHC theory of cognitive abilities.

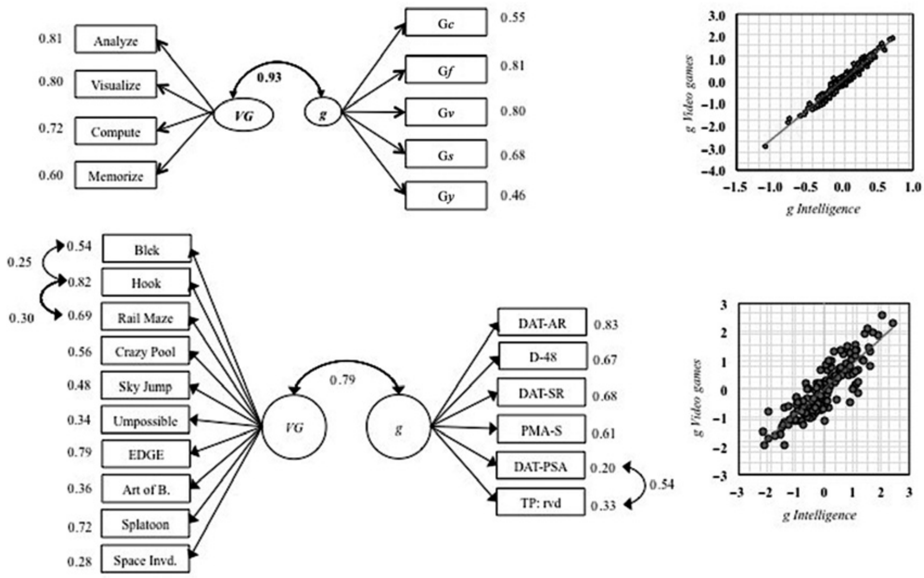

Figure 26.2 *Correlations between the latent factors representing general video game performance (VG) and the general factor of intelligence (g) from SEM model with brain games (upper panels) and from SEM model with non-brain games (bottom panels) (after Quiroga et al., 2015, 2019).*

performance factor are closely similar. This opens the door to the design of intelligence assessment batteries using video games (see Figure 26.2).

Regarding G*f*, when video games are very novel, raw correlations with cognitive ability are low at the beginning and increase until reaching the 0.65/0.74 range (Kranz et al., 2017; Quiroga et al., 2016; Rabbitt et al., 1989). This increased correlation demonstrates that video game performance is far from automated across practice (Ackerman, 1988; Quiroga et al., 2011).

Studies include players and non-players. In this regard, the study by Foroughi and colleagues(2016) shows that previous experience with the game hardly changes the correlation between fluid intelligence and video game performance. New items were designed using the mod that *Portal 2* includes (Buford & O'Leary, 2015; Foroughi et al., 2016). *Portal 2* consists of chambers containing puzzles to be solved. The mod allows researchers to build their own chambers (each chamber is usually like an item in a test) and so remove the effect of previous experience to solve the new game. Note that these results support measurement invariance for video games related to G*f*, or in other words, the video game is measuring the same construct irrespective of the experience players have.

More recent studies have introduced the assessment of playing habits and self-perceived skill when playing for identifying profiles of video game players across different genres. The first studies simply selected participants without any experience, but this is unfeasible nowadays.

There is a lack of studies regarding predictive validity. The few studies considering this crucial issue focused on the association between academic success and playing habits assessed by hours per week devoted to playing (Drummond & Sauer, 2014; Posso, 2016). Higher scores in the 2012 Program for International Student Assessment (PISA) were observed in students playing more hours per week. Specifically, students who played online games almost every day scored 15 points above the average in math and reading, and 17 points above the average in science. This advantage was absent in those using social networks. In fact, students using online social media on a daily basis scored 4 percent lower than the average on math, reading, and science.

There are no studies relating video game performance and job performance, but Chiang (2010) enumerated ten ways video games might boost occupational achievement using *World of Warcraft* (a role-playing [RPG] and massive multiplayer online [MMO] game; see Table 26.1 for details). Chiang enumerated several facets (leadership, dealing and learning from failure, teamwork, developing talent, flexibility [learning to improvise], being performance driven, living for challenge, competitiveness, entrepreneurship, and managing information) but this still requires formal research.

In conclusion, video game performance correlates with cognitive abilities. However, more systematic research is required using a clear theoretical framework regarding the cognitive abilities considered along with the superficial features and mental requirements of the analyzed video games.

The Measurement of Cognitive Processes Associated with Intelligence Using Video Games

Cognitive processes involve (1) the acquisition and understanding of knowledge, (2) decision-making, and (3) problem-solving. There are two processes extensively analyzed with respect to video game performance: perception and attention (Bediou et al., 2018).

In this regard, video game research has been focused on first- and third-person shooters, usually referred as action video games. The key features of these games are: fast pace, high cognitive load requiring updating, systematic switching between local and global fields of action, and selective attention to detect relevant items among distractors.

The fast pace of these games is ideal for youngsters. Findings usually show that video game experience is associated with more efficient cognitive processes: visuospatial cognition g[3] $= 0.75$; perception $g = 0.78$; top-down attention $g = 0.63$; multitasking/switching $g = 0.55$; inhibition $g = -0.31$ and verbal cognition $g = 0.30$. Obtained effect sizes for video game training are lower than those for video game experience (more than 30 hours of training are required for achieving noticeable

[3] This is Hedge's g, which is equivalent to Cohen's d but especially suited for small sample sizes in meta-analysis. It estimates effect size correcting for positive bias.

improvements). However, follow-up data are required for confirming these positive effects. Furthermore, it is quite possible that children showing higher cognitive ability levels from the outset are more prone to play. Unfortunately, these research studies do not measure ability baseline levels.

For real-time strategy (RTS) games, video game experience is related to the set size that can be followed accurately in a multiple object tracking (MOT) task. Also, RTS players are less affected by task switching than non-players (Dobrowolski et al., 2015).

In a study of video game training by Glass and colleagues (2013), involving forty hours training on *Starcraft*, a gaming condition that emphasized rapid switching between multiple sources of information and action (the player commands and controls two separate bases in multiple battles against two different opponent bases) led to a large increase (Stroop $d = 0.70$) in cognitive flexibility compared to playing *The Sims* (a life simulator game) for the same amount of time. Interestingly, an even larger effect ($d = 1.44$) in cognitive flexibility has been obtained after training for only two hours with a customized game that requires switching between competing tasks (Parong et al., 2017).

Intelligence, Video Games, and the Brain

Playing video games is usually intensive and extensive. We have already highlighted the relevance of different playing habits with respect to the measurement invariance of video games. Now we discuss some neural correlates of video game playing. These neural correlates will be related with (1) how intensive and extensive the practice has been, (2) the video game genre, and (3) the players' cognitive profile. Table 26.3 summarizes the published studies.

Analyzing the same group of participants who had completed *Professor Layton and the Curious Village* (see Tables 26.1 and 26.2 for a comprehensive description), which took sixteen hours on average (four hours per week during four weeks), structural and functional brain changes were observed when compared with a control group. Regarding brain structural responsiveness to practice, Colom and colleagues (2012) analyzed cortical gray-matter volume, cortical surface area, cortical thickness, and white matter integrity. Gray-matter changes were mainly circumscribed to frontal regions, but there were also some findings in the temporal and parietal lobes. White matter integrity increased in the hippocampal cingulum and the inferior longitudinal fasciculus.

The study by Martínez and colleagues (2013) computed group-independent component analyses applying multi-session temporal concatenation on test-retest resting state fMRI along with a dual-regression approach. The key finding revealed increased correlated activity in parietal-frontal networks after playing the game (Figure 26.3) (the video animation showing the regions involved on the identified networks can be seen here: www.youtube.com/watch?v=jj3eaMm-Frc).

The functional changes occurred mainly in left temporal, parietal, and frontal networks involved in varied memory and executive functions presumably relevant

Table 26.3 *Main neural correlates of playing video games (after Palaus et al., 2017; Colom et al., 2012; Martínez et al., 2013)*

Variable	Video game	Neuroimaging method	Results
Intelligence and working memory			
Video game (VG) performance	*Rise of Nations* (real-time strategy)	Magnetic resonance imaging (MRI) with optimized voxel-based morphometry (VBM)	Volumetric changes in dorsolateral prefrontal cortex (dlPFC).
	Warship Commander Task (action)	NIRS (infrared spectroscopy)	Higher activation of prefrontal regions associated to game difficulty (dlPFC).
	Neuroracer (3D customized game)	-	Stimulating left dlPFC using tDCS obtained improvement in multitasking performance.
	Tank Attack 3D (action game); *Sushi Go Round* (strategy without action)	Diffusion tensor imaging (DTI) scans	White matter FA in the right fornix/stria correlated with action game learning whereas white matter FA in the left cingulum/hippocampus correlated with strategy game learning.
VG experience	*Starcraft* (real-time strategy)	Cortical thickness (FreeSurfer software)	Increased cortical thickness in parietal cortex correlated with winning rates of the league.
	League of Legends; Dota	MRI and fMRI (functional magnetic resonance imaging)	Consolidate connectivity between executive regions (dlFC and PPC) and the salience network (anterior insula and the ACC).
	Guilty Gear (third-person shooter)	VBM and statistical parametric mapping (SPM) analysis	Structural gray matter change in posterior parietal. VGPS higher right inferior parietal lobe. ROI analysis increased gray matter volume in left caudate nucleus.
VG training	*Brain Fitness* (puzzle); *Space Fortress* (shoot 'em up); *Rise of Nations* (real-time strategy)	Diffusion-derived white matter integrity; functional connectivity	Puzzle game: changes in integrity occipitotemporal white matter. Puzzle and action games: decrease functional connectivity between SPC and ITL compared to *Rise of Nations*.

Table 26.3 (*cont.*)

Variable	Video game	Neuroimaging method	Results
	Space Fortress (shoot 'em up)	EEG; ERSPS (event-related spectral perturbations)	Frontal alpha power and alpha and delta ERSPS predicted subsequent learning and performance.
	Space Fortress (shoot 'em up)	fMRI	Changes in functional activity in SPL.
	Professor Layton and the Pandora's Box (puzzle)	Connectivity-wise Resting state	Resting-state functional connectivity changes in frontal, parietal, and temporal areas.
	Professor Layton and the Pandora's Box (puzzle)	MRI-optimized VBM; cortical surface; cortical thickness; white matter integrity	Volumetric changes in frontal, parietal, and temporal lobes, bilateral. White matter: volumetric changes in hippocampal cingulum and inferior longitudinal fasciculus.
	Super-Mario 64 (action adventure)	MRI VBM8 toolbox	Gray-matter increases in right hippocampus. RdlPFC and bilateral cerebellum. Hippocampal increase related to changes from egocentric to allocentric navigation.
Visuospatial ability			
VG training	*Super Mario 64*	Cortical thickness (FreeSurfer)	Increased hippocampal volumes.
	Space Fortress (Shoot 'em up)	fMRI	Decreased activation in occipitoparietal regions linked to improved visuomotor task performance.
VG experience	Hours per week without specifying types of games	Cortical thickness (FreeSurfer)	Structural volume enlargements in the right hippocampus.
	Puzzle, action and role games	MRI VBM	Entorhinal cortex was positively correlated with lifetime experience in logic/puzzle VGs but negatively with action-based role-playing games.
	Expert gamers (more than 8 years playing more than 20 hours/ week last 6 months)	EEG	Earlier N100 latencies in visual pathways.

Table 26.3 (*cont.*)

Variable	Video game	Neuroimaging method	Results
Attention			
VG experience	*Halo; Counterstrike*; *Gears of War; Call of Duty* (first-person shooter)	MRI FMRIB Software Library	In non-gamers, a frontoparietal network of areas showed greater recruitment as attentional demands increased. Gamers barely engaged this network as attentional demands increased.
	Action games	Steady-state visual evoked potentials	P300 larger amplitude in VGPs than in NVGPs.
VG training	*Space Fortress* (Shoot 'em up)	fMRI (FSL 4.1 and FEAT)	After training, participants showed a reduction of activation of the right middle frontal gyrus, right superior frontal gyrus and ventral medial prefrontal cortex while control group continued to engage these areas.
VG performance	*Mario Power Tennis* (sports)	EEG (spectral analysis of theta and alpha waves)	Increment of the midline theta rhythm that increases with practice and decrease of the parietal alpha wave activity followed by a slow increase.

for the game (co-activated during video game playing). Playing the game may, therefore, feed the interaction between prefrontal and posterior memory-related regions for cognitive control of encoding and retrieval processes when the information stored in the short-term is monitored and manipulated within the working-memory system.

We discuss next three examples related with video game (1) training, (2) experience, and (3) performance.

Kühn and colleagues (2013) analyzed gray-matter volume changes after two months (thirty minutes per day) of practice with *Super Mario 64* (an action-adventure game) in young adults with little or no game experience in the past six months and who had not previously played *Super Mario 64*. The results obtained showed significant increase in gray-matter volume in the right hippocampus, right dorsolateral prefrontal cortex, and bilateral cerebellum. Regarding number of playing hours, Kühn and colleagues (2014) found a positive association between cortical thickness and two brain areas that belong to the frontoparietal network. They did not report the genres played and, therefore, their results can be interpreted as a brain mean effect of playing video games in general.

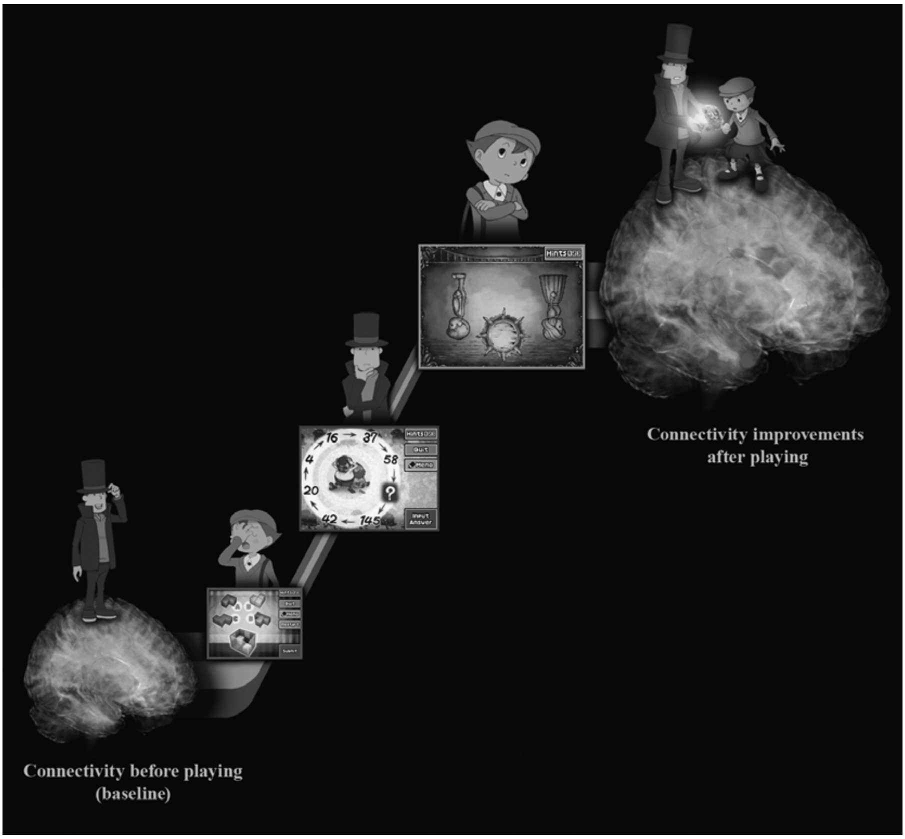

Figure 26.3 *Regions showing increased functional connectivity at rest, after playing* Professor Layton and the Curious Village *four hours per week over four weeks (Martínez et al., 2013).*

The comparison of individuals who play shooter video games (at least five hours per week playing video games like *Call of Duty, Halo, Counterstrike*, or *Gears of War* in the previous twelve months) and non-players (less than one hour per week playing the aforementioned video games in the previous twelve months, but playing other games such as puzzle, card, or strategy games) has revealed clear differences between those groups when completing selective attention tasks (Bavelier et al., 2012). Functional MRI showed higher frontoparietal activation in non-players with increased attention requirements, whereas this was not the case for experienced players. Therefore, experienced players seem more efficient in filtering irrelevant information.

In the third study, Nikolaidis and colleagues (2014) used *Space Fortress* to analyze whether changes observed in some brain areas while playing predict changes in nontrained working memory tasks. Participants were nonfrequent players (less than four hours per week). Results showed that activity changes in the superior parietal lobe, the paracingulate gyrus, and the precuneus predicted 37 percent of the individual differences observed in a nontrained working memory task, but not in a change-detection task.

The studies described support the association between practice with commercial video games (puzzles and shooters) and brain changes. The conclusion is reinforced by a recent meta-analysis (Palaus et al., 2017) but, again, it is crucial to have clear frameworks for orienting research efforts to avoid wasting time and resources (Colom & Román, 2018).

What's Next?

We have seen that people can be ranked according to their video game performance. This parallels ranking using standardized intelligence tests. The correlation between cognitive ability tests and video game performance is medium to high at the test level (Baniqued et al., 2013; McPherson & Burns, 2007, 2008; Quiroga et al., 2016), but the values are extremely high at the latent level (Quiroga et al., 2015, 2019) (see Figure 26.2). These results apply to quite heterogeneous genres, from puzzles to MOBAs.

Video games are also useful to test for intelligence in both players and non-players (they show measurement invariance; Foroughi et al., 2016). If (and only if) video games show medium cognitive complexity, are relatively consistent, and avoid transfer, extensive practice does not change their correlation with standard intelligence tests (Quiroga et al., 2009, 2011).

We are now ready to ask the next question: Is it time to use video games for measuring intelligence and related cognitive abilities?

Yes, it is. We strongly endorse the message contained in the quote that opens this chapter.

However, several issues must be addressed.

First, psychologists must be involved in the steps required for designing a video game: content, mechanics, complexity levels, variables to be saved, and scores to be computed. Commercial video game creators don't care about the information researchers and practitioners want.

Using commercial video games for research is inefficient because of time spent and research assistants needed, and this may explain why researchers do not test video game performance but rather video game experience (with a questionnaire). However, as already noted, these two measures tell different stories. Furthermore, when psychologists are there from the very beginning (i.e., at the development of the game, as in the case of McPherson & Burns' 2007, 2008 research), correlations with paper-and-pencil tests increase because the video game is oriented toward tapping the cognitive ability of interest.

Second, video games can be designed as adaptive tests by broadening their scope. Video games can easily include individualized pathways with different endings depending on the difficulty levels achieved. If a player cannot overcome a certain difficulty level, an exit pathway can be provided to avoid negative feelings.

Intelligence models may guide these pathways. This double adaptive approach will allow implementing, in (say) a video game designed to measure fluid reasoning, the rules and components for inductive reasoning considered by Primi (2014): (1) quantitative pairwise progression; (2) figure addition and subtraction; (3) distribution of three values in which the elements are instances of a conceptual attribute; (4) attribute addition; (5) distribution of two values. This may allow the design of criterion-referenced assessment tools avoiding arbitrary metrics based on normative scores.

Third, video games would be a useful way for estimating the average IQ level of populations (Kirkegaard, 2018). They might contribute to assessing inaccessible groups (video games can be implemented in cell phones) and also to systematically analyze the link between intelligence and health in real time, as suggested by Kokkinakis and colleagues (2017).

Fourth, video games may allow testing for response processes. They can record the continuous "flow" of behaviors. This would increase ecological validity. In everyday life settings, the same result can be achieved by using different pathways.

Furthermore, emotions can be manipulated to test their influence (or lack of) over cognitive performance. *Forgotten Depths* (downloadable for free from www.quirogas.net) is a customized game designed for achieving this goal. The game taps working memory with or without an environment that evokes fear. The software provides accuracy and time data scores for the primary (processing) and secondary (storing) tasks (to find the exit to each labyrinth and to collect all the required gems, respectively).

Available results show a correlation of 0.70 between standard working memory tasks and video game performance within neutral labyrinths (without emotion – no spiders). However, the correlation is decreased within the emotional labyrinths ($r = 0.50$).

Forgotten Depths also provides data about (1) time invested in "safe" or "risky" places within each labyrinth, (2) number of clicks to exit, (3) number of times spiders killed the player, (4) number of times the player used their weapon, (5) number of spiders killed, and so on. Using these variables, results have shown that fearful people, even if they have the same working memory ability level on the standard tasks, perform worse on the video game that contains spiders than non-fearful people ($d = 0.48$) because they invest more time in finding the exit of the labyrinth ($d = 0.44$), although they collect the same amount of gems (achievement measure $d = 0.05$). Also, they stay longer in risky areas than non-fearful players ($d = -0.54$) and spiders bite them more frequently than non-fearful players ($d = -0.59$). Fearful people seem to experience greater levels of fear while solving the game, obtain worse working memory scores, are easily disoriented (more time on risky areas and more clicks to exit) and show lower reaction behaviors (more spiders bite them).

This customized video game includes a mod for researchers to elaborate their own labyrinths as needed. *Forgotten Depths* is a good example of the type of video game required for measuring cognitive abilities properly.

In closing, systematic research is needed. The available evidence is highly promising, but funds are greatly required. Commercial video games must be substituted with games designed by scientists from the very beginning if they are to be used in both research and practice.

References

Ackerman, P. J. (1988). Individual differences and skill acquisition. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and practice* (pp. 165–217). New York: W. H. Freeman and Company.

Adams, D., & Mayer, R. (2012). Examining the connection between dynamic and static spatial skills and video game performance. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 34. https://escholarship.org/uc/item/8vc391r3

Baniqued, P. L., Lee, H., Voss, M. W., Basak, C., Cosman, J. D., DeSouza, S., et al. (2013). Selling points: What cognitive abilities are tapped by casual video games? *Acta Psychologica*, *142*, 74–86. http://dx.doi.org/10.1016/j.actpsy.2012.11.009

Bavelier, D., Achtman, R. L., Mani, M., & Föcker, J. (2012). Neural bases of selective attention in action video game players. *Vision Research*, *61*, 132–143. https://doi.org/10.1016/j.visres.2011.08.007

Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green, C. S., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological Bulletin*, *144*(1), 77–110. http://dx.doi.org/10.1037/bul0000130

Bonny, J. W., Castaneda, L. M., & Swanson, T. (2016). Using an international gaming tournament to study individual differences in MOBA expertise and cognitive skills. In *Proceedings of the SIGCHI conference on human factors in computing systems* (3473–3484). San José, CA. http://dx.doi.org/10.1145/2858036.2858190

Buford, C. C., & O'Leary, B. J. (2015). Assessment of fluid intelligence utilizing a computer simulated game. *International Journal of Gaming and Computer-Mediated Simulation*s, *7*, 1–17. http://dx.doi.org/10.4018/IJGCMS.2015100101

Chiang, O. (2010). Ten ways games can boost your careers. www.forbes.com/2010/07/19/career-leadership-strategy-technology-videogames_slide.html#2b2f83c05cb6

Colom, R., Quiroga, M. A., Solana, A. B., Burgaleta, M., Román, F. J., Privado, J., et al. (2012). Structural changes after videogame practice related to a brain network associated with intelligence. *Intelligence*, *40*, 479–489.

Colom, R., & Román, F. J. (2018). Enhancing intelligence. From the group to the individual. *Journal of Intelligence*, *6*(1), 11. https://doi.org/10.3390/jintelligence6010011

Dobrowolsky, P., Hanusz, K., Sobczyk, B., Skorko, M., & Wiatrow, A. (2015). Cognitive enhancement in video game players: The role of video game genre. *Computers in Human Behavior*, *44*, 59–63. http://dx.doi.org/10.1016/j.chb.2014.11.051

Drummond, A., & Sauer, J. D. (2014). Video-games do not negatively impact adolescent academic performance in science, mathematics or reading. *PLoS One*, *9*(4), e87943.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of Factor-Referenced Cognitive Tests*. Princeton: Educational Testing Service.

Eysenck, H. J. (1993). Meta-analysis and its problems. *British Medical Journal*, *309*, 789–792.

Foroughi, C. K., Serraino, C., Parasuraman, R., & Boehm-Davis, A. (2016). Can we create a measure of fluid intelligence using Puzzle Creator within Portal 2? *Intelligence*, *56*, 58–64. http://dx.doi.org/10.1016/j.intell.2016.02.011

Glass, B. D., Maddox, W. T., & Love, B. C. (2013). Real time strategy game training: Emergence of a cognitive flexibility trait. *PLoS One*, *8*(8), e70350. http://dx.doi.org/10.1371/journal.pone.0070350

Gnambs, T., & Appel, M. (2017). Is computer gaming associated with cognitive abilities? A population study among German adolescents. *Intelligence*, *61*, 19–28. http://dx.doi.org/10.1016/j.intell.2016.12.004

Gottfredson, L. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*(1), 13–23.

Gottfredson, L. (1997b). Why g matters: The complexity of everyday life. *Intelligence*, *24*, 79–132.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*, 534–537.

Green, C. S., Kattner, F., Eichenbaum, A., Bediou, B., Adams, D. M., et al. (2017). Playing some video games but no others is related to cognitive abilities: A critique of Unsworth et al. (2015). *Psychological Science*, *28*(5), 679–682. http://dx.doi.org/10.1177/0956679761664837

Green, C. S., Pouget, A., & Bavelier, D. (2010). Improved probabilistic inference as a general learning mechanism with action video games. *Current Biology*, *20*, 1573–1579. http://dx.doi.org/10.1016/j.cub.2010.07.040

Heim, A. W. (1968). *AH4 Test*. Windsor, UK: Nfer-Nelson.

Hunt, E. B. (2011). Where are we? Where are we going?: Reflections on the current and future state of research on intelligence. In R. J. Sternberg & S. B. Kauffman (Eds.), *Cambridge handbook of intelligence* (pp. 864–885). New York: Cambridge University Press.

Hunt, E., Pellegrino, J. W., Frick, R. W., Farr, S. A., & Alderton, D. (1988). The ability to reason about movement in the visual field. *Intelligence*, *12*, 77–100.

Jensen, A. (1998). *The* g *factor: The science of mental ability*. Westport, CT: Praeger.

Jones, M. B., Dunlap, W. P., & Bilodeau, I. M. (1986). Comparison of video game and conventional test performance. *Simulation and Game*s, *17*(4), 435–446.

Kirkegaard, E. O. W. (2018). Is national mental sport ability a sign of intelligence? An analysis of the top players of 12 mental sports. https://psyarxiv.com/9qnwy

Kokkinakis, A. V., Cowling, P. I., Drachen, A., & Wade, A. R. (2017). Exploring the relationship between video game expertise and fluid intelligence. *PLoS One*, *12* (11), e0186621. https://doi.org/10.1371/journal.pone.0186621

Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, *27*, 151–177. https://doi.org/10.1080/1047840X.2016.1153946

Kranz, M. B., Baniqued, P. L., Voss, M. W., Lee, H., & Kramer, A. F. (2017). Examining the roles of reasoning and working memory in predicting casual game performance across extended gameplay. *Frontiers in Psychology*, *8* (203), 1–13. http://dx.doi.org/10.3389/fpsyg.2017.00203

Kühn, S., Gleich, T., Lorenz, R. C., Lindenberger, U., & Gallinat, J. (2013). Playing Super Mario induces structural brain plasticity: Gray matter changes resulting from training with a comercial video game. *Molecular Psychiatry*. https://doi.org/10.1038/mp.2013.120

Kühn, S., Lorenz, R., Banaschewski, T., Barker, G. J., Büchel, C., Conrod, P. J., et al. (2014). Positive association of video game playing with left frontal cortical thickness in adolescents. *PLoS One*, *9*(3), e91506.

Lim, J., & Furnham, A. (2018). Can commercial games function as intelligence tests? A pilot study. *Computer Games Journal*, *7*(1), 27–37. https://doi.org/10.1007/s40869-018–0053-z

Macnamara, B., Hambrick, D., & Oswald, F. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, *25*(8), 1608–1618. http://dx.doi.org/10.1177/0956797614535810

Martínez, K., Solana, A. B., Burgaleta, M., Hernández-Tamames, J. A., Alvarez-Linera, J., Román, F. J., et al. (2013). Changes in resting-state functionally connected parieto-frontal networks after videogame practice. *Human Brain Mapping*, *34*, 3143–3157. http://dx.doi.org/10.1002/hbm.22129

Martinovic, D., Ezeife, C. I., Whent, R., Reed, J., Burgess, G. H., Pomerleau, C. M., et al. (2014). "Critic-proofing" of the cognitive aspects of simple games. *Computers and Education*, *72*, 132–144. http://dx.doi.org/10.1016/j.compedu.2013.10.017

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10. http://dx.doi.org/10.1016/j.intell.2008.08.004

McPherson, J., & Burns, N. R. (2007). Gs invaders: Assessing a computer game-like test of processing speed. *Behavior Research Methods*, *39*, 876–883. http://dx.doi.org/10.3758/BF03192982

McPherson, J., & Burns, N. R. (2008). Assessing the validity of computer-game-like tests of processing speed and working memory. *Behavior Research Methods*, *40*, 969–981. http://dx.doi.org/10.3758/BRM.40.4.969

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Palaus, M., Marron, E. M., Viejo-Sobera, R., & Redolar-Ripoll, D. (2017). Neural basis of video gaming: A systematic review. *Frontiers in Human Neuroscience*, *11*(248), 1–40. http://dx.doi.org/10.3389/fnhum.2017.00248

Parong, J., Mayer, R. E., Fiorella, L., MacNamara, A., Homer, B. D., & Plass, J. L. (2017). Learning executive function skills by playing focused video games. *Contemporary Educational Psychology*, *51*, 141–151. http://dx.doi.org/10.1016/j.cedpsych.2017.07.002

Posso, A. (2016). Internet usage and educational outcomes among 15-year-old Australian students. *International Journal of Communication*, *10*, 3851–3876.

Primi, R. (2014). Developing a fluid intelligence scale through a combination of Rasch modeling and cognitive psychology. *Psychological Assessment*, *26*(3), 774–788. http://dx.doi.org/10.1037/a0036712

Quiroga, M. A., Aranda, A., Román, F. J., Privado, J., & Colom, R. (2019). Intelligence can be measured with video games other than "brain-games." *Intelligence*, 75, 85–94.

Quiroga, M. A., Escorial, S., Román, F. J., Morillo, D., Jarabo, A., Privado, J. et al. (2015). Can we reliably measure the general factor of intelligence (*g*) through commercial video games? Yes, we can! *Intelligence*, *53*, 1–7. http://dx.doi.org/10.1016/j.intell.2015.08.004

Quiroga, M. A., Herranz, M., Gómez-Abad, M., Kebir, M., Ruiz, J., & Colom, R. (2009). Video-games: Do they require general intelligence? *Computers and Education*, *53*, 414–418. http://dx.doi.org/10.1016/j.compedu.2009.02.017

Quiroga, M. A., Román, F. J., Catalán, A., Rodríguez, H., Ruiz, J., Herranz, M., et al. (2011). Videogame performance (not always) requires intelligence. *International Journal of Online Pedagogy and Course Design*, *1*, 18–32. http://dx.doi.org/10.4018/ijopcd.2011070102

Quiroga, M. A., Román, F. J., De la Fuente, J., Privado, J., & Colom, R. (2016). The measurement of intelligence in the XXI century using video games. *Spanish Journal of Psychology*, *19*, 1–13.

Rabbitt, P., Banerji, N., & Szymanski, A. (1989). Space Fortress as an IQ test? Predictions of learning and of practiced performance in a complex interactive video game. *Acta Psychologica*, *71*, 243–257.

Sajjadi, P., Vlieghe, J., & De Troyer, O. (2017). Exploring the relation between the theory of multiple intelligences and games for the purpose of player-centered game design. *Electronic Journal of e-Learning*, *15*(4), 320–334. www.ejel.org/main.home

Sala, G., Tatlidil, K. S., & Gobet, F. (2018). Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin*, *144* (2), 111–139. http://dx.doi.org/10.1037/bul0000139

Sedig, K., Haworth, R., & Corridore, M. (2015). Investigating variations in gameplay: Cognitive implications. *International Journal of Computer Games Technology*, Article ID 208247. http://dx.doi.org/10.1155/2015/208247

Shute, V. J., Ventura, M., & Ke, F. (2015). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers and Education*, *80*, 58–67. http://dx.doi.org/10.1016/j.compedu.2014.08.013

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*(2), 201–292.

Torre-Tresols, J. J. (2017). Clasificación de géneros de videojuegos [Classification of video games genres]. Laboratory of Intelligence (Faculty of Psychology), Universidad Complutense de Madrid: Laboratorio de Inteligencia y videojuegos. www.quirogas.net

Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2015). Is playing video games related to cognitive abilities? *Psychological Science*, *26*, 759–774. http://dx.doi.org/10.1177/0956797615570367

Ventura, M., Shute, V. J., Wright, T., & Zhao, W. (2013). An investigation of the validity of the virtual spatial navigation assessment. *Frontiers in Psychology*, *4*, 852. http://dx.doi.org/10.3389/fpsyg.2013.00852

West, G. L., Konishi, K., Diarra, M., Benady-Chorney, J., Drisdelle, B. L., Dahmani, L., et al. (2017). Impact of video games on plasticity of the hippocampus. *Molecular Psychiatry*, *23*, 1566–1574. http://dx.doi.org/10.1038/mp.2017.155

PART VI

Kinds of Intelligence

27 The Theory of Multiple Intelligences

Mindy L. Kornhaber

Introduction

The theory of multiple intelligences (MI) was established in 1983 by Howard Gardner, then a research psychologist at the Boston Veterans Administration Medical Center and co-director of Project Zero, a research and development group at the Harvard Graduate School of Education (Gardner, 1983). Under MI theory, an intelligence is "a biopsychological potential to process information" (Gardner, 2009, p. 5). Further, all individuals have several, relatively autonomous intelligences, which they use in varying combinations to solve problems or create products that are valued in one or more cultures. Together, the intelligences underlie the range of adult endstates – occupations, social roles, and practitioners within domains of organized knowledge and skill – found across cultures (Gardner, 1983, 1993, 1999a, 2009). MI thus diverges from theories rooted at the start of the twentieth century that argue general intelligence, *g*, is central to all human problem-solving (Spearman, 1904). It also stands as a challenge to classical stage theory (Piaget, 1983) and, more generally, to measuring intelligence through psychometric methods.

In this chapter, I first consider the origins of MI. Following that, I present the evidence and criteria used to develop the theory, provide clarifications of the theory, examine critiques of the theory, consider its educational implementations, and briefly look to its future.

Origins of Multiple Intelligences (MI)

Since the early 1970s, Gardner had conducted numerous studies both of children's cognitive development and of adults' neuropsychological impairments (Gardner, 1983, 1999a). This research generated numerous contradictions to the conception of intelligence as a single, general intelligence. For example, Gardner and his colleagues found that children's developmental trajectories differed with different symbol systems – for example, for those entailing language, music, and gesture (e.g., Gardner & Wolf, 1983). However, if all problem-solving were governed by one underlying mental capacity, then the developmental trajectory across different symbols should proceed at the same pace. The unevenness of development across symbol systems also ran counter to the work of Jean Piaget (Gardner, 1983), who held that stages of children's cognitive development from sensorimotor to formal operations, in realms such as number, causality, and volume, occurred in a well-coordinated, if not entirely lockstep, manner (Piaget, 1983).

In neuropsychological investigations, Gardner and his colleagues documented patients who had lost communicative skills in language but who could learn to communicate through visual symbols (Gardner et al., 1976). He documented patients with language impairment, who varied with regard to their comprehension of humorous cartoons (Gardner et al., 1975), and the case of an artist whose brain damage spared his language abilities, while leaving him unable to recognize or identify objects presented visually (Wapner, Judd, & Gardner, 1978). Such disjunctions run counter to the notion of *g* and related psychometric conceptions under which all problem-solving abilities are positively correlated (e.g., Carroll, 1993; Spearman, 1904) and can be arrayed in a positive manifold (Spearman, 1904).

The contradictions to mainstream theory about intelligence were further catalyzed by Gardner's involvement in the Project on Human Potential at the Harvard Graduate School of Education. Gardner's charge by the project's funder, the Bernard Van Leer Foundation, was to synthesize what was then known about human cognition in the biological and behavioral sciences. Thus, Gardner undertook a detailed exploration of what was known about the development of cognition in normal and gifted children and the breakdown of cognitive capacities among adults (Gardner, 1983, 2011a). The product of his synthesis, MI theory, took issue "with the assumption inherent in *g* that an individual who has a high *g* could be equally accomplished in any intellectual area. MI theory is an extended argument against this all-purpose view of intellect" (Gardner, 2006a, p. 69). That said, MI is not intended to eradicate *g*-based views but, rather, to question their "province and explanatory power" (Gardner, 2006a, p. 69). In essence, Gardner argues that *g*'s role is likely more limited than its proponents have asserted and that other cognitive capacities contribute markedly to human problem-solving.

Evidence and Criteria for the Intelligences

Psychometric conceptions of intelligence are frequently used to explain and predict a range of measurable outcomes but particularly those closely associated with school: grades, achievement tests, other intelligence tests, and occupational status (e.g., Jensen, 1998; Sorjonen et al., 2013). Such findings are enabled partly by test instruments that rely on discrete items and decontextualized, standardized conditions from which *g* and its descendants have commonly been derived and which they also purportedly explain. In contrast, MI seeks to identify the intellectual capacities that enable human beings to assume the range of adult endstates that are valued within and across cultures. Therefore, the theory might contribute to understanding not only the traditionally measured cognitive capacities that are the hallmarks of scientists and lawyers but also the cognitive capacities of artists, entrepreneurs, musicians, animal trainers, peace makers, athletes, or pilots.

To identify the intelligences that might explain the range of adult endstates, Gardner reviewed diverse bodies of research literature. These included psychometric studies as well as studies in anthropology, neuroscience, developmental psychology, and evolutionary biology and studies of special populations including savants and prodigies. During this investigation, Gardner also formulated a set of eight criteria against which candidate intelligences could be screened. The criteria included characteristic developmental trajectory and selective sparing or breakdown among brain-damaged individuals. In addition, an intelligence should be distinguishable by its neural structures and functions, core information processing, experimental tasks, characteristic forms of symbolic representation, and evolutionary biology. Gardner stipulated that a candidate intelligence should meet all, or nearly all, of the criteria (see Table 27.1).

These criteria can be illustrated by considering two potential intelligences: linguistic and bodily-kinesthetic. The former is clearly manifested in psychometric testing. Its typical developmental trajectory, from incipient expression to proficient usage, is rapid and distinct from that of mathematics or music. In addition, linguistic intelligence is selectively lost or spared among stroke victims. Linguistic intelligence is also distinguished by neural structures associated with it (e.g., Broca's and Wernicke's areas) and entails core information processes for syntax and phonology. Evidence for this intelligence is found in experimental tasks; for example, neonates respond differently to speech versus nonspeech sounds with similar temporal and spectral qualities

Table 27.1 *Criteria for the identification of an intelligence (after Davis et al., 2011; Gardner, 1983; Kornhaber et al., 2004).*

- It should be seen in relative isolation in prodigies, autistic savants, stroke victims, or other exceptional populations. In other words, certain individuals should demonstrate particularly high or low levels of a particular capacity in contrast to other capacities.
- It should have a distinct neural representation – that is, its neural structure and functioning should be distinguishable from that of other major human faculties.
- It should have a distinct developmental trajectory – that is, different intelligences should develop at different rates and along paths that are distinctive.
- It should have some basis in evolutionary biology. In other words, an intelligence ought to have a previous instantiation in primate or other species and putative survival value.
- It should be susceptible to capture in symbol systems, of the sort used in formal or informal education.
- It should be supported by evidence from psychometric tests of intelligence.
- It should be distinguishable from other intelligences through experimental psychological tasks.
- It should demonstrate a core, information-processing system – that is, there should be identifiable mental processes that handle information related to each intelligence.

(Vouloumanos & Werker, 2007). The criterion of evolutionary biology is met in the form of gesturing among apes and their use of specific rumblings to indicate the presence of dangerous animals. Linguistic intelligence is also captured in varied symbol systems, from cuneiform to sign language to Morse code.

In contrast, bodily-kinesthetic capacities are not measured within psychometric models. However, evidence for bodily-kinesthetic intelligence exists in a developmental trajectory from infancy to mature use that is more variable than that of language, with some uses (e.g., neurosurgery) taking many years of practice. Bodily-kinesthetic intelligence is selectively lost among brain-damaged individuals suffering apraxias, even as such individuals may retain speech comprehension or compose music. Bodily-kinesthetic intelligence has distinct neural structures – for example, the motor strip, which supports the execution of coordinated movement. Evidence from evolutionary biology can be seen in the form of tool use by primates. It is captured in various symbol systems – for example, choreography diagrams and play sheets used in American football.

By applying these criteria to wide-ranging research literature, Gardner initially identified seven intelligences: linguistic, logical-mathematical, spatial, musical, bodily-kinesthetic, interpersonal, and intrapersonal. In 1995, he identified an additional intelligence, naturalist: "the ability to make consequential distinctions among organisms and entities in the natural world" (Gardner, 2011a, p. xiv; see Table 27.2). Gardner has considered other potential intelligences, among these, existential,

Table 27.2 *Gardner's eight intelligences (Davis et al., 2011; Gardner, 1983, 1999a).*

Intelligence	Description
Linguistic	An ability to analyze information and create products involving oral and written language such as speeches, books, and memos.
Logical-Mathematical	An ability to develop equations and proofs, make calculations, and solve abstract problems.
Spatial	An ability to recognize and manipulate large-scale and fine-grained spatial information
Musical	An ability to produce, remember, and make meaning of different patterns of sound.
Naturalist	An ability to identify and distinguish among different types of plants, animals, and weather formations that are found in the natural world.
Bodily-Kinesthetic	An ability to use one's own body to create products or solve problems.
Interpersonal	An ability to recognize and understand other people's moods, desires, motivations, and intentions.
Intrapersonal	An ability to recognize and understand one's own moods, desires, motivations, and intentions.

pedagogical, and moral, though none has adequately met the criteria for inclusion in the theory (Gardner, 1999a).

Certainly, other psychologists have argued that intelligence is not unitary. However, unlike Gardner, their theories rested largely on psychometric data. Among contemporary theorists, Robert Sternberg has posited a triarchic theory of intelligence, comprised initially of componential, experiential, and contextual subtheories, the evidence for which came from tests as well as analyses of questionnaire data (Sternberg, 1985). Later versions posited analytical, practical, and creative intelligences, each supported by evidence from a test developed for the theory (e.g., Sternberg et al., 2001). The roots of a multifaceted versus general intelligence can be found in Binet and Simon (1916/1973), who stated, "One [child] succeeds best in test 'A' and fails in test 'B'; another, of the same age, fails in 'A' and on the other hand succeeds in 'B.' . . . intelligence is therefore not made nor can it be made as one measures height" (p. 243). Thurstone (1938) posited seven primary mental abilities; Guilford (1967) argued that intelligence was comprised of 120 and, later, 150, and, then, 180 factors arrayed in three categories. In contrast to Thurstone and Guilford, researchers using factor analysis have typically arrived at structures of intelligence that place *g* atop narrower factors or abilities (e.g., Carroll, 1993), a structure that has gained broad acceptance (e.g., Demetriou & Spanoudis, 2017; Willingham, 2004).

However structured, it is not clear what *g* and the group or narrower factors presented in hierarchical models actually represent: Is *g,* per Spearman, a single "mental energy" (Spearman, 1923, p. 5) or are *g* and group factors primarily syntheses of the test data and correlates of school and test performance? Are they based in neurobiology and, if so, how? Does *g* capture processing speed or flexibility? Does it reflect motivation to perform in, or familiarity with, decontextualized tasks, a willingness to follow instructions, or persistence (Gardner, 2006b)? Given that psychometric models have drawn data from atypical problem-solving contexts (e.g., test settings or labs that lack resources, collaborators, and time to reflect and revise), it may be that these models represent contrived forms of cognition (Davis et al., 2011a; Gardner, 2006b). In contrast, MI attends to the development of intelligences within and toward culturally valued domains.

Clarifications and Caveats

The intelligences Gardner posited in MI theory have sometimes been conflated with other concepts or have been misunderstood. These misunderstandings are addressed and clarified in this section.

First, an intelligence is not dependent on a given sensory system. For example, spatial intelligence is not dependent on the visual system, though such conflation is captured in mislabelings, such as visual-spatial intelligence. In fact, spatial problem-solving is often keenly developed among the blind, as is evident from their ability to navigate without sight. Even musical intelligence can be manifested among those with severe hearing loss (Kolb, 2017). Gardner differentiates conceptually between sensory systems, which allow for perceptual input from the world, and intelligences, which are biopsychological computational capacities that operate on information that can be conveyed from various sensory systems.

Second, given that intelligences are computational capacities, they are not synonymous with domains, though an intelligence and a domain may have similar monikers – for example, musical intelligence and music. A domain consists of a body of knowledge and skills that exists within one or more societies and for which there are practitioners who vary from novice to expert (Gardner, 2006a).

Third, while the intelligences can be isolated using the criteria Gardner has stipulated, they are rarely applied in isolation to real-world problem-solving. For example, a mathematician will clearly draw on logical-mathematical intelligence. However, she also likely draws on linguistic intelligence to read papers written by other mathematicians as well as on interpersonal intelligence in collaborating with others to tackle complex problems. A clinical psychologist depends on interpersonal intelligence but also relies on linguistic intelligence to interpret clients' speech and likely uses bodily-kinesthetic intelligence in interpreting clients' body language. Indeed, only in brain-damaged patients or individuals who present with symptoms like autism can one observe intelligences spared or destroyed in isolation.

Fourth, MI theory has been mistakenly conflated with ideas about learning styles (Gardner, 2013). In contrast to MI, the evidence and criteria for positing different learning styles theories rest largely on psychometric test batteries. In addition, learning styles theories often include constructs pertaining to both cognition and personality. Unlike MI, learning styles theorists often hold that individuals tend to employ a consistent style in their problem-solving, taking, for example, a visual approach to whatever cognitive task lies before them (see Zhang, 2011). There is no assumption under MI that individuals rely on any individual intelligence or cluster of intelligences across the range of real-world problem-solving.

Fifth, while the intelligences are universally present among the neurologically intact, individuals vary with regard to their "profiles of intelligence." Gardner further claims that no two individuals, not even identical twins, will have exactly the same profile. That is, one individual may be relatively strong in spatial intelligence and linguistic intelligence, less strong in naturalist, mathematical, interpersonal, and intrapersonal intelligences, and weak in musical intelligence. Another individual may have roughly the opposite profile. Such heterogeneity stands in contrast to the positive manifold arising from test data found by Spearman (1904) and later proponents of *g*.

Sixth, although individuals will manifest relative strengths among the intelligences, this does not indicate that all individuals are gifted in one or more of the intelligences. In the same vein, relative weaknesses do not indicate that each person suffers from at least one intellectual deficit. These are romantic views, which lack a factual basis.

Seventh, intelligences are modifiable. Environmental opportunities and supports will influence the development and expression of intelligences and the profiles of intelligences.

Eighth, profiles of intelligence or a given strength or weakness among the intelligences should not be associated with claims of innate racial or ethnic characteristics or differences. Gardner has taken particular care to undermine such assertions and has described such claims as “offensive” (Gardner, 2011a).

Further Developments of the Theory

The set of intelligences Gardner has identified has been stable since 1995. However, Gardner does not claim that the intelligences identified thus far are the definitive set. His primary claim is that human intelligence is not singular or general but rather multifaceted. Additional intelligences may occur as new data and evidence arise. At the time the theory was first proposed, it was grounded in notions of domain performance and the set of intellectual abilities that would explain endstates valued across cultures. Should some valued roles emerge that cannot be understood in terms of the intelligences, it may be that further investigation of such roles will prompt reconsideration.

Further, MI theory was not grounded in neuroimaging studies, technology for which was being advanced and made more accessible during the 1980s when the theory was first published. Such studies are now possible. Their preliminary findings support the existence of distinguishable neural structures for each of the intelligences (Shearer, 2018). Yet it could also be the case that additional neuroimaging may find different patterns and a reconsideration of the distinctions among intelligences would be justified.

Critiques of the Theory

MI theory has been criticized on varied fronts. Critics have argued that the theory suffers from conceptual murkiness, that it is not grounded in evidence or scientific, and that its claim of relative autonomy across the intelligences is incorrect (see Schaler, 2006).

Critics of the theory argued that the use of the term intelligences undermined theoretical clarity about the construct of intelligence (e.g., Barnett, Ceci, & Williams, 2006) and that “it is not at all clear what we gain by referring to such skills, competences, and abilities as ‘intelligences’” (Kanazawa, 2010, p. 281). It is common for these critics to hold that the term intelligence ought to refer to abstract reasoning, planning, and the ability to grasp complex ideas readily that might be applicable in broad contexts but especially valued in school, as well as be predictive of occupational status and assessed by intelligence tests or tests of processing speed (Brody, 2006; Gottfredson, 1997). Gardner asserts such boundaries and tests partly reflect a narrow Western stance on what intelligence is and does not account for cross-cultural research that has repeatedly found low-scoring individuals functioning at high levels in their own societies (Gardner, 2006b). The construct of intelligence has yet to secure an agreed-on definition, even among psychologists investigating it (e.g., Hauser, 2010; Neisser et al., 1996; Sternberg & Detterman, 1986).

White (2006) has taken issue with the claim that the set of intelligences enables problem-solving that is valued across cultures. He states it is not possible to know all the types of problem-solving that have been valued. This issue does not undermine the theory. The theory speaks to the set of cognitive capacities that, together, can tackle problems or fashion products that are known to be valued. Should a domain arise for which the current set of intelligences is insufficient to explain, then the theory may be modified. Indeed, such a conundrum helped give rise to the naturalist intelligence (Gardner, 2006a). In addition, White (2006) argues that some intelligences may not have been equally valued across cultures. To wit, not all cultures equally value expressions of music. However, Gardner did not stipulate that all the intelligences must be equally valued across cultures. He does stipulate that individuals across cultures will possess the full set of intelligences. The contribution of the environment to the development of the intelligences will vary depending in part on what a culture values.

Critics have also asserted that the intelligences cannot be relatively autonomous, since psychometric studies continually find positive correlations among aptitudes (Visser, Ashton, & Vernon, 2006; Waterhouse, 2006a; Willingham, 2004). As noted, Gardner has responded that such correlations arise in part from the atypical tasks and settings psychometric studies feature. In any event, the existence of a positive manifold does not in itself invalidate the utility of positing and assessing intellectual capacities that may have considerable autonomy.

Critics have argued that MI is not based on evidence. Yet, as already noted, Gardner drew on evidence from a variety of disciplines. More central to this critique is the claim that MI ignores or misinterprets psychometric evidence (Brody, 2006; van der Ploeg, 2016; Willingham, 2004). The theory clearly does not ignore psychometrics, since evidence from psychometrics is one of the criteria for evaluating a candidate intelligence.

Relatedly, the theory has been criticized as unscientific, particularly because studies that might test the theory's validity or overthrow it via experimental methods are absent or inadequate (Jensen, 2008; van der Ploeg, 2016; Waterhouse, 2006a, 2006b). This critique holds that psychometric data are central to such tests of the theory. Per Jensen (2008), "Because Gardner's theory of mental abilities remains aloof from research based on measurement and analysis in the tradition of the natural sciences, it has no means for proving itself to be more correct than the model of intelligence that has emerged from the London School" (p. 97). Gardner has responded in varied ways to critiques that the theory is untestable. He notes that, while he does not think the theory is subject to an up or down testing regime, psychometric evidence could be brought to bear on the theory (Gardner, 2011a). Some studies (e.g., Visser, Ashton, & Vernon, 2006) using factor analysis of tasks reflecting Project Spectrum (detailed in the section "Efforts to Assess Individuals' Multiple Intelligences") find that the intelligences are not relatively autonomous and instead can be explained by *g*. Other factor analytic studies using similar data find support for relatively autonomous intelligences (Plucker, Callahan, & Tomchin, 1996). Still others take a middling position, finding that the intelligences are neither as distinct from *g* as Gardner claims nor wholly explainable by *g* (Almeida et al., 2009; Castejón, Perez, & Gilar, 2010). Such divergent results are unsurprising given decades of debate among psychometric researchers about the structure of intelligence or intelligences.

While psychometric tests of MI are likely to be inconclusive, Gardner (2011a) has noted that evidence from neuroscientific research might disconfirm or modify the theory. For example, if such studies find a particular configuration of neural structures provide for equally strong performance across the intelligences, this would support *g* and undermine MI. In contrast, if distinct profiles are associated with characteristic neural structures, then this would provide support for the theory at the expense of *g* (Davis et al., 2011). Though this point is debated (see Gardner & Moran, 2006; Waterhouse, 2006a, 2006b), there is increasing, not decreasing, evidence from neuroscience to support distinctions in the neural processing of capacities identified by MI (Gardner, 2011a; Gardner & Moran, 2006). For example, Shearer (2018) has examined more than 500 studies (primarily fMRI experiments) and concludes that the core processing skills are well supported by neuroscientific evidence. In a separate investigation of more than 420 studies, he found that neural activation of regions associated with the different intelligences varied among individuals who were impaired, typical, or skilled. The theory could be supported or refuted by analogous research into the genetics of intelligence (Gardner, 2011a): Should some pattern of genes support equally high functioning across intelligences, then the argument for *g* over MI would be supported. If patterns of genes were associated with particular intelligences or profiles of intelligence, then the evidence for MI, rather than *g*, would be strengthened.

Even as Gardner has illustrated how empirical studies could test MI, he has also stated,

> I've never felt that MI theory was one that could be subjected to an "up and down" kind of test, or even series of tests. Rather, it is and has always been fundamentally a work of synthesis; and its overall fate will be determined by the comprehensiveness of the synthesis, on the one hand, and its utility to both scholars and practitioners, on the other. (Gardner, 2011a, pp. xix–xx)

Gardner's stance and that of MI's critics regarding how the theory might be undermined or sustained reflect philosophical differences in epistemology and conceptions of science. MI was not constructed through formal hypothesis testing and experimental design (Sternberg, 2012). Instead, it is what Einstein called a *constructive theory*, one that offers a reasonable model for understanding a given phenomenon (e.g., variation in human intelligence as manifested in domains across cultures) versus a *principle theory*, which is built on confirmed, empirical generalizations (Howard, 2017). For MI, the latter is becoming possible via neuroscience (Gardner, 2011a; Shearer, 2018) and potentially through new research in genetics.

Efforts to Assess Individuals' Multiple Intelligences

Gardner argues that psychometric methods and laboratory tasks are not adequate for examining multiple intelligences. Psychometric tests require individuals to channel their intelligences via media (words, numbers, pictures) or tools (keyboards, pencils, speech) that are inadequate or narrow. Moreover, they can be deceptive. For example, musical intelligence enables individuals to hear and analyze a chord progression. However, individuals who have little by way of musical intelligence or training might learn to analyze a written chord progression without being able to recognize or perform the sounds or to appreciate how a given chord progression might evoke responses. Analogous challenges exist for most of the other intelligences that are not typically tested and even for spatial intelligence, which frequently is via items requiring the spatial rotation of two-dimensional figures. A two-dimensional rotation item also cannot evaluate whether individuals can interpret or create images intended to represent concepts or to navigate among locations when digital maps are unavailable.

Gardner does not believe testing of the intelligences is necessary in most cases – a view not unlike Binet, who said that intelligence testing should be limited to those individuals who struggle within typical classrooms (Binet & Simon, 1916/1973). To the extent that it is important to assess intelligences, Gardner argues that such assessments need to be "intelligence-fair." That is, they should allow for the direct expression of the intelligence, using appropriate resources (e.g., musical instruments, other people, objects that can be taken apart and reassembled), and take place in environmental settings that are familiar (vs. isolated, test-administrator–controlled settings). Such assessments are both intelligence-fair and authentic – they mirror the kinds of problem-solving that are valued in the world beyond the test or school (Gardner, 1995).

Via Project Spectrum, a research endeavor within Project Zero, Gardner and colleagues developed a more intelligence-fair approach to assessing young children's "profiles of intelligence." Spectrum assessment tasks were embedded in the regular classroom and available for students to explore and interact with on a regular basis. Thus, they broke down the barrier between assessment and ongoing curricular and classroom activities. This approach is in line with Spectrum's emphasis on both identifying children's profiles of intelligences and drawing on that knowledge to foster individualized learning opportunities. Spectrum relied on hands-on, engaging activities in language, mathematics, visual arts, music, science, movement, and social understanding.

For example, in the assessment of social understanding, students made use of a scale model of their own classroom and the students within it, whose photographic likenesses were affixed to movable wooden pegs. Spectrum researchers, who spent time in the classroom with the students and instructors, could then interact with students using the model to assess students' understanding of social relationships within the class. An example of a mathematics activity was the bus game, in which a model bus was moved about several bus stops at which passengers got on or off (Chen, Krechevsky, & Viens, 1998; Krechevsky, 1998). Spectrum tasks demonstrate reliability (Chen, 1998; Chen, Krechevsky, & Viens, 1998). Spectrum researchers also found that young children exhibited distinctive profiles of intelligences and that the intelligences were relatively autonomous (Gardner & Hatch, 1989). Hatch (1997) reported that profiles exhibited in a small group of Spectrum kindergarteners manifested similarities and differences six years later. Hatch (1997) held that some change in observed profiles of intelligences should be expected, given that intelligences are demonstrated in activities and social contexts and that these will undergo change in the six years following kindergarten. Such change also reflects the role of context in the development and expression of intelligences.

Project Spectrum was the only assessment of intelligences in which Gardner was directly involved but many other assessments have been developed outside of Project Zero. Many have no credible basis whatsoever or consist of brief paper-and-pencil tests, which are at odds with intelligence-fair, authentic, or ecologically valid approaches.

One assessment that does incorporate these approaches is Web-Observation (Nicolini, 2011). "Web-Ob" is an online platform that facilitates teachers' daily observations of students during ordinary classroom activities. Web-Ob enables teachers to organize, manage, store, and retrieve records of observations, as well as to generate reports related to frequency and about examples of children's manifestation of the different intelligences within the classroom and report profiles of intelligence manifested at different points in time (Nicolini, 2011; Nicolini, Alessandri, & Bilancioni, 2010). Web-Ob is accompanied by ongoing professional development of teachers about MI and about producing descriptive observations of their students' activities (Nicolini, 2011; Nicolini, Alessandri, & Bilancioni, 2010).

The most well-known assessment developed outside of Project Zero is the Multiple Intelligences Developmental Assessment Scale (MIDAS). The MIDAS provides a structured self-report in which individuals can use both quantitative and qualitative descriptions of their skills and abilities (Shearer, 2012). It has been translated into several languages, administered worldwide, and used in a variety of research projects (Shearer, 2007). Self-reports are not necessarily accurate appraisals of one's own cognitive abilities. However, Shearer has undertaken research to compare individuals' self-reports with reports by informants who are knowledgeable about the individual. He found the inter-reliability between individuals' self-reports and reports by knowledgeable informants to be fairly strong. Of 742 paired comparisons of assessments of strengths of individuals' multiple intelligences on the MIDAS, 46 percent of comparisons generated exactly the same rating on a five-point scale and 92 percent of ratings were within one scale point (Shearer, 2012).

Educational Influence of the Theory of Multiple Intelligences

Within a few years of the publication of Gardner's (1983) book *Frames of Mind: The Theory of Multiple Intelligences*, educators in the United States began drawing on MI (Armstrong, 2017; Campbell, Campbell, & Dickinson, 2003; Hoerr, 2000; Kunkel, 2009). By the mid-1990s, the theory was being widely used across US schools and in colleges, museums, and other settings. The theory has been used with diverse learners, including those with special needs (Hearne & Stone, 1995; Takahashi, 2013), gifted students (Callahan et al., 1995; Hernández-Torrano et al., 2014; Maker, 2005), and adults (Kallenbach & Viens, 2002). It has spread to educators in diverse parts of the world (e.g., Chen, Moran, & Gardner, 2009; Nicolini, 2011; Pienaar, Nieman, & Kamper, 2011). A science theme park in Denmark drew on MI to communicate about science and to enable visitors to explore their intelligences (Sahl-Madsen & Kyed, 2009). More recently, game designers have investigated MI's utility for adapting learning games for different players (Sajjadi, Vlieghe, & De Troyer, 2017).

Over time, and given MI's broad implementation, Gardner elaborated on the theory's educational implications, which were only briefly sketched in *Frames of mind*. Two main implications are *individuation* and *pluralization*. The former entails knowing each student well and using knowledge of their profiles of intelligence to provide varied ways for each to learn and demonstrate their understanding (Gardner, 1999a). Pluralization entails conveying what it is that is important for students to learn and understand in a variety of ways. Each may be enabled by digital resources as well as other means (Gardner, 2015). In addition, pluralization may be fostered via curricula and instruction that engage different "entry points" (Gardner, 1991, 1999b). Gardner (1991, 1999b) developed the entry points framework as a way of bridging the richness of disciplines and the complexity of individuals' profiles of intelligence. He has argued that worthwhile curricular topics can be approached through each entry point and thereby made accessible to all learners (Table 27.3). In addition, Gardner stated that the theory should not be an end in and of itself but rather serve as a means to pursue and achieve valued educational ends established by the cultures surrounding schools (Gardner, 1999a).

One powerful end for which MI could provide a useful means is disciplinary understanding. This entails the acquisition of knowledge and skills and forms of analyses pertinent to science, math, history, psychology, and other domains, as well as the application of skills and knowledge to new material and questions (Gardner, 1999a, 2006a).

Research over the course of ten years has surfaced five explanations about why educators adopt MI. First, the theory resonates with educators' everyday experiences and observations that students manifest different capacities and sets of capacities. Put otherwise, MI serves as a *constructive theory* for representing variation among students. Second, it provides a vocabulary for educators to think more systematically about differences, strengths, and needs among their students. Third, MI enables richer communications with colleagues and students' families about learners' strengths and needs, especially relative to the communications enabled by test results (e.g., "proficient," "at the 40th percentile," "two standard deviations about the mean"). Fourth, the theory provides a framework for educators to reflect on their own practice – a kind of mental closet organizer for the many different activities, materials, and instructional strategies educators employ. Finally, educators' reflection via MI fosters their efforts to develop learning environments that support the varied learners with whom they work (Kornhaber, 2004; Kornhaber & Krechevsky, 1995).

There is little doubt that implementations of the theory are widely variable. There are no permissions needed to adopt the theory and no clearing house for reporting

Table 27.3 *The entry points framework and an illustration of application to the topic of evolution (Gardner 1991, 1999b; Kornhaber et al., 2004).*

Narrative	The narrative entry point deals with the story or stories that are central to a topic. Typically, a rich or "generative" topic will offer several possible narrative entry points, some of which may be recounted or performed as dramatic narratives. For instance, for the topic of evolution, there is the narrative involving Darwin's own life, his voyage to the Galapagos Islands, or even various traditional folk stories about how different animals and plants came to have their unique form.
Logical-Quantitative	The logical-quantitative entry point focuses on numerical aspects of a topic and/or on deductive, logical reasoning, of the sort that can often be captured by if-then syllogisms. A more quantitative entry point for the topic of evolution might entail looking at Darwin's effort to map the distribution of different species across different islands. A logic-focused entry point might pose syllogisms for the students to explore: If there were no variation within a species, then what might happen when its environment changed?
Aesthetic	The aesthetic entry point engages artistic aspects of, or representations of, a topic. An aesthetic entry point for evolution might be to examine different drawings Darwin made of finches or other species he studied on the Galapagos and to describe how their shapes/morphologies differ.
Experiential ("Hands-on")	This entry point provides students opportunities to do work involving the physical "stuff" of the topic. For example, for the topic of evolution, students might breed fruit flies, or do virtual simulations of evolutionary processes, and document what they observe.
Interpersonal	The interpersonal entry point involves working with others to learn about a topic. One way to incorporate the interpersonal entry point in the topic of evolution is to form research teams to carry out real or simulated experiments in breeding fruit flies.
Existential/Foundational	This entry point deals with fundamental, philosophical questions about the nature of the topic, why it exists, and/or what is its meaning or purpose. For the topic of evolution, this entry point might explore questions such as "Why are new species created and others die out?" and "What is the purpose of variation within species?"

how the theory is implemented. This raises questions about whether the theory can be associated with any particular practices among teachers or with any changes among students. These issues have been most extensively investigated during the Schools Using MI Theory (SUMIT) study (Kornhaber, Fierros, & Veenema, 2004). As part of

Table 27.4 *SUMIT's Compass point practices (Kornhaber et al., 2004).*

Readiness	Educators took time to study the theory and explore how it might be applied. On average, it took about eighteen months of such study to build readiness to implement MI in classrooms.
Culture	The culture of the school was marked by beliefs in students' strengths and potential, advocacy of care and respect among all members of the school community, a sense of excitement about learning, and by persistence, dedication, and hard work by educators.
MI as Means to High-Quality Work	Educators use MI as a means to help learners acquire knowledge and skills in the disciplines.
Collaboration	Educators came to see variations in strengths, knowledge, and skills among their colleagues as resources for improving curricula and teaching. This supported ongoing informal and formal collaboration among teachers.
Meaningful Choice	Students were given meaningful curricular and assessment options. Such options provided routes for students to draw on their profiles of intelligences to produce high-quality work and demonstrate their understanding.
Arts	The arts played a vital role, both in formal studies of arts disciplines and as a means of fostering students' understanding of the range of other disciplines.

SUMIT, researchers conducted qualitative interviews among an intentional sample of forty-one public schools with diverse populations in eighteen US states and one Canadian province that used MI for three or more years. Of these schools, 49 percent associated improved test scores with MI; 54 percent associated improvements in student discipline with MI; 60 percent reported improvement in parent participation associated with MI; 78 percent associated the theory with improvements for students with learning disabilities; and 2 percent reported improvements for that population not associated with MI (Kornhaber, Fierros, & Veenema, 2004).

SUMIT researchers also conducted school case studies, including classroom observations, interviews with teachers, and documentation of student work. This, together with the interviews from forty-one schools, enabled them to identify five "compass point practices" that were common to schools using MI (see Table 27.4).

The SUMIT study has been critiqued for not reporting statistical significance, using control groups, offering causal claims, or accounting for changes that might be due to other factors (van der Ploeg, 2016; Willingham, 2004), though the latter were clearly reported (see Kornhaber, Fierros, & Veenema, 2004, pp. 13–16). These critiques evince a limited understanding of qualitative research, which typically does not seek to generate causal explanations via control groups, and the goals of the SUMIT research in particular, which focused on identifying the practices educators used to implement MI and the changes they associated with the implementation of MI.

There are studies of MI whose designs meet calls for the theory's educational utility to be demonstrated via a randomized control design. For example, Nguyen (2000) investigated learning outcomes in one school that participated in a pilot site visit for the SUMIT study. That school randomly assigned students to clusters of faculty that were or were not implementing MI. Because Nguyen found no difference in average achievement scores between the two groups, Van der Ploeg (2016) claimed the study showed MI makes no difference. Yet there were substantial differences in the spread of scores, with much less variance around the mean among students in the MI treatment (Nguyen, 2000). One reasonable inference from this finding is that, when MI is used thoughtfully, it can help foster more equitable achievement.

Certainly, additional quantitative and qualitative studies could be valuable in examining how MI is used and the variations in implementations that are most beneficial and whether any may have downsides. However, it is not the case that only randomized controlled trials can answer such questions. Indeed, demands that research on educational interventions must employ such designs carry problems of their own (see, e.g., Erickson & Gutierrez, 2002; Ginsburg & Smith, 2016).

It is also useful to consider MI against educational practices prompted by *g*-based conceptions of intelligence. Allowing that IQ testing may help identify and then also provide enriched educational opportunities to students who may have learning disabilities – a practice Binet and Simon (1916/1973) said should be the primary reason for intelligence testing – the track record of *g*-based interventions is problematic. For much of the twentieth century, IQ testing fostered markedly different learning opportunities along socioeconomic and racial lines (Callahan, 1962; Oakes, 1985). It also promoted the view that intelligence was unmodifiable – a belief shown to undermine student learning (Blackwell, Trzesniewski, & Dweck, 2007) – while frequently reinforcing harmful stereotypes of intellectual inferiority among students from historically disadvantaged groups (Herrnstein & Murry, 1994; Jensen, 1969; Terman, 1916).

The Future of MI in Education and Psychology

While MI was posited and first implemented in the United States, in the last few decades government policies requiring schools to raise math and English language test scores have often narrowed curricula and pedagogy in ways that undermine educators' use of the theory (Kornhaber, 2009). Yet the theory is still being used by US educators and is continuing to be adopted around the world. In 2017 and 2018, an interactive online MI course launched from the Harvard Graduate School of Education enrolled teams of pre-K through professional school educators from five continents and some twenty nations, among them Cypress, Vietnam, Lithuania, China, Indonesia, Peru, Turkey, and India as well as the United States. Thus, it appears that MI continues to be a constructive theory for educators.

Within psychology, MI has stood as a provocation to normative psychometric conceptions of intelligence. The theory's foundations are empirical, though not restricted to psychometric evidence. Gardner himself has largely moved on from the theory to other research (e.g., Gardner, 2011b, 2018; Gardner & Davis, 2013). Others have tested the theory's claims by using psychometric methods and have come to disparate conclusions regarding the relative autonomy of the intelligences (Almeida et al., 2009; Castejón, Perez, & Gilar, 2010; Visser, Ashton, & Vernon, 2006; Plucker, Callahan, & Tomchin, 1996; Waterhouse, 2006a; Willingham, 2004). New research involving neuroscience has begun to test such claims. Research involving genetics could also be brought to bear. Through such research, MI may be refuted. Or it may allow the theory to come to rest as both a constructive and a principle theory of intelligence.

References

Almeida, L. S., Prieto, M. D., Ferreira, A. I., Bermejo, M. R., Ferrando, M., & Ferrándiz, C. (2009). Intelligence assessment: Gardner multiple intelligence theory as an alternative. *Learning and Individual Differences*, *30*(3), 225–230.

Armstrong, T. (2017). *Multiple intelligences in the classroom* (4th ed.). Alexandria, VA: ASCD.

Barnett, S. M., Ceci, S. J., & Williams, W. M. (2006). Is the ability to make a bacon sandwich a mark of intelligence?, and other issues: Some reflections on Gardner's theory of multiple intelligences. In J. Schaler (Ed.), *Howard Gardner under fire: The rebel psychologist faces his critics* (pp. 95–114). Chicago: Open Court.

Binet, A., & Simon, T. (1916/1973). *The development of intelligence in children. (The Binet-Simon Scale*) (trans. Elizabeth S. Kite). New York: The Arno Press.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C.S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263.

Brody, N. (2006). Geocentric theory: A valid alternative to multiple intelligences. In J. Schaler (Ed.), *Howard Gardner under fire: The rebel psychologist faces his critics* (pp.73–94). Chicago: Open Court.

Callahan, C. M., Tomlinson, C. A., Moon, T. R., Tomchin, E. M., & Plucker, J. A. (1995). Project START: Using a multiple intelligences model in identifying and promoting talent in high-risk students (RM 95136). Charlottesville: University of Virginia, National Research Center on the Gifted and Talented.

Callahan, R. (1962). *Education and the cult of efficiency.* Chicago: University of Chicago Press.

Campbell, L., Campbell, B., & Dickenson, D. (2003). *Teaching and learning through multiple intelligences*. Boston: Allyn & Bacon.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Castejón, J. L., Perez, A. M., & Gilar, R. (2010). Confirmatory factor analysis of Project Spectrum activities. A second-order g factor or multiple intelligences? *Intelligence*, 38(5), 481–496.

Chen, J.-Q., Isberg, E., & Krechevsky, M. (1998). *Project Zero frameworks for early childhood education: Volume 2, Project Spectrum early learning activities*. New York: Teachers College Press.

Chen, J.-Q, Krechevsky, M., Viens, J., & Isberg, E. (1998). *Building on children's strengths: The experience of Project Spectrum*. New York: Teachers College Press.

Chen, J.-Q., Moran, S., & Gardner, H. (2009). *Multiple intelligences around the world*. San Francisco: Jossey-Bass.

Davis, K., Christodoulou, J., Seider, S., & Gardner, H. (2011). The theory of multiple intelligences. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence*. Cambridge, UK: Cambridge University Press.

Demetriou, A., & Sanoudis, G. (2017). From cognitive development to intelligence: Translating developmental mental milestones into intellect. *Journal of Intelligence*, *5*(3), 30. https://doi.org/10.3390/jintelligence5030030

Erickson, F., & Gutierrez, K. (2002). Culture, rigor, and science in educational research. *Educational Researcher*, *31*(8), 22–24.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gardner, H. (1991). *The unschooled mind*. New York: Basic Books.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.

Gardner, H. (1995). Reflections on multiple intelligences: Myths and messages. *Phi Delta Kappan*, *77*, 200–203, 206–209.

Gardner, H. (1999a). *Intelligence reframed: Multiple intelligences for the 21st Century*. New York: Basic Books.

Gardner, H. (1999b). *The disciplined mind: What all students should understand*. New York: Simon & Schuster.

Gardner, H. (2006a). *Multiple intelligences: New horizons*. New York: Basic Books.

Gardner, H. (2006b). Replies to my critics. In J. Schaler (Ed.), *Howard Gardner under fire: The rebel psychologist faces his critics* (pp. 277–344). Chicago: Open Court.

Gardner, H. (2009). Birth and the spreading of a "meme." In J. Q. Chen, S. Moran, & H. Gardner (Eds.). *Multiple intelligences around the world* (pp. 3–16). San Francisco: Jossey-Bass.

Gardner, H. (2011a). Multiple intelligences: The first thirty years. In H. Gardner (Ed.), *Frames of mind: The theory of multiple intelligences* (3rd ed., pp. ix–xxv). New York: Basic Books.

Gardner, H. (2011b). *Truth, beauty, and goodness: Educating for the virtues in the age of truthiness and Twitter.* New York: Basic Books.

Gardner, H. (2013). Multiple intelligences are not learning styles. *Washington Post*, October 16. https://washingtonpost.com/news/answer-sheet/wp/2013/10/16/howard-gardner-multiple-intelligences-are-not-learning-styles/?noredirect=on&utm_term=.018cc5ae35c8

Gardner, H. (2015). *Espacio de Pensamiento e Innovación Educativa* (*Space for Thinking and Educational Innovation*). https://howardgardner01.files.wordpress.com/2015/10/12070suplemento-eniac.pdf

Gardner, H. (2018). "Why are you doing that research?" https://howardgardner.com/category/life-long-learning-a-blog-in-education

Gardner, H., & Davis, K. (2013). *The app generation: How today's youth navigate identity, intimacy, and imagination in a digital world*. New Haven, CT: Yale University Press.

Gardner, H., & Hatch, T. (1989). Educational implications of the theory of multiple intelligences. *Educational Researcher*, *18*(8), 4–10. https://doi.org/10.3102/0013189X018008004

Gardner, H., Ling, K., Flamm, L., & Silverman, J. (1975). Comprehension and appreciation of humor in brain-damaged patients. *Brain*, *98*, 399–412.

Gardner, H., & Moran, S. (2006). The science in multiple intelligences: A response to Lynn Waterhouse. *Educational Psychologist*, *41*, 227–232.

Gardner, H., & Wolf, D. (1983). Waves and streams of symbolization. In D. R. Rogers & J. A. Sloboda (Eds.), *The acquisition of symbolic skills*. London: Plenum Press.

Gardner, H., Zurif, E. B., Berry, T., & Baker, E. (1976). Visual communication in aphasia. *Neuropsychologia*, *14*, 275–292.

Ginsburg, A., & Smith, M. S. (2016). Do randomized controlled trials meet the gold standard? Washington, DC: American Enterprise Institute.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*, 13–23.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Hatch, T. (1997). Getting specific about multiple intelligences. *Educational Leadership*, *54* (6), 26–29.

Hauser, R. M. (2010). Causes and consequences of cognitive functioning across the life course. *Educational Researcher*, *9*, 95–109. https://doi.org/10.3102/0013189x10363171

Hearne, D., & Stone, S. (1995). Multiple intelligences and underachievement: Lessons from individuals with learning disabilities. *Journal of Learning Disabilities*, *28*(7), 439–448.

Hernández-Torrano, D., Ferrándiz, C., Ferrando, M., Prieto, L., & Fernández, M.C. (2014). The theory of multiple intelligences in the identification of high-ability students. *Anales de Psicología*, *30*(1). http://dx.doi.org/10.6018/analesps.30.1.148271

Herrnstein, R., & Murry, C. A. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Hoerr, T. (2000). *Becoming a multiple intelligences school*. Alexandria, VA: ASCD.

Howard, D. A. (2017). Einstein's philosophy of science. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. https://plato.stanford.edu/archives/fall2017/entries/einstein-philscience

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement. *Harvard Educational Review*, *39*(1), 1–123. https://doi.org/10.17763/haer.39.1.l3u15956627424k7

Jensen, A. R. (1998). *The* g *factor: The science of mental ability*. Westport, CT: Praeger.

Jensen, A. R. (2008). Review of "Howard Gardner under fire: The rebel psychologist faces his critics." *Intelligence*, *36*(1), 96–97.

Kallenbach, S., & Viens, J. (2002). *Open to interpretation: Multiple intelligences theory in adult literacy education*. Cambridge, MA: National Center for the Study of Adult Learning and Literacy, Harvard Graduate School of Education.

Kanazawa, S. (2010). Evolutionary psychology and intelligence research. *American Psychologist*, *65*(4), 279–289. https://doi.org/10.1037/a0019378

Kolb, R. (2017). Sensations of sound: On deafness and music. *New York Times*, November 3. https://nytimes.com/2017/11/03/opinion/cochlear-implant-sound-music.html

Kornhaber, M. L. (2004). Multiple intelligences: From the ivory tower to the dusty classroom – But why? *Teachers College Record*, *106*(1), 67–76.

Kornhaber, M. L. (2009). *What's policy got to do with it? A policy perspective on the theory of multiple intelligences*. In J. Chen, S. Moran, & H. Gardner (Eds.), *Multiple intelligences around the world* (pp. 374–385). San Francisco: Jossey-Bass.

Kornhaber, M. L, Fierros, E. G., & Veenema, S. A. (2004). *Multiple intelligences: Best ideas from research and practice*. Boston: Allyn & Bacon and Pearson.

Kornhaber, M. L., & Krechevsky, M. L. (1995). Expanding definitions of teaching and learning: Notes from the MI underground. In P. Cookson & B. Schneider (Eds.), *Transforming schools*. New York: Garland.

Krechevsky, M. (1998). *Project Zero frameworks for early childhood education: Volume 3, Project Spectrum preschool assessment handbook*. New York: Teachers College Press.

Kunkel, C. (2009). The world's first multiple intelligences school. In J.-Q. Chen, S. Moran, & H. Gardner (Eds.), *Multiple intelligences around the world*. San Francisco: Jossey-Bass.

Maker, C. J. (2005). The DISCOVER Project: Improving assessment and curriculum for diverse gifted learners. Storrs: The National Research Center on the Gifted and Talented, University of Connecticut.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Nguyen, T. T. (2000). Differential effects of a multiple intelligences curriculum on student performance, Unpublished doctoral dissertation, Harvard University.

Nicolini, P. (2011). How to assess intelligences through the observational method: The Italian experience. *Procedia Social and Behavioral Sciences*, *11*, 87–91. https://doi.org/10.1016/j.sbspro.2011.01.039

Nicolini, P., Alessandri, G., & Bilancioni, G. (2010). Web-Ob for multiple intelligences observation. *Procedia Social and Behavioral Sciences*, *2*(2), 728–732. https://doi.org/10.1016/j.sbspro.2010.03.092

Oakes, J. (1985). Keeping track: How schools structure inequality. New Haven, CT: Yale University Press.

Piaget, J. (1983). Piaget's theory. In P. Mussen (Ed.), *Handbook of child psychology*, vol. 1. New York: John Wiley.

Pienaar, H. C., Nieman, M. M., & Kamper, G. D. (2011). Implementing a teaching approach based on the multiple intelligence theory in a South African school: A case study. *Journal of Africa Education Review*, *2*, 267–285. https://doi.org/10.1080/18146627.2011.603235

Plucker, J. A., Callahan, C. M., & Tomchin, E. M. (1996). Wherefore art thou, Multiple Intelligences? Alternative assessments for identifying talent in ethnically diverse and low income students. *Gifted Child Quarterly*, 40(2), 81–91.

Sahl-Madsen, C., & Kyed, P. (2009). The explorama. In J.-Q. Chen, S. Moran, & H. Gardner (Eds.), *Multiple intelligences around the world* (pp. 169–183). San Francisco: Jossey-Bass.

Sajjadi, P., Vlieghe, J., & De Troyer, O. (2017). Exploring the relation between the theory of multiple intelligences and games for the purpose of player-centred game design. *Electronic Journal of E-Learning: EJEL*, *15*(4), 320–334.

Schaler J. A. (Ed.) (2006). *Howard Gardner under fire: The rebel psychologist faces his critics*. Chicago: Open Court.

Shearer, C. B. (2007). *The MIDAS: Professional manual* (rev. ed.). Kent, OH: MI Research and Consulting.

Shearer, C. B. (2012). An inter-rater reliability study of a self-assessment for the multiple intelligences. *International Journal of Psychological Studies*, *4*(3), 131–138.

Shearer, C. B. (2018). Multiple intelligences in teaching and education: Lessons learned from neuroscience. *Journal of Intelligence*, *6*(3), 38. https://doi.org/10.3390/jintelligence6030038

Sorjonen, K., Hemmingsson, T., Deary, I. J., & Bo, M. (2013). Intelligence moderates how education mediates the effect of social background on own attained occupational position. *Intelligence*, *41*(5), 529–536.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition*. London: MacMillan.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge: Cambridge University Press.

Sternberg, R. J. (2012). Intelligence. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*, 501–511. https://doi.org/10.1002/wcs.1193

Sternberg, R. J., Castejón, J. L., Prieto, M. D., Hautamäki, J., & Grigorenko, E. L. (2001). Confirmatory factor analysis of the Sternberg Triarchic Abilities Test in three international samples: An empirical test of the triarchic theory of intelligence. *European Journal of Psychological Assessment*, *17*(1), 1–16

Sternberg, R. J., & Detterman, D. K. (Eds.) (1986). *What is intelligence?* Norwood, NJ: Ablex.

Takahashi, J. (2013). Multiple intelligence theory can help promote inclusive education for children with intellectual disabilities and developmental disorders: Historical reviews of intelligence theory, measurement methods, and suggestions for inclusive education. *Creative Education*, *4*(9), 605–610.

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale*. Boston: Houghton Mifflin.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

van der Ploeg, P. (2016). Multiple Intelligences and pseudo-science. https://academia.edu/24174224/Multiple_intelligences_and_pseudo-_science

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2006). Beyond *g*: Putting multiple intelligences theory to the test. *Intelligence*, *45*(5), 487–502.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159–164. https://doi.org/10.1111/j.1467–7687.2007.00549.

Wapner, W., Judd, T., & Gardner, H. (1978). Visual agnosia in an artist. *Cortex*, *14*, 343–364.

Waterhouse, L. (2006a). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. *Educational Psychologist*, *41*, 207–225.

Waterhouse, L. (2006b). Inadequate evidence for multiple intelligences, Mozart effect, and emotional intelligence theories. *Educational Psychologist*, *41*(4), 247–255.

White, J. (2006). Multiple invalidities. In J. Schaler (Ed.), *Howard Gardner under fire: The rebel psychologist faces his critics* (pp. 45–71). Chicago: Open Court.

Willingham, D. T. (2004). Reframing the mind. *Education Next*, *4*(3), 19–24.

Zhang, L.-F. (2011). The developing field of intellectual styles: Four recent endeavours. *Learning and Individual Differences*, *21*, 311–318.

28 The Augmented Theory of Successful Intelligence

Robert J. Sternberg

My mother once told me that, although I was smart in school, I lacked common sense. Although her judgment of me was woefully mistaken – I hope – she, in effect, adumbrated and perhaps prompted what I have come to call the "theory of successful intelligence." In this chapter, I describe the theory of successful intelligence (Sternberg, 1997, 2003c, 2005, 2010b, 2015b). The history of the theory presented here has been documented, to some extent, in two earlier theoretical articles in *Behavioral and Brain Sciences* (Sternberg, 1980b, 1984), one in *Trends in Cognitive Sciences* (Sternberg, 1999c), and one in an edited volume (Sternberg, 2018b), all of which are related to this presentation.

In the first article (Sternberg, 1980b), a theory of components of intelligence ("componential subtheory of intelligence") was presented, with the argument that intelligence could be understood in terms of a set of elementary information-processing components that contributed to people's intelligence and individual differences in it. In the second article (Sternberg, 1984), the theory was expanded ("triarchic theory of intelligence") to include not just the analytical aspect of intelligence, which had been the emphasis of the earlier article, but also the creative and practical aspects of intelligence. By the third article (Sternberg, 1999c), the "theory of successful intelligence" was emphasizing not only levels of abilities but also how one capitalizes on one's strengths in abilities and compensates for or corrects one's weaknesses. The theory also emphasized the critical importance of adaptivity in intelligence rather than the importance of psychometric tests. In the fourth publication (Sternberg, in press), the theory is augmented and adds wisdom to the mix of skills central to intelligence. In particular, wisdom is the use of creative, analytical, and practical skills, as well as knowledge, toward a common good, by balancing one's own, others', and higher order interests, over the long as well as the short term, through the infusion of positive ethical values.

The Nature of Intelligence

There are many and diverse definitions of intelligence. Nevertheless, intelligence is usually defined in terms of an individual's ability to adapt to the

This chapter draws on, updates, and expands on Sternberg (2011), my previous Cambridge handbook chapter on successful intelligence.

environment and to learn from experience in life (Sternberg & Detterman, 1986). The definition of intelligence presented in this chapter is more elaborated. It is based on my (Sternberg, 1997, 1998a, 1999c) theory of successful intelligence. Successful intelligence is (1) the ability to formulate, strive for, and, to the extent possible, achieve one's goals in life, given one's sociocultural context, (2) by capitalizing on strengths and correcting or compensating for weaknesses (3) in order to adapt to, shape, and select environments (4) through a combination of analytical, creative, and practical abilities. In recent years, I have emphasized that intelligence best serves individuals and societies when it is augmented by wisdom (Sternberg, 1998a, 2003b, 2008, in press) – the utilization of our abilities and knowledge, through the infusion of positive ethical values, toward a common good, by balancing one's own, others', and higher order interests over the long as well as the short term.

Consider first Item 1. Intelligence involves formulating, striving for, and achieving, to the extent possible, a meaningful and coherent set of goals. Successful intelligence requires one to have the skills and dispositions to reach those goals, if possible. Environmental circumstances sometimes, perhaps often, make the reaching of one's goals impossible and part of successful intelligence is reformulating one's goals, throughout life, in order to keep them realistic with respect to what is possible for one's life. The important question typically is not so much what career or personal goals individuals have chosen but rather whether those goals make sense for the person, what he or she has done to be able to realize those goals in a meaningful way, and reformulating goals as necessary. Thus, this item actually includes four subitems: (1) identifying meaningful goals; (2) coordinating those goals in a meaningful way so that they form a coherent story of what one is seeking in life; (3) moving a substantial distance along the path toward realizing those goals, given one's opportunities and general environmental circumstances; and (4) reformulating goals as one realizes that earlier goals were unrealistic or even ill-chosen.

This first item recognizes that "intelligence" means a somewhat different thing to each individual. The individual who wishes to become a distinguished judge in a courtroom will be taking a different path from the individual who wishes to become a successful musician – but both individuals will have formulated a set of personally meaningful and internally coherent goals to reach. A complete evaluation of intelligence needs to focus not on the particular goals that have been chosen but rather on (1) whether the individual has chosen a personally meaningful and worthwhile set of goals compatible with the knowledge, skills, and personality dispositions he or she possesses that are required to achieve those goals and (2) whether the individual has ascertained feasible ways of achieving those goals.

Item 2 acknowledges that, despite the fact that psychologists sometimes talk of a "general" factor of intelligence (Jensen, 1998; Spearman, 1927; see essays in Sternberg, 2000; Sternberg & Grigorenko, 2002b; Sternberg & Kaufman, 2011), really, virtually no one excels at everything or fails at everything. People who make the world a better place have figured out their strengths and weaknesses and, further, have found ways to work effectively within that pattern of strengths and weaknesses.

There is no single way to succeed in any given career. Consider, as an example, teaching. Educational researchers often try to identify the characteristics of expert teachers (see Sternberg & Horvath, 1995; Sternberg & Williams, 2010). Indeed, the researchers have distinguished some such characteristics. But, in the end, teachers can excel in many and diverse ways. Some teachers excel at presenting large lectures; others excel in small seminars; and still others excel in one-on-one mentoring. There is no single uniform formula that works for each and every teacher. Excellent teachers infer their strengths in teaching and make an effort to arrange their teaching responsibilities, if possible, so that they can optimize on their strengths and, at the same time, either compensate for or fix their weaknesses. Team teaching is one way of capitalizing and compensating, as one teacher can compensate for what the other teacher does not do well. The same options for capitalization and compensation might be available for people in other careers as well.

Item 3 recognizes that intelligence broadly defined refers to more than just "adapting to the environment," which is the mainstay of conventional definitions of intelligence. The theory of successful intelligence distinguishes among adaptation, shaping, and selection of environments in which one lives or may live.

Item 4 recognizes that goals change throughout one's life. For example, one might have hoped, earlier in one's life, to be the world's greatest scientist, artist, or business person. Eventually, one may have to moderate such goals to render them more realistic for one's life. Or one might have hoped to be the father or mother of some number of children and discovered that it just was not possible. The person who approaches the end of life with exactly the same goals as those with which he or she started may have missed something important in life along the way. Many of us hope to be just like our professional mentors. At some point, we discover we will never become them – we only can become us.

When one adapts to the environment, one modifies oneself to fit one or more environments. Historically, adaptability always has been viewed as a key skill in almost any definition of intelligence.

In life, however, adaptation to the environment often is not enough. Adaptation must be balanced with the shaping of the environment. In shaping, one modifies the environment to attain what one seeks from that environment, rather than modifying oneself, as in adaptation, to fit the environment. Truly successful people in any field of endeavor are not just adaptors; they also are shapers of the environment. Such people realize that they cannot change everything in the world but also realize that, if they want to have a meaningful impact on the world, they have to change some things that are not right in the world. An important part of successful intelligence is deciding what needs to be changed and, then, how to change it (Sternberg, 2003a).

Sometimes, one tries unsuccessfully to adapt to a given environment and then also fails in shaping that very same environment. In such instances, one may find that, no matter what one tries to do to make the environment work out, nothing in fact seems to work out. In such cases, the optimal action may be to select a different environment.

Many of the most successful people in any given field are individuals who started off their careers in another field and discovered that their first field of endeavor was not the one in which they truly had the most to contribute to the world. Rather than spend their lives in some pursuit that turned out not to match their individual pattern of strengths and weaknesses, they had the good sense to find something else to do with their lives where they really had a significant contribution to make. They selected a new environment in which to make a difference.

Successful intelligence involves a wider range of human abilities than is typically assessed by tests of intellectual or academic skills. Most of these tests assess largely memory and analytical abilities. As concerns memory, the tests assess primarily the abilities one uses in recalling and recognizing information. As concerns analytical abilities, the tests assess the skills one uses when one analyzes, compares and contrasts, critiques, evaluates, and judges. These skills are important and indeed essential during the years of schooling as well as in later life. But they are not the only skills that are important for success in school and in life. A person needs not only to recall and, when appropriate, to analyze concepts but to be able to generate and apply these concepts. Memory is involved in essentially all analytic, creative, and practical thinking and is necessary for acting on such thinking; but it is far from sufficient for action.

According to the augmented theory of successful intelligence and its development (Sternberg, 1980a, 1984, 1985a, 1985b, 1986, 1988, 1990a, 1997, 1999a, 2003c, 2004, 2015b), a common set of mental processes is involved in all aspects of intelligence. These mental processes are viewed as universal, bridging across cultures (Sternberg, 2004), racial and ethnic groups (Sternberg, Grigorenko, & Kidd, 2005), and other groupings as well. On the one hand, the problem solutions that are viewed as intelligent in one culture may be quite different from the problem solutions viewed as intelligent in a different culture; on the other hand, the need to define what problems are and to develop strategies to solve these problems is to be found in any culture. Even within a given culture, there may exist differences in what various groups of people mean by "intelligence" (Grigorenko et al., 2001; Okagaki & Sternberg, 1993; Sternberg, 1985c).

Metacomponents, or executive processes, are used by a person to plan what to do, to monitor tasks as they are being done, and to evaluate performance on tasks after the tasks are completed. Metacomponents include recognizing the existence of a problem, defining the nature of the problem, deciding on a strategy to solve the problem, monitoring the solution of the given problem, and evaluating the solution after the particular problem is solved.

Performance components execute the instructions furnished by the metacomponents. For instance, inference is executed to decide how two stimuli are related to each and application is used to apply what one has inferred to a new situation (Sternberg, 1977). Other examples of performance components include comparison of stimuli, justification of a given response as adequate although less than perfect, and actually providing the response.

Knowledge-acquisition components are utilized to acquire declarative knowledge or to learn how to solve problems (Sternberg, 1985a). First, selective encoding is used by a learner to decide what information is relevant in the context of one's learning. Second, selective comparison is used by a learner to bring old information to bear on novel problems. Third, selective combination is used by a learner to put together the information that has been selectively encoded and compared into a unified solution to a problem.

Although the same mental processes are used universally for all three aspects of intelligence, these mental processes are applied with different kinds of problems and problem situations as a function of whether a given problem requires analytical thinking, creative thinking, practical thinking, wise thinking, or a combination of these different kinds of thinking. In particular, analytical thinking is needed when information-processing components are applied to relatively familiar kinds of abstract problems somewhat removed from the challenges of everyday life. Creative thinking is needed when the information-processing components are used in the solution of relatively novel kinds of problems and problem situations. Practical thinking is needed when the information-processing components are applied to one's experience in adapting to, shaping, and selecting environments. Wise thinking is needed to utilize one's thinking so as to create a common good. One requires creative skills and attitudes in order to generate ideas, analytical skills and attitudes to decide if the ideas are good ones, and practical skills and attitudes to put one's ideas into practice and to persuade others of the worth of these ideas. Because the theory of successful intelligence comprises three subtheories – a componential subtheory concerning the information-processing components of intelligence; an experiential subtheory concerning the importance of coping with relative novelty and of the automatization of information processing; and a contextual subtheory concerning processes of adaptation, shaping, and selection – the theory has been referred to from time to time as *triarchic* (e.g., Sternberg, 1988).

Intelligence is not merely what intelligence tests test (as claimed by Boring, 1923). Tests of intelligence and of other cognitive and academic skills measure a segment of the range of intellectual skills. They do not measure anywhere close to the whole range. One should not infer that a person who does not test well on these tests is not intelligent. Instead, one should look at scores on tests as one of many possible indicators of a person's intellectual skills. Furthermore, the kinds of skills posited by hierarchical theories (e.g., Carroll, 1993; Cattell, 1971; Vernon, 1971) of intelligence are viewed by the theory of successful intelligence only as a narrow subset of the skills that are important in a broader conception of intelligence.

The Assessment of Successful Intelligence

Our assessments of intelligence have been organized around the analytical, creative, and practical aspects of intelligence, broadly defined. I discuss those assessments here, both singly and collectively.

Analytical Intelligence

Analytical intelligence is called on when the information-processing components of intelligence are applied so as to analyze, judge, evaluate, or compare and contrast. It typically is required when components of information processing are applied to relatively familiar kinds of problems where the judgments to be made are of a fairly abstract nature.

Analytical types of problems, such as analogies or syllogisms, can be analyzed componentially – that is, into their underlying information-processing components (Guyote & Sternberg, 1981; Sternberg, 1977, 1980b, 1983; Sternberg & Gardner, 1983; Sternberg & Turner, 1981). With these problems, response times or error rates can be decomposed to understand the information-processing components that underlie them. The goal of componential research on "intelligence test" types of problems is to uncover the information-processing origins of individual differences in the analytical aspect of human intelligence. Through componential analysis, a researcher could identify sources of individual differences underlying a subtest or factor score such as for "inductive reasoning" or "spatial visualization." For example, response times for the solution of analogies (Sternberg, 1977) and linear syllogisms (Sternberg, 1980a) were decomposed into the elementary performance components that underlay them. The general strategy of such research is to (1) propose an information-processing model of task performance; (2) specify a parameterization of this model, such that each information-processing component is linked to a mathematical parameter corresponding to its latency (and another linked to its error rate); and (3) construct cognitive tasks administered so that it is possible, using mathematical modeling, to isolate the parameters of the statistical (usually regression) model. In this way, it is possible to specify, for the solution of various kinds of problems, several sources of important individual or developmental differences: (1) What performance components are utilized? (2) How much time does it takes to execute each component? (3) How susceptible is each component to error? (4) How are the individual components combined into strategies? (5) What are the underlying mental representations on which the components act?

For example, through componential analysis, it was possible to decompose inductive-reasoning performance into a set of basic underlying information-processing components (Sternberg, 1977). An analogy, *A : B : C : D1, D2, D3, D4*, will be used here as an example to illustrate the components. These components are (1) *encoding*, which is the amount of time needed to register each stimulus (*A, B, C, D1, D2, D3, D4*); (2) *inference*, which is the amount of time needed to figure out the basic relation between given stimuli (*A to B*); (3) *mapping*, which is the amount of time needed to transfer the inferred relation from one set of stimuli to another (needed in analogical reasoning) (*A to C*); (4) *application*, which is the amount of time needed to apply the relation as inferred (and sometimes as mapped) to a new set of stimuli (*A to B to C to ?*); (5) *comparison*, which is the amount of time needed to compare the relative validity of the response options (*D1, D2, D3, D4*); (6) *justification*, which is the amount of time needed to justify one of the answer options as the best of those available (e.g., *D1*); and (7) *preparation-response*, which is the amount of time needed to prepare for problem solution and to respond.

Studies of reasoning do not have to use artificial formats. In one study, David Kalmar and I examined predictions for everyday kinds of situations, such as when milk will spoil (Sternberg & Kalmar, 1997). In this study, we looked at both predictions about the future and postdictions (hypotheses about the past where information about the past is unknown). We found that postdictions took longer to make than did predictions.

Research on the underlying information-processing components of human intelligence yielded some interesting results. Consider a few examples. First, execution of early components (e.g., inference and mapping) in inductive reasoning tends exhaustively to consider the attributes of the stimuli, whereas execution of later components (e.g., application) tends to consider the attributes of the stimuli in self-terminating fashion; only those stimulus attributes are processed that are necessary for reaching a solution (Sternberg, 1977). Second, in a study of the development of figural analogical reasoning in children, we found that, although children generally become quicker in information processing with age, not all components are executed more rapidly with increasing age (Sternberg & Rifkin, 1979). In particular, the encoding component first shows a decrease with age in component time and then an increase. It appears that older children realize that their optimal strategy is to spend more time in encoding the terms of a problem so that they later would be able to devote less time to operating on these encodings. A related, third finding was that superior reasoners tend to spend relatively more time than do poorer reasoners in global, up-front metacomponential (executive) planning, when they solve difficult reasoning or other problems. Poorer reasoners, on the other hand, tend to spend relatively more time in local planning – the planning involved once one has started solving a problem (Sternberg, 1981). It appears that the better reasoners recognize that it is better to invest more of their time up front in problem-solving so as to be able to process information more efficiently later on. Fourth, we found in verbal analogical reasoning that, as children developed, their strategies shifted so that they relied less on word association and more on abstract relations (Sternberg & Nigro, 1980).

Some of the componential studies focused on knowledge-acquisition components rather than on performance components or metacomponents. In one set of studies, for example, my colleagues and I investigated sources of individual differences in people's levels of vocabulary (Sternberg & Powell, 1983; Sternberg, Powell, & Kaye, 1983; see also Sternberg, 1987a, 1987b). We did not view these individual differences merely as ones in declarative knowledge because we especially hoped to understand why some people acquired more of this declarative knowledge and others less. We identified multiple sources of individual and developmental differences. The three main sources of individual differences were in use of knowledge-acquisition components, use of context clues, and use of mediating variables. For instance, in the sentence, "The blen rises in the east and sets in the west," selective comparison, a knowledge-acquisition component, is used to connect prior knowledge about a known concept, the sun, to an unknown word (neologism) in the sentence, "blen." Multiple context cues appear in the sentence: the fact that a blen rises, the fact that it also sets, and the information regarding where it rises and sets. A mediating variable is that the information can occur following (as well as before) the presentation of the unknown word.

My colleagues and I carried out the research described above because we believed that conventional psychometric research sometimes incorrectly attributed individual and developmental differences merely to structural variables, such as "verbal ability." For example, a verbal analogies test might appear, on its surface, to assess verbal reasoning but, in fact, it might measure primarily possession of vocabulary and general information (Sternberg, 1977; Sternberg & Gardner, 1983). In some populations, verbal reasoning might hardly be a source of individual or developmental differences. And, of course, it is important to realize that some children had much more frequent and better opportunities to learn word meanings than did others, depending on the environments in which they were raised.

In the componential-analysis work described earlier in this section, correlations were calculated between latency-component scores of individuals and (percentage-correct) scores on tests of different kinds of psychometric abilities. First, in studies of inductive reasoning (Sternberg, 1977; Sternberg & Gardner, 1982, 1983), we found that although latencies for inference, mapping, application, comparison, and justification tended to correlate with scores on psychometric tests, the highest correlation typically was between the psychometric tests and the latency for the preparation-response component. This result was surprising at first, because the preparation-response component latency was estimated as the regression constant in the predictive regression equation. This result was what originally gave birth to the concept of metacomponents: higher order executive processes used to plan, monitor, and evaluate task performance. We also found, second, that the correlations obtained for all the component latencies demonstrated convergent-discriminant validation: The correlations tended to be statistically significant with psychometric tests of reasoning skills but not with psychometric tests of perceptual-speed skills (Sternberg, 1977; Sternberg & Gardner, 1983). Third, we obtained significant correlations with vocabulary only for the encoding of verbal stimuli (Sternberg, 1977, Sternberg & Gardner, 1983), which makes sense, as encoding is the component whereby vocabulary is recognized. Fourth, we found in studies of linear-syllogistic reasoning (e.g., *Joe is taller than Mike. Mike is taller than Bill. Who is tallest?*) that latency-based components of a proposed (mixed linguistic-spatial) model that were supposed to correlate with verbal ability did so but did not correlate with spatial ability; latency-based components that were supposed to correlate with spatial ability did so but did not correlate with verbal ability. Put another way, we validated the proposed model of linear-syllogistic reasoning not only in terms of the fit of response-time or error data to the predictions of the alternative information-processing models but also in terms of the correlations of latency-based component scores with scores on psychometric tests of verbal and spatial abilities (Sternberg, 1980a). Fifth, and finally, we found individual differences in subjects' strategies for solving linear syllogisms. Some people use a largely linguistic model for solving the problems, others use a largely spatial model, and most use the proposed linguistic-spatial mixed model. Thus, sometimes a less than perfect fit of a proposed model to group data may reflect individual differences in strategies among reasoners.

In later work, discussed below in the sections on the Rainbow and Kaleidoscope Projects (Sternberg, 2009, 2010a, 2016; Sternberg et al., 2012; Sternberg & Coffin, 2010; Sternberg & The Rainbow Project Collaborators, 2006), my colleagues and I studied analytical intelligence using analytical essays as well as multiple-choice items – for example, asking examinees to analyze ideas in a book. We have found, as have most others, that almost all analytical tests tend to correlate highly with each other – they are measures of so-called general intelligence – although essays introduce some variation beyond what is found in typical multiple-choice assessments.

Creative Intelligence

Intelligence tests involve a range of problems, some of them more novel, others less. In some of the componential work my colleagues and I have done, we have shown that when one goes beyond the range of unconventionality of the conventional tests of intelligence, one starts to tap sources of individual differences that are not well measured by the tests. According to the augmented theory of successful intelligence, creative intelligence can be measured by assessing how well an individual can cope with relative novelty.

For example, in an early study, we presented participants with novel kinds of reasoning problems that were convergent – they had a single best answer. For example, the participants were told that some objects are green and others blue; but still other objects might be *grue*, meaning green until the year 2000 and blue thereafter, or bleen, meaning blue until the year 2000 and green thereafter (see Goodman, 1974). Or they might be told of four kinds of people on the planet Kyron – blens, who are born young and die young; kwefs, who are born old and die old; balts, who are born young and die old; and prosses, who are born old and die young (see Sternberg, 1982; Tetewsky & Sternberg, 1986). The task of the participants was to predict future states of the world from past states, given only incomplete information. In another set of studies, participants were presented with more conventional kinds of inductive-reasoning problems, in particular, analogies, series completions, and classifications. However, the inductive-reasoning problems were nonstandard: They had premises preceding them that were either conventional (dancers wear shoes) or novel (dancers eat shoes). The participants had to solve the problems as if the counterfactuals were true (Sternberg & Gastel, 1989a, 1989b).

In these studies, we found that correlations between latencies on these tasks and scores on conventional kinds of psychometric ability tests depended on how novel, or nonentrenched, the conventional tests were. The more novel were the items, the higher were the correlations of our nonstandard kinds of tests with scores on successively more novel conventional tests. Thus, the information-processing components isolated for relatively novel items would tend to correlate more highly with scores on more unusual tests of fluid abilities (e.g., the test of Cattell & Cattell, 1973) than with tests of crystallized abilities. My colleagues and I also discovered that, when response times on the relatively novel problems were componentially analyzed, some latency-based components better measured the creative aspect of intelligence than did others. For example, in the above-mentioned "grue-bleen" task, the information-processing component requiring people to switch from conventional green-blue thinking to grue-bleen thinking and then back to green-blue thinking again was an especially strong measure of the ability to cope with novelty.

In our original work with divergent-reasoning problems with no one best answer, we asked participants to create various kinds of products (Lubart & Sternberg, 1995; Sternberg & Lubart, 1991, 1995, 1996), where an infinite variety of responses was theoretically possible. Individuals were asked to create products in four domains: writing, art, advertising, and science. In writing, they were asked to write very short stories. We gave them a choice of titles, such as

"Beyond the Edge" or "The Octopus's Sneakers." In art, the participants were asked to produce art compositions, with titles including "The Beginning of Time" or "Earth from an Insect's Point of View." In advertising, they were asked to produce advertisements for boring products, such as a brand of bow tie or a new brand of doorknob. In science, the participants were asked to solve problems such as a problem asking them how people might detect extraterrestrial aliens among us who are seeking to escape detection. Participants created two products in each domain.

First, we found that creativity appears to comprise, at least in part, the components proposed by Sternberg and Lubart's (1995) investment model of creativity: intelligence, knowledge, thinking styles (see Zhang & Sternberg, 1998), personality, and motivation, as well as the environment's support of creativity. Second, we found that creativity is largely although not entirely domain-specific. Correlations of ratings of the creative quality of products across domains were lower than correlations of ratings within domains – suggesting strong domain effects – and generally were at about the 0.4 level. Thus, there was some level of relation across domains, at the same time that it was clear that some people were strong in one or more domains but not in others. Third, we found a wide range of correlation coefficients of measures of creative performance with conventional tests of abilities. As had been the case for the correlations obtained with convergent problems, the correlations were higher to the extent that problems on the conventional tests were novel or nonentrenched. In particular, correlations of scores on the creativity measures were higher with scores on fluid ability tests than with scores on crystallized ability tests. In general, correlations were higher, the more novel the fluid ability test was. These results suggest that tests of creative intelligence have some overlap with conventional psychometric tests of intelligence (e.g., in requiring verbal skills or the ability to analyze one's own ideas; Sternberg & Lubart, 1995); but the tests of creative intelligence also tap skills beyond those measured even by relatively novel kinds of items on the conventional tests of intelligence.

Creativity goes beyond creative intelligence (Sternberg, 1999b, 2018a). It largely involves an attitude toward life (Sternberg & Lubart, 1995). In particular, creative people are willing to defy the crowd, to defy themselves and go beyond their past ideas, and to defy the Zeitgeist, not accepting conventional presuppositions just because others do.

Practical Intelligence

Practical intelligence is involved when individuals apply their abilities to the kinds of problems that confront them in daily life, such as in the home or on the job. Practical intelligence involves applying the information-processing components of intelligence to experience in order to (1) adapt to, (2) shape, and (3) select environments. People differ in their relative balance of adaptation, shaping, and selection and in the care and competence with which they balance among the three possible courses of action.

Much of our research on practical intelligence has focused on the concept of tacit knowledge. We have defined this construct as what one needs to know, that one is not explicitly taught and that often is not even verbalized, in order to work effectively in an environment (Sternberg et al., 2000; Sternberg & Hedlund, 2002; Sternberg & Wagner, 1993; Sternberg, Wagner, & Okagaki, 1993; Sternberg, Wagner, Williams, & Horvath, 1995; Wagner, 1987; Wagner & Sternberg, 1986; Williams et al., 2002). We have represented tacit knowledge in the form of production systems, or sequences of "if-then" statements that describe procedures that one follows in various kinds of everyday situations.

We usually have measured tacit knowledge using work-related situations that present problems one might encounter in one's job. We have assessed tacit knowledge for both children and adults. These assessments have been among adults, for people in more than two dozen occupations, including but not limited to management, sales, school and college teaching, school administration, secretarial work, and the military. In a typical tacit knowledge problem, participants are asked to read a story about a problem an individual faces and to rate, for each statement in a larger set of statements, how satisfactory a solution each statement represents. For instance, in a paper-and-pencil measure of tacit knowledge for sales, one of the problems concerns sales of photocopy machines. A relatively inexpensive machine has not been moving out from the showroom and has become overstocked. The participant is asked to rate the quality of each of the various solutions for moving the particular model out from the showroom. In a performance-based measure for sales people, the participant makes a simulated phone call to a supposed customer, who is actually the examiner. The participant then tries to sell advertising space over the phone. The examiner raises various objections to buying the advertising space. The participant is evaluated for the quality, rapidity, and fluency of their responses on the telephone.

In the tacit knowledge studies, we have learned some things about practical intelligence. First, practical intelligence as embodied in tacit knowledge generally increases with experience. But what matters more than sheer amount of experience is the extent to which an individual *profits* from that experience. Some individuals can have been in a job for years and yet have acquired relatively little tacit knowledge. Second, we also discovered that subscores on tests of tacit knowledge – such as for managing oneself, managing others, and managing tasks – correlate significantly and substantially with each other. Third, scores on various tests of tacit knowledge, such as for academics and managers, also correlated substantially (at about the 0.5 level) with each other. Fourth, therefore, tests of tacit knowledge may yield a general factor across these tests. Fifth, however, scores on tacit knowledge tests do not correlate more than modestly with scores on conventional tests of intelligence, whether the measures used are single-score measures or multiple-ability batteries. Thus, any general factor obtained from the tacit knowledge tests is not the same as a general factor from standardized tests of academic abilities (suggesting that neither kind of *g*-factor is truly general but rather general only across a somewhat limited range of measuring instruments). Sixth, despite the lack of correlation between scores on practical-intellectual tests with scores on conventional measures, the scores on tacit knowledge tests predict performance on the job about as well as or even better than do conventional psychometric intelligence tests. Seventh, in one study done at the Center for Creative Leadership in North Carolina, we further found that scores on our tests of tacit knowledge for management were the best single predictor of executive performance on a managerial simulation. In a hierarchical regression, scores on conventional tests of intelligence, personality, styles, and interpersonal orientation were entered into the regression first and then scores on the test of tacit knowledge were entered last. Scores on the test of tacit knowledge provided the single best prediction of the managerial simulation score. Moreover, these scores also contributed significantly to the prediction of performance on the simulation even after everything else was entered first into the equation. Eighth, in work on military leadership (Hedlund et al., 2003; Sternberg & Hedlund, 2002; Sternberg et al., 2000), we found that soldiers' scores on tests of tacit knowledge for military leadership predicted ratings of military leadership effectiveness. In contrast, scores on a conventional standardized test of intelligence and on a tacit knowledge test for managers did not significantly predict the ratings of military leadership effectiveness. In work with Yup'ik Eskimos (Grigorenko et al., 2004), we found that low academic achievers in school nevertheless can have exceptionally high practical adaptive skills at home.

We also have done studies of social intelligence, which, in the theory of successful intelligence, is viewed as a part of practical intelligence. In these studies, we presented participants with photos and asked them to make judgments about the photos. There were two kinds of photos. In the first kind, we asked participants to evaluate whether a male-female couple posing in a photo was a genuine couple (i.e., truly involved in a romantic relationship) or, rather, a phony couple posed by the experimenters just for the photo. In the second kind of photo, we asked participants to indicate which of two individuals was the other's supervisor (Barnes & Sternberg, 1989; Sternberg & Smith, 1985). We found that females were superior to males on these social-intelligence tasks. Scores on the two social-intelligence tasks did not correlate with scores on conventional cognitive-ability tests, nor did they correlate with each other, suggesting a substantial degree of domain specificity in the photo-identification task.

We also have tested the theory overseas. In a study in Usenge, Kenya, near the city of Kisumu, we investigated school-age children's ability to adapt to their indigenous environment. We devised a test of practical intelligence for adaptation to the environment (see Sternberg & Grigorenko, 1997; Sternberg, Nokes, et al., 2001; for more examples of cultural work relevant to the theory, see Sternberg, 2004, 2007, 2014). Our test of practical intelligence assessed children's informal tacit knowledge of natural herbal medicines that the villagers believe can be used to fight various types of parasitic illnesses, such as whipworm, schistosomiasis, or malaria. Most villagers believe in the efficacy of these medicines, as shown by the fact that children in the villages use their knowledge of these medicines an average of once a week, both in medicating themselves and in medicating others. Thus, tests of what these medicines are and of how to use the medicines constitute effective measures of one aspect of practical intelligence, as defined by the villagers as well as by the environmental demands of their life circumstances. Middle-class Westerners might find it quite a daunting challenge to thrive or even to survive in these harsh environmental contexts or, for that matter, in the contexts of urban ghettos that sometimes are not distant from the comfortable homes of the middle class.

We measured the rural Kenyan children's ability to identify the medicines, where they come from, what they are used for, and what appropriate doses are under given circumstances. Based on the theory of successful intelligence and on research we had done elsewhere, we expected that scores on this test would be uncorrelated with scores on conventional tests of intelligence. To test this hypothesis, we also administered to the children a test of fluid or abstract-reasoning–based abilities, Raven's Coloured Progressive Matrices Test, as well as a measure of crystallized or formal-knowledge–based abilities, the Mill Hill Vocabulary Scale. In addition, we administered to the children a comparable test of vocabulary, which was couched in their own Dholuo language. Typically, the Dholuo language is spoken in the children's homes, English in the schools.

As predicted, we found no significant correlation between scores on the test of indigenous tacit knowledge and scores on the fluid-ability tests. But, to our astonishment, we found statistically significant *negative* correlations of the tacit knowledge tests with the tests of crystallized abilities. Put another way, the higher the children scored on the test of tacit knowledge, on average, the lower they scored, on average, on the tests of crystallized abilities. This unexpected result can be interpreted in various ways. Based on the ethnographic observations of the two anthropologists on the team, Geissler and Prince, we concluded that a plausible scenario takes into account the expectations of families and the society in general for their children.

At least at the time we did the study, many of the rural Kenyan children in the village of Usenge dropped out of school before graduating, even from elementary school. They may have dropped out for financial or other reasons, such as that many families in the village do not particularly value formal Western schooling and that, in the ideal, the children would be hired into apprenticeships that ultimately would bring them a steady income. The relatively low value attached to Western schooling stems from the fact that children of many families will, for the most part, spend their lives farming or engaged in other occupations that make little or no use of Western schooling. These families typically emphasize teaching their children the indigenous informal knowledge that will lead to successful adaptation to the environments in which the children really live. Children who spend their time learning the indigenous practical knowledge of the community environment generally do not invest themselves heavily in school achievement, whereas children who do well in school generally do not invest themselves as heavily in learning the indigenous knowledge – hence the negative correlations. Indeed, the more "successful" children drop out of school to take the apprenticeships.

The results of the study in Kenya suggest that, if we identify a general factor of human intelligence, this general factor may tell us more about how abilities interact with patterns of (Westernized) schooling and the environment in general than it does about the structure of human abilities. In Western schooling, children, from an early age, typically study a variety of kinds of subject matter and thus develop skills in a variety of academic areas. This kind of Western schooling well prepares the children to take a test of intelligence, which typically measures skills in a variety of academic areas. Often, intelligence tests measure skills that children taking the tests were expected to acquire some years before taking the intelligence test. But, as Rogoff (1990) and others have shown, this pattern of schooling is far from universal and has not even been common during much of the history of humankind. Throughout history, and in many locations even today, schooling, especially for boys, takes the form of apprenticeships, through which children learn a craft from an early age. The children learn what they will need to know to succeed in a trade but not a lot more. Hence it is less likely that one would observe a large general factor in their scores, much as we discovered in Kenya.

We have considered each of the aspects of successful intelligence separately. How do these elements fare when they are assessed together?

All Three Aspects of Successful Intelligence Together

Internal validity studies. Several factor-analytic investigations support the internal validity (structural soundness) of the theory of successful intelligence.

In one study (Sternberg et al., 1999), my collaborators and I used an early assessment, the so-called Sternberg Triarchic Abilities Test (STAT; Sternberg, 1993), to investigate the internal validity of the theory. High school students, almost all from diverse parts of the United States, took the test, which was composed of twelve subtests in all, with four subtests measuring each of analytical, creative, and practical abilities. For each type of ability, the four subtests comprised three multiple-choice tests and one essay test. The multiple-choice tests each involved, respectively, verbal, quantitative, and figural content for their items. Consider the content of each test, as taken from Sternberg and colleagues (1999):

1. Analytical-Verbal: Figuring out meanings of neologisms (artificial words) from natural contexts. Students see a novel word embedded in a paragraph and have to infer its meaning from the context.
2. Analytical-Quantitative: Number series. Students have to say what number should come next in a series of numbers.
3. Analytical-Figural: Matrices. Students see a figural matrix with the lower right entry missing. They have to say which of the options fits into the missing space.
4. Practical-Verbal: Everyday reasoning. Students are presented with a set of everyday problems in the life of an adolescent and have to select the option that best solves each problem.
5. Practical-Quantitative: Everyday math. Students are presented with scenarios requiring the use of math in everyday life (e.g., buying tickets for a ball game) and have to solve math problems based on the scenarios.
6. Practical-Figural: Route planning. Students are presented with a map of an area (e.g., an entertainment park) and have to answer questions about navigating effectively through the area depicted by the map.
7. Creative-Verbal: Novel analogies. Students are presented with verbal analogies preceded by counterfactual premises (e.g., money falls off trees). They have to solve the analogies as though the counterfactual premises were true.
8. Creative-Quantitative: Novel number operations. Students are presented with rules for novel number operations, for example "flix," which involves numerical

manipulations that differ as a function of whether the first of two operands is greater than, equal to, or less than the second. Participants have to use the novel number operations to solve presented math problems.

9. Creative-Figural: In each item, participants are first presented with a figural series that involves one or more transformations; they then have to apply the rule of the series to a new figure with a different appearance and complete the new series.
10. Analytical-Essay: This essay requires students to analyze the use of security guards in high schools: What are the advantages and disadvantages and how can these be weighed to make a recommendation?
11. Practical-Essay: Give three practical solutions to a problem you are currently having in your life.
12. Creative-Essay: Describe the ideal school.

Confirmatory factor analysis on the data was supportive of what was then called the triarchic theory of human intelligence. The analysis yielded separate and orthogonal analytical, creative, and practical factors. The lack of correlation among factors was due in large part to the inclusion of essay as well as multiple-choice subtests. This was because, although multiple-choice tests tended to correlate moderately to highly with other multiple-choice tests, their correlations with essay tests were much lower. The multiple-choice analytical subtest loaded most highly on the analytical factor but the essay-based creative and practical subtests loaded most highly on their respective (creative and practical) factors. Thus, measurement of creative and practical abilities need ideally to be effected with testing instruments that complement but are different from multiple-choice instruments.

In another study, in the United States, Finland, and Spain, my colleagues and I used the multiple-choice section of the STAT to compare five alternative models of intelligence, again using confirmatory factor analysis. A factorial model featuring a general factor of intelligence fit the data only relatively poorly. The triarchic model, allowing for intercorrelations among the analytic, creative, and practical factors (as predicted by the theory), provided the best fit to the factorial data (Sternberg, Castejón et al., 2001).

In a further study, Grigorenko and I (Grigorenko & Sternberg, 2001) tested Russian schoolchildren (ranging in age from eight to seventeen years) as well as mothers and fathers of these children, as available. We measured analytical, creative, and practical intelligence. Consider, for instance, the tests used for adults; similar tests were used for the children in the studies.

Fluid (analytical) intelligence was assessed by two subtests of a test of nonverbal (figural) intelligence. The test we used, the Test of *g*: Culture Fair, Level II (Cattell & Cattell, 1973), is a test of fluid intelligence that was designed to reduce, to the extent possible, influences of vocabulary, culture, and educational level, although no test successfully eliminates or even substantially reduces such influences.

The test of crystallized intelligence that we used was adapted from existing traditional tests used in Russia measuring analogical reasoning and vocabulary. We used adaptations of Russian tests rather than translations of American tests because the vocabulary used in Russia differs substantially from that used in the United States and other English-speaking countries.

The measure of creative intelligence also contained two parts. The first part asked the participants to describe the world through the eyes of insects (see Sternberg & Lubart, 1995). The second part of the test asked participants to describe who might live, and what might happen, on the (imaginary) planet "Priumliava." No further information on the nature of the planet was specified. Three different scores were assigned: novelty, quality, and sophistication of essay. The first part of the measure of practical intelligence was a self-report instrument and assessed skills in the social domain, skills in the family domain, and skills for dealing with problems that arise suddenly. The second part of the measure of practical intelligence presented four vignettes, based on themes that appeared in popular Russian magazines in the context of discussion of adaptive skills in society that was current at the time. The four themes were, respectively, (1) how to maintain the value of one's savings in the face of inflation, (2) what to do when one makes a purchase and discovers that the item one has purchased is defective, (3) how to locate medical assistance when it is needed, and (4) how to manage a salary bonus one has received for exemplary work. Consensus scoring was used.

Exploratory principal-components analysis for the responses of both children and adults yielded very similar component structures. Both varimax and oblimin rotations yielded clear-cut analytical, creative, and practical components for the tests. Thus, with a sample of a different nationality (Russian), a different set of tests, and a different method of analysis (exploratory rather than confirmatory analysis), we obtained support for the theory of successful intelligence.

We used the analytical, creative, and practical tests to predict mental and physical health among the Russian adults. We measured mental health by widely used paper-and-pencil tests of depression and anxiety and physical health by self-report. The best predictor of both mental and physical health was the practical-intelligence measure. Analytical intelligence came in second and creative intelligence came in third. All three kinds of intelligence contributed incrementally to prediction of mental and physical health, however. Thus, we again concluded that the theory of successful intelligence provides better prediction of success in life than does a narrower conventional theory of intelligence.

External validity studies. We have also examined the external validity of tests assessing successful intelligence.

The Rainbow Project. In a study supported by the College Board (Sternberg & The Rainbow Project Collaborators, 2006), we used an expanded set of tests, including the multiple-choice STAT described in the "Internal validity studies" section, plus three additional measures of creative skills and three of practical skills:

Creative skills. The three additional tests were as follows:

1. *Cartoons*. Participants captioned cartoons.
2. *Written Stories*. Participants were asked to write two short stories, given a choice of titles.
3. *Oral Stories*. Participants orally told a story based on a chosen pictorial collage.

Practical skills. The three additional tests were as follows:

1. *Everyday Situational Judgment Inventory (Movies)*. Participants watched movies of students their age who had problems and had to solve the problems for the students. (The movies cut off before the participants could see how the students themselves solved the problems.)
2. *Common Sense Questionnaire*. Participants were presented with situational judgment tests containing general business-related situations, such as managing tedious tasks or handling a competitive work situation.
3. *College Life Questionnaire*. This was similar to the Common Sense Questionnaire, except that the problems were ones encountered in general college-related situations, such as paying college bills or dealing with a difficult roommate.

We found that our assessments significantly and substantially improved on the validity of the SAT for predicting first-year college grades (Sternberg & The Rainbow Project Collaborators, 2006) (actually doubling prediction). The assessments also improved equity: Using the test to admit a class would have resulted in greater ethnic diversity than would have using just the SAT or just the SAT plus grade point average (GPA).

The Kaleidoscope Project. The Kaleidoscope Project (Sternberg, 2009, 2010a; Sternberg & Coffin, 2010) was made available for a number of years to all 15,000 plus undergraduate applicants to Tufts University. The students were given a selection of essays and projects assessing analytical, creative, practical, and (based on the augmented theory) wisdom-based skills. The applicants were given the opportunity to complete one of the essays and projects. They then could demonstrate analytical, creative, practical, and wisdom-based skills through these essays and projects. Other aspects of the application also were rated for these skills.

The exact Kaleidoscope prompts varied from year to year. The appendix to Sternberg (2010a) contains a comprehensive selection. The essays and projects covered a wide range of talents: writing creative short stories, drawing something creative, designing a science experiment, solving a practical problem, analyzing a favorite book, and speculating on how one might, later in life, be able to make the world a better place.

The questions differed in the skills they emphasized. No question was a "pure" measure of any single component of successful intelligence. Scoring of the exercises was holistic and was done by admissions officers using rubrics. With training, admissions officers could achieve very good interrater reliability (consistency) in their evaluations of the materials.

In the year immediately after Tufts introduced the Kaleidoscope pilot project, applications remained roughly steady or increased slightly and the mean SAT scores of accepted and enrolling students increased to new peaks. Moreover, Tufts did not detect statistically meaningful ethnic group differences on the Kaleidoscope measures. Controlling for the academic rating given to applicants by admissions officers (which combines information from the transcript and standardized tests), students who participated in the Kaleidoscope Project achieved significantly higher GPAs in their undergraduate work than did students who did not participate in Kaleidoscope.

Furthermore, students with higher ratings on Kaleidoscope were more involved in, and reported getting more benefit out of, extracurricular, active-citizenship, and leadership activities in their freshman year at Tufts.

The Panorama Project. The Panorama Project was implemented at Oklahoma State University. The challenge here was to transfer some of the principles of Kaleidoscope so as to apply them in a land-grant institution. In such an institution, the goal is to offer acceptance to as many students as possible, given their reaching the necessary qualifications to attend the institution, rather than to reject large numbers of students. The ideas of Kaleidoscope showed themselves to be at least as applicable at Oklahoma State as at Tufts. Through the Panorama Project, it was possible to accept students into the university whose grades and scores on conventional standardized tests did not fully reflect their capabilities.

The Graduate-Admissions Project. The ideas of the augmented theory of successful intelligence can be applied at the graduate as well as the undergraduate level. My colleague Wendy Williams and I (Sternberg & Williams, 1997) found that the Graduate Record Examination (GRE) is not a particularly meaningful predictor of graduate performance in psychology, at least at Yale University. But we did not propose an alternative measure. We later (Sternberg & Sternberg, 2017; Sternberg, Sternberg, & Todhunter, 2017) proposed an alternative measure, one that can be used in the behavioral and brain sciences. It measures not generalized cognitive skills but rather the research skills one actually needs to succeed in graduate school in psychological sciences and related disciplines – generating hypotheses, generating experiments, and evaluating experiments. We found that scores on our measures correlated well with each other but did not correlate much with scores on measures of fluid intelligence or with scores on the SAT. Some of the correlations with the SAT, at least among Cornell undergraduates, even were negative. In a follow-up, we found that scores on a further measure of evaluating teaching correlated with scores on our assessments but also did not correlate with the inductive-reasoning or SAT measures. These data suggest that measures of creative, analytical, and practical skills measured in a specific domain are distinct from and may be more relevant to future and especially graduate success than generalized analytical measures.

Instruction for Successful Intelligence

Instructional studies have provided an additional means of testing the theory of successful intelligence (Sternberg, Grigorenko, & Zhang, 2008; Sternberg, Jarvin, & Grigorenko, 2009). My colleagues and I have used instruction both in cognitive skills, in general, and in academic skills, in particular.

Cognitive Skills

The kinds of analytical, creative, practical, and wisdom-based skills discussed in this chapter are not fixed but rather modifiable (Sternberg & Pretz, 2005).

Analytical skills are teachable. In one study, for example, I (Sternberg, 1987a) explored whether it is possible to teach people to improve their skills in figuring out the meanings of unknown words presented in natural verbal contexts. I found that theory-based instruction was better than no instruction at all or just practice without formal instruction in improving people's decontextualization skills.

Creative thinking skills also can be taught, at least to some degree. My colleagues and I have devised various programs for teaching them (Sternberg & Williams, 1996; see also Sternberg & Grigorenko, 2007; Sternberg, Jarvin, & Grigorenko, 2009). In one study, we found that children taught how to solve insight problems using knowledge-acquisition components gained more from a pretest to a posttest than did students who were not taught in a theory-based way how to solve the problems (Davidson & Sternberg, 1984).

Practical intelligence skills also are teachable. My colleagues and I developed a program for teaching practical intellectual skills, targeted at middle school students, that explicitly teaches students "practical intelligence for school" in the contexts of doing homework, taking tests, reading, and writing (Gardner et al., 1994; Williams et al., 1996, 2002). We evaluated the program in a variety of settings (Gardner et al., 1994; Sternberg, Okagaki, & Jackson, 1990). We found that students taught using the program academically outperformed students in control groups that did not receive the instruction.

Individuals' use of practical intelligence can be for their own gain and benefit in addition to or instead of the gain and benefit of others. People can exercise practical intelligence for themselves at the expense of others. It is for this reason that wisdom needs to be studied (Baltes & Staudinger, 2000; Sternberg, 1990b, 1998b) and be part of an augmented theory of successful intelligence.

I view intelligence as a form of developing expertise (Sternberg, 1998a, 1999a, 2003a). On this view, tests of intelligence are essentially tests of achievement or of developing expertise (see Ericsson, 1996; Howe, Davidson, & Sloboda, 1998), not merely of intelligence. There is no clear-cut distinction between the two constructs of intelligence and achievement (Sternberg, 1998a, 1999a). Indeed, all measures of intelligence, at some level, assess a form of developing expertise. In some instances, expertise can actually undermine creative thinking (Frensch & Sternberg, 1989).

An example of how tests of intelligence assess developing expertise comes out of work we did in Tanzania. A study we did in Tanzania (see Sternberg & Grigorenko, 1997; Sternberg et al., 2002) points out the risks of giving tests, scoring them, and interpreting the results as static measures of some latent, unmodifiable intellectual ability or abilities. We administered to school children between the ages of eleven and thirteen years near Bagamoyo, Tanzania, assessments of three kinds. They included a form-board classification test, a linear syllogisms test, and a Twenty Questions Test. These kinds of assessments measure the sorts of skills required on conventional tests of intelligence. We administered the tests dynamically rather than statically so that we could assess modifiability (Brown & Ferrara, 1985; Budoff, 1968; Day et al., 1997; Feuerstein, 1979; Grigorenko & Sternberg, 1998; Guthke, 1993; Haywood & Tzuriel, 1992; Lidz, 1987, 1991; Sternberg & Grigorenko, 2002a; Tzuriel, 1995; Vygotsky, 1978). In dynamic testing, children are provided with some kind of feedback to help them improve their skills and their test scores. Vygotsky (1978) suggested that the children's ability to profit from the guided instruction they received during a testing session could serve as a measure of what he called the children's zone of proximal development (ZPD), or the difference between the children's developed abilities and their latent capacities. In other words, Vygotsky considered testing and instruction as being of one integrated piece rather than as being distinct processes.

This integration of teaching and testing makes sense in terms of conceptions of intelligence as involving the ability to learn ("Intelligence and Its Measurement," 1921; Sternberg & Detterman, 1986). A dynamic test directly assesses processes of learning in the context of testing rather than assessing these processes indirectly as the product of past learning. Such assessment is especially important when some children have lacked equal opportunities relative to other children to learn in the past.

In our assessments, children were first given our ability tests. In an experimental group, the children then were provided with a brief period of instruction in which they were able to learn skills that potentially would enable them to improve their scores. In a control group, the children were not provided with this intervention. Then the children were tested again. Because the instruction for each test was brief, one would not expect large gains in test scores. Nevertheless, on average, the gains were both statistically significant in the experimental group and statistically greater than the gains in the control group. In the control group, pretest and posttest scores correlated at the 0.8 level, meaning that the pretest and posttest measured pretty much the same skills. In the experimental group, however, scores on the pretest showed only weak although significant correlations (at the 0.3 level) with scores on the posttest. These correlations suggest that, when tests are administered in static fashion to children living in developing countries, the test scores may be relatively unstable and easily subject to influences of training. The reason might be that the children are not used to taking Western-style tests; as a result, they profit quickly even from relatively small quantities of instruction as to what is expected from them when they take tests. The more important question, however, is not whether the children's test scores changed or even correlated with each other but rather how the scores correlated with scores on other cognitive measures. Put another way, which test was a better predictor of transfer to other cognitive performances, the pretest score or the posttest score? In fact, the posttest score was the better predictor.

Academic Skills

We carried out several sets of studies investigating instruction for academic skills. Five sets are briefly described here.

In a first set of studies, my colleagues and I addressed the question of whether conventional education in school systematically disadvantages children with creative and practical strengths (Sternberg & Clinkenbeard, 1995; Sternberg et al., 1996; Sternberg et al., 1999). Behind this work was the view that the educational systems in most schools clearly tend to favor children with particular strengths in memory and analytical abilities.

We used the STAT, measuring analytical, creative, and practical abilities, in some of our instructional work. The test was administered to children throughout the United States and in some other countries who were identified by their schools, by any standard whatsoever, as gifted. Children were chosen for a summer program in (college-level) psychology if they fell into one of five ability groupings: high analytical, high creative, high practical, high balanced (high in all three abilities), or low balanced (low in all three abilities). Students who were chosen and then came to Yale, where the summer program was held, were divided into four instructional groups. Students in all four instructional groups used the identical introductory psychology textbook (a preliminary version of Sternberg, 1995) and listened to the same psychology lectures, which lasted throughout each morning of instruction. What differed among the groups was the type of afternoon discussion section to which the individuals in the groups were assigned. They were assigned to one of four instructional conditions. The conditions emphasized memory, analytical, creative, or practical instruction. For example, in the memory condition, the students might be asked to describe the main ideas behind a major theory of depression. In the analytical condition, the students might be asked to compare and contrast two theories of depression. In the creative condition, the students might be asked to formulate their own original theory of depression. In the practical condition, the students might be asked how they could use what they had learned in the course about depression to help a friend of theirs who was depressed.

Students in all four instructional conditions were evaluated in various ways, specifically in terms of their performance on homework, a midterm exam, a final exam, and an independent project. Each of these types of work was evaluated for quality with respect to memory, analytical, creative, and practical performance. Thus, all students were evaluated in the same way.

Our results suggested the usefulness of the theory of successful intelligence. This usefulness showed itself in at least two ways.

First, we discovered that all three ability tests – analytical, creative, and practical – significantly predicted success in course performance. Using multiple-regression analysis, we found that at least two of these ability measures contributed significantly to the prediction of each of the measures of achievement. One of the significant predictors was always the analytical score, suggesting that, even with our new teaching methods, analytical skills were still important.

Second, we found an aptitude-treatment interaction whereby students who were placed in instructional conditions that better matched their pattern of abilities outperformed students who were mismatched for their pattern of abilities. Put in other words, when students are taught in a way that fits how they think, they perform better in school. Unfortunately, children who excel in creative and practical abilities are almost never taught or assessed in a way that matches their pattern of abilities. They thus may be at a disadvantage in course after course, year after year.

We performed a follow-up study (Sternberg, Torff, & Grigorenko, 1998a, 1998b) examining learning of social studies and science by third graders and eighth graders. The third graders were students in a very low-income neighborhood in North Carolina. The eighth graders were students who were largely middle to upper middle class studying in Baltimore, Maryland, and Fresno, California. In this study, students were each placed in one of three instructional conditions. In the first instructional condition, they were taught the course that basically they would have been taught had there been no instructional intervention. The emphasis in the course was on memorization of material. In a second condition, students were taught in a way that stressed analytical thinking. In the third condition, students were taught in a way that stressed analytical, creative, and practical thinking in equal measure. Performance of all students was assessed for memory learning (through multiple-choice assessments) as well as for analytical, creative, and practical learning (through performance assessments).

As expected, students in the successful-intelligence (analytical, creative, practical) condition outperformed the students in the other conditions in terms of the performance assessments. Students in the successful-intelligence condition outperformed the other students even on the multiple-choice memory-based tests.

We extended these ideas to reading curricula at the middle school and the high school level. In a study of middle school students and high school students, we taught reading either analytically, creatively, and practically or through the regular curriculum. At the middle school level, reading was explicitly taught. At the high school level, reading was infused into instruction in other subject matter – mathematics, physical sciences, social sciences, English, history, foreign languages, and the arts. In all of the settings we studied, students who were taught for successful intelligence substantially outperformed students who were taught in standard ways (Grigorenko, Jarvin, & Sternberg, 2002).

A much larger-scale study, described in Sternberg, Grigorenko, and Zhang (2007) and later in Sternberg and colleagues (2014), was conducted with a very large sample. The study covered 4 years, 9 states, 14 school districts, and 110 schools. Initially, the results looked promising but later proved to be disappointing. The results showed no significant differences among methods. It proved to be extremely difficult to maintain control of teacher training and implementation of curricula. We lacked the personnel carefully to monitor teaching practices in every experimental site. As a consequence, some teachers in the analytical and successful-intelligence conditions reverted, over time, to their more usual and comfortable memory-based methods of teaching. The result was that, in the end, the instruction was not nearly as different and targeted across conditions as we had hoped and expected.

Conclusions

This chapter has presented the augmented theory of successful intelligence. Some believe the theory departs too much from the conventional theory of general intelligence proposed by Spearman (1904): Some disagree with aspects of the theory (e.g., Brody, 2003a, 2003b) and others disagree with the whole thing, vehemently (Gottfredson, 2003a, 2003b). Others believe the theory does not depart sufficiently from conventional *g* theory (Gardner, 2006, 2011). Still other psychologists have proposed theories that are more compatible, at least in spirit, with the theory proposed here (Ceci, 1996). Certainly the theory is not wholly correct – scientific theories so far have not been. But I would hope, at the same time, it might provide a broader basis for future theories than Spearman's theory of general intelligence.

The educational system in the United States, as in many other countries around the world, favors instruction and assessments that tap into two important cognitive skills: memory and analysis. Students who are adept at these two skills are a good fit for the educational system because the ability tests, instruction, and achievement tests based on them all largely assess products and sometimes processes deriving from these two kinds of skill. The problem, however, is that children whose strengths are in other kinds of skill may be shortchanged and even severely penalized by this system. These children might learn and test better if only they were given a chance to show their strengths, not just their weaknesses.

A society can create a closed system that provides advantages to only certain types of students and that disadvantages the rest. Students who excel in memory and analytical skills may perform well on ability tests and achievement tests and therefore find the doors of opportunity opening to them. Students who excel in other abilities may do poorly on the tests and find the doors shutting tight. By treating students with alternative patterns of abilities as lacking ability to succeed, we may end up creating harmful self-fulfilling prophecies.

During the twentieth century, IQs rose around the world by roughly thirty points – two standard deviations (Flynn, 2016). That is an incredible increase and the rise is continuing in the United States (whereas, in other countries, it has ceased or even reversed). But when one looks at how people function, does their functioning reflect a two standard-deviation rise in intelligence? Have we made substantial progress, since the beginning of the twentieth century, in combating global warming, pollution, income disparities, poverty, hunger, violence, or any of the other challenges that today face the world? When one looks at the 2016 presidential election in the United States, one hardly sees signs of wondrous intelligence. If we do not start emphasizing wisdom and ethics more in our schooling (Sternberg & Hagen, 2018, 2019), what hope is there for any of our societies, or for the world (Sternberg, 2015a)?

References

Baltes, P. B., & Staudinger, U. M. (2000). Wisdom: A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. *American Psychologist*, *55*, 122–135.

Barnes, M. L., & Sternberg, R. J. (1989). Social intelligence and decoding of nonverbal cues. *Intelligence*, *13*, 263–287.

Boring, E. G. (1923). Intelligence as the tests test it. *New Republic*, June 6, 35–37.

Brody, N. (2003a). What Sternberg should have concluded. *Intelligence*, *31*(4), 339–342.

Brody, N. (2003b). Construct validation of the Sternberg Triarchic abilities test: Comment and reanalysis. *Intelligence*, *31*(4), 319–329.

Brown, A. L., & Ferrara, R. A. (1985). Diagnosing zones of proximal development. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives* (pp. 273–305). New York: Cambridge University Press.

Budoff, M. (1968). Learning potential as a supplementary assessment procedure. In J. Hellmuth (Ed.), *Learning disorders* (Vol. 3, pp. 295–343). Seattle, WA: Special Child.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.

Cattell, R. B., & Cattell, H. E. P. (1973). *Measuring intelligence with the Culture Fair Tests*. Champaign, IL: Institute for Personality and Ability Testing.

Ceci, S. J. (1996). *On intelligence* (rev. ed.). Cambridge, MA: Harvard University Press.

Davidson, J. E., & Sternberg, R. J. (1984). The role of insight in intellectual giftedness. *Gifted Child Quarterly*, *28*, 58–64.

Day, J. D., Engelhardt, J. L., Maxwell, S. E., & Bolig, E. E. (1997). Comparison of static and dynamic assessment procedures and their relation to independent performance. *Journal of Educational Psychology*, *89*(2), 358–368.

Ericsson, K. A. (Ed.) (1996). *The road to excellence*. Mahwah, NJ: Lawrence Erlbaum Associates.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device theory, instruments, and techniques*. Baltimore, MD: University Park Press.

Flynn J. R. (2016). *Does your family make you smarter? Nature, nurture, and human autonomy*. New York: Cambridge University Press.

Frensch, P. A., & Sternberg, R. J. (1989). Expertise and intelligent thinking: When is it worse to know better? In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 157–188). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gardner, H. (2006). Multiple intelligences: New horizons in theory and practice. New York: Basic Books.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gardner, H., Krechevsky, M., Sternberg, R. J., & Okagaki, L. (1994). Intelligence in context: Enhancing students' practical intelligence for school. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp.105–-127). Cambridge, MA: MIT Press.

Goodman, N. (1974). *Fact, fiction, and forecast*. Indianapolis: Bobbs-Merrill.

Gottfredson, L. S. (2003a). Discussion: On Sternberg's "Reply to Gottfredson." *Intelligence*, *31*(4), 415–424.

Gottfredson, L. S. (2003b). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence*, *31*(4), 343–397.

Grigorenko, E. L., Geissler, P. W., Prince, R., Okatcha, F., Nokes, C., Kenny, D. A., Bundy, D. A., & Sternberg, R. J. (2001). The organization of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavioral Development*, *25*(4), 367–378.

Grigorenko, E. L., Jarvin, L., & Sternberg, R. J. (2002). School-based tests of the triarchic theory of intelligence: Three settings, three samples, three syllabi. *Contemporary Educational Psychology*, *27*, 167–208.

Grigorenko, E. L., Meier, E., Lipka, J., Mohatt, G., Yanez, E., & Sternberg, R. J. (2004). Academic and practical intelligence: A case study of the Yup'ik in Alaska. *Learning and Individual Differences*, *14*, 183–207.

Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, *124*, 75–111.

Grigorenko, E. L., & Sternberg, R. J. (2001). Analytical, creative, and practical intelligence as predictors of self-reported adaptive functioning: A case study in Russia. *Intelligence*, *29*, 57–73.

Guthke, J. (1993). Current trends in theories and assessment of intelligence. In J. H. M. Hamers, K. Sijtsma, & A. J. J. M. Ruijssenaars (Eds.), *Learning potential assessment* (pp. 13–20). Amsterdam: Swets & Zeitlinger.

Guyote, M. J., & Sternberg, R. J. (1981). A transitive-chain theory of syllogistic reasoning. *Cognitive Psychology*, *13*, 461–525.

Haywood, H. C., & Tzuriel, D. (1992). Epilogue: The status and future of interactive assessment. In H. C. Haywood & D. Tzuriel (Eds.), *Interactive assessment* (pp. 38–63). New York: Springer-Verlag.

Hedlund, J., Forsythe, G. B., Horvath, J. A., Williams, W. M., Snook, S., & Sternberg, R. J. (2003). Identifying and assessing tacit knowledge: Understanding the practical intelligence of military leaders. *Leadership Quarterly*, *14*, 117–140.

Howe, M. J., Davidson, J. W., & Sloboda, J. A (1998). Innate talents: Reality or myth? *Behavioral & Brain Sciences*, *21*, 399–442.

"Intelligence and its measurement": A symposium (1921). *Journal of Educational Psychology*, *12*, 123–147, 195–216, 271–275.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger/Greenwood.

Lidz, C. S. (Ed.). (1987). *Dynamic assessment*. New York: Guilford Press.

Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. New York: Guilford Press.

Lubart, T. I., & Sternberg, R. J. (1995). An investment approach to creativity: Theory and data. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach*. Cambridge, MA: MIT Press.

Okagaki, L., & Sternberg, R. J. (1993). Parental beliefs and children's school performance. *Child Development*, *64*(1), 36–56.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* New York: Oxford University Press.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*(2), 201–293.

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1980a). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General*, *109*, 119–159.

Sternberg, R. J. (1980b). Sketch of a componential subtheory of human intelligence. *Behavioral and Brain Sciences*, *3*, 573–584.

Sternberg, R. J. (1981). Intelligence and nonentrenchment. *Journal of Educational Psychology*, *73*, 1–16.

Sternberg, R. J. (1982). Natural, unnatural, and supernatural concepts. *Cognitive Psychology*, *14*, 451–488.

Sternberg, R. J. (1983). Components of human intelligence. *Cognition*, *15*, 1–48.

Sternberg, R. J. (1984). Toward a triarchic theory of human intelligence. *Behavioral and Brain Sciences*, *7*, 269–287.

Sternberg, R. J. (1985a). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (Ed.). (1985b). *Human abilities: An information-processing approach*. San Francisco: Freeman.
Sternberg, R. J. (1985c). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, *49*(3), 607–627.
Sternberg, R. J. (1986). Inside intelligence. *American Scientist*, *74*, 137–143.
Sternberg, R. J. (1987a). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89–105). Hillsdale, NJ: Lawrence Erlbaum Associates.
Sternberg, R. J. (1987b). The psychology of verbal comprehension. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 3, pp. 97–151). Hillsdale, NJ: Lawrence Erlbaum Associates.
Sternberg, R. J. (Ed.). (1988). *The triarchic mind*. New York: Viking.
Sternberg, R. J. (1990a). *Metaphors of mind*. New York: Cambridge University Press.
Sternberg, R. J. (1990b). Understanding wisdom. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 3–9). New York: Cambridge University Press.
Sternberg, R. J. (1993). *Sternberg Triarchic Abilities Test*. Unpublished test.
Sternberg, R. J. (1995). *In search of the human mind*. Orlando, FL: Harcourt Brace College.
Sternberg, R. J. (1997). What does it mean to be smart? *Educational Leadership*, *54*(6), 20–24.
Sternberg, R. J. (1998a). Abilities are forms of developing expertise. *Educational Researcher*, *27*, 11–20.
Sternberg, R. J. (1998b). A balance theory of wisdom. *Review of General Psychology*, *2*, 347–365.
Sternberg, R. J. (1999a). Intelligence as developing expertise. *Contemporary Educational Psychology*, *24*, 359–375.
Sternberg, R. J. (1999b). A propulsion model of types of creative contributions. *Review of General Psychology*, *3*, 83–100.
Sternberg, R. J. (1999c). Successful intelligence: Finding a balance. *Trends in Cognitive Sciences*, *3*, 436–442.
Sternberg, R. J. (Ed.). (2000). *Handbook of intelligence*. New York: Cambridge University Press.
Sternberg, R. J. (2001). Why schools should teach for wisdom: The balance theory of wisdom in educational settings. *Educational Psychologist*, *36*(4), 227–245.
Sternberg, R. J. (Ed.). (2003a). *The anatomy of impact: What has made the great works of psychology great?* (pp. 223–228). Washington, DC: American Psychological Association.
Sternberg, R. J. (Ed.). (2003b). Psychologists defying the crowd: Stories of those who battled the establishment and won. Washington, DC: American Psychological Association.
Sternberg, R. J. (2003c). *WICS: A theory of wisdom, intelligence, and creativity, synthesized*. New York: Cambridge University Press.
Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist*, *59*(5), 325–338.
Sternberg, R. J. (2005). The theory of successful intelligence. *Interamerican Journal of Psychology*, *39*(2), 189–202.
Sternberg, R. J. (2007). Culture, instruction, and assessment. *Comparative Education*, *43*(1), 5–22.
Sternberg, R. J. (2008). Schools should nurture wisdom. In B. Z. Presseisen (Ed.), *Teaching for intelligence* (2nd ed., pp. 61–88). Thousand Oaks, CA: Corwin.
Sternberg, R. J. (2009). The Rainbow and Kaleidoscope projects: A new psychological approach to undergraduate admissions. *European Psychologist*, *14*, 279–287.

Sternberg, R. J. (2010a). *College admissions for the 21st century*. Cambridge, MA: Harvard University Press.

Sternberg, R. J. (2010b). WICS: A new model for cognitive education. *Journal of Cognitive Education and Psychology*, *9*, 34–46.

Sternberg, R. J. (2011). The theory of successful intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 504–527). New York: Cambridge University Press.

Sternberg, R. J. (2014). The development of adaptive competence. *Developmental Review*, *34*, 208–224.

Sternberg, R. J. (2015a). Archaic testing and teaching in the United States: Why do they persist? In M. J. Feuer, A. I. Berman, & R. C. Atkinson (Eds.), *Past as prologue: The National Academy of Education at 50* (pp. 265–269). Washington, DC: National Academy of Education.

Sternberg, R. J. (2015b). Successful intelligence: A new model for testing intelligence beyond IQ tests. *European Journal of Education and Psychology*, *8*, 76–84. https://doi.org/10.1016/j.ejeps.2015.09.004.

Sternberg, R. J. (2016). *What universities can be: A new model for preparing students for active concerned citizenship and ethical leadership*. Ithaca, NY: Cornell University Press.

Sternberg, R. J. (2018a). A triangular theory of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *12*, 50–67.

Sternberg, R. J. (2018b). The triarchic theory of successful intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed.). New York: Guilford Press.

Sternberg, R. J. (in press). *Adaptive intelligence*. New York: Cambridge University Press.

Sternberg, R. J., Bonney, C. R., Gabora, L., & Merrifield, M. (2012). WICS: A model for college and university admissions. *Educational Psychologist*, *47*(1), 30–41.

Sternberg, R. J., Castejón, J. L., Prieto, M. D., Hautamäki, J., & Grigorenko, E. L. (2001). Confirmatory factor analysis of the Sternberg triarchic abilities test in three international samples: An empirical test of the triarchic theory of intelligence. *European Journal of Psychological Assessment*, *17*(1), 1–16.

Sternberg, R. J., & Clinkenbeard, P. R. (1995). The triarchic model applied to identifying, teaching, and assessing gifted children. *Roeper Review*, *17*(4), 255–260.

Sternberg, R. J., & Coffin L. A. (2010). Admitting and developing "new leaders for a changing world." *New England Journal of Higher Education*, Winter, *24*, 12–13.

Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence?* Norwood, NJ: Ablex.

Sternberg, R. J., Ferrari, M., Clinkenbeard, P. R., & Grigorenko, E. L. (1996). Identification, instruction, and assessment of gifted children: A construct validation of a triarchic model. *Gifted Child Quarterly*, *40*(3), 129–137.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J., Snook, S., Williams, W. M., Wagner, R. K., & Grigorenko, E. L. (2000). *Practical intelligence in everyday life*. New York: Cambridge University Press.

Sternberg, R. J., & Gardner, M. K. (1982). A componential interpretation of the general factor in human intelligence. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 231–254). Berlin: Springer-Verlag.

Sternberg, R. J., & Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, *112*, 80–116.

Sternberg, R. J., & Gastel, J. (1989a). Coping with novelty in human intelligence: An empirical investigation. *Intelligence*, *13*, 187–197.

Sternberg, R. J., & Gastel, J. (1989b). If dancers ate their shoes: Inductive reasoning with factual and counterfactual premises. *Memory and Cognition*, *17*, 1–10.

Sternberg, R. J., & Grigorenko, E. L. (1997). The cognitive costs of physical and mental ill health: Applying the psychology of the developed world to the problems of the developing world. *Eye on Psi Chi*, *2*(1), 20–27.

Sternberg, R. J., & Grigorenko, E. L. (2007). *Teaching for successful intelligence* (2nd ed.). Thousand Oaks, CA: Corwin.

Sternberg, R. J., & Grigorenko, E. L. (2002a). *Dynamic testing*. New York: Cambridge University Press.

Sternberg, R. J., & Grigorenko E. L. (Eds.). (2002b). *The general factor of intelligence: How general is it?* Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J., Grigorenko, E. L., Ferrari, M., & Clinkenbeard, P. (1999). A triarchic analysis of an aptitude-treatment interaction. *European Journal of Psychological Assessment*, *15*(1), 1–11.

Sternberg, R. J., Grigorenko, E. L., & Kidd, K. K. (2005). Intelligence, race, and genetics. *American Psychologist*, *60*(1), 46–59.

Sternberg, R. J., Grigorenko, E. L., & Zhang, L.-F. (2008). Styles of learning and thinking matter in instruction and assessment. *Perspectives on Psychological Science*, *3*(6), 486–506.

Sternberg, R. J., Grigorenko, E. L., Ngrosho, D., Tantufuye, E., Mbise, A., Nokes, C., Jukes, M., & Bundy, D. A. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence*, *30*, 141–162.

Sternberg, R. J., & Hagen, E. S. (2018). Wisdom. In M. Bornstein (Ed.), *Sage encyclopedia of lifespan human development*. Thousand Oaks, CA: Sage.

Sternberg, R. J., & Hagen, E. S. (2019). Teaching for wisdom. In R. J. Sternberg & J. Glueck (Eds.), *Cambridge handbook of wisdom* (pp. 372–406). New York: Cambridge University Press.

Sternberg, R. J., & Hedlund, J. (2002). Practical intelligence, *g*, and work psychology. *Human Performance*, *15*(1/2), 143–160.

Sternberg, R. J., & Horvath, J. A. (1995). A prototype view of expert teaching. *Educational Researcher*, *24*(6), 9–17.

Sternberg, R. J., Jarvin, L., Birney, D., Naples, A., Stemler, S., Newman, T., Otterbach, R., Randi, J., & Grigorenko, E. L. (2014). Testing the theory of successful intelligence in teaching grade 4 language arts, mathematics, and science. *Journal of Educational Psychology*, *106*, 881–899.

Sternberg, R. J., Jarvin, L., & Grigorenko, E. L. (2009). *Teaching for wisdom, intelligence, creativity, and success*. Thousand Oaks, CA: Corwin.

Sternberg, R. J., & Kalmar D.A. (1997). When will the milk spoil? Everyday induction in human intelligence. *Intelligence*, *25*(3), 185–203.

Sternberg, R. J., & Kaufman, S. B. (Eds.) (2011). *Cambridge handbook of intelligence*. New York: Cambridge University Press.

Sternberg, R. J., & Lubart, T. I. (1991). An investment theory of creativity and its development. *Human Development*, *34*(1), 1–31.

Sternberg, R. J., & Lubart, T. I. (1995). Defying the crowd: Cultivating creativity in a culture of conformity. New York: Free Press.

Sternberg, R. J., & Lubart, T. I. (1996). Investing in creativity. *American Psychologist*, *51*(7), 677–688.

Sternberg, R. J., & Nigro, G. (1980). Developmental patterns in the solution of verbal analogies. *Child Development*, *51*, 27–38.

Sternberg, R. J., Nokes, K., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., & Grigorenko, E. L. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence*, *29*, 401–418.

Sternberg, R. J., Okagaki, L., & Jackson, A. (1990). Practical intelligence for success in school. *Educational Leadership*, *48*, 35–39.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, *38*, 878–893.

Sternberg, R. J., Powell, J. S., & Kaye, D. B. (1983). Teaching vocabulary-building skills: A contextual approach. In A. C. Wilkinson (Ed.), *Classroom computers and cognitive science* (pp. 121–143). New York: Academic Press.

Sternberg, R. J., & Pretz, J. E. (Eds.). (2005). *Cognition and intelligence: Identifying the mechanisms of the mind*. New York: Cambridge University Press.

Sternberg, R. J., & The Rainbow Project Collaborators. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical and creative skills. *Intelligence*, *34*(4), 321–350.

Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, *27*, 195–232.

Sternberg, R. J., & Smith, C. (1985). Social intelligence and decoding skills in nonverbal communication. *Social Cognition*, *2*, 168–192.

Sternberg, R. J., & Sternberg, K. (2017). Measuring scientific reasoning for graduate admissions in psychology and related disciplines. *Journal of Intelligence*, www.mdpi.com/2079–3200/5/3/29/pdf.

Sternberg, R. J., Sternberg, K., & Todhunter, R. J. E. (2017). Measuring reasoning about teaching for graduate admissions in psychology and related disciplines. *Journal of Intelligence*, www.mdpi.com/2079–3200/5/4/34/pdf.

Sternberg, R. J., Torff, B., & Grigorenko, E. L. (1998a). Teaching for successful intelligence raises school achievement. *Phi Delta Kappan*, *79*, 667–669.

Sternberg, R. J., Torff, B., & Grigorenko, E. L. (1998b). Teaching triarchically improves school achievement. *Journal of Educational Psychology*, *90*, 374–384.

Sternberg, R. J., & Turner, M. E. (1981). Components of syllogistic reasoning. *Acta Psychologica*, *47*, 245–265.

Sternberg, R. J., & Wagner, R. K. (1993). The *g*–ocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science*, *2*(1), 1–4.

Sternberg, R. J., Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: The nature and role of tacit knowledge in work and at school. In H. Reese & J. Puckett (Eds.), *Advances in lifespan development* (pp. 205–227). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. *American Psychologist*, *50*(11), 912–927.

Sternberg, R. J., & Williams, W. M. (1996). *How to develop student creativity*. Alexandria, VA: Association for Supervision and Curriculum Development.

Sternberg, R. J., & Williams, W. M. (1997). Does the graduate record examination predict meaningful success in the graduate training of psychologists? A case study. *American Psychologist*, *52*, 630–641.

Sternberg, R. J., & Williams, W. M. (2010). *Educational psychology* (2nd ed.). Boston: Allyn & Bacon.

Tetewsky, S. J., & Sternberg, R. J. (1986). Conceptual and lexical determinants of nonentrenched thinking. *Journal of Memory and Language*, *25*, 202–225.

Tzuriel, D. (1995). Dynamic-interactive assessment: The legacy of L. S. Vygotsky and current developments. Unpublished manuscript.

Vernon, P. E. (1971). *The structure of human abilities*. London: Methuen.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Wagner, R. K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality & Social Psychology*, *52*(6), 1236–1247.

Wagner, R. K., & Sternberg, R. J. (1986). Tacit knowledge and intelligence in the everyday world. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 51–83). New York: Cambridge University Press.

Williams, W. M., Blythe, T., White, N., Li, J., Sternberg, R. J., & Gardner, H. I. (1996). *Practical intelligence for school: A handbook for teachers of grades 5–8*. New York: HarperCollins.

Williams, W. M., Blythe, T., White, N., Li, J., Gardner, H., & Sternberg, R. J. (2002). Practical intelligence for school: Developing metacognitive sources of achievement in adolescence. *Developmental Review*, *22*(2), 162–210.

Zhang, L. F., & Sternberg, R. J. (1998). Thinking styles, abilities, and academic achievement among Hong Kong University students. *Hong Kong Educational Research Association Educational Research Journal*, *13*, 41–62.

29 Emotional Intelligence

Susan E. Rivers, Isaac J. Handley-Miner, John D. Mayer, and David R. Caruso

The term "emotional intelligence" first appeared in the psychological literature in 1990 and was defined as "the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions" (Salovey & Mayer, 1990, p. 189). However, it was a trade book published in 1995 called *Emotional Intelligence* that introduced the concept of an emotional intelligence to the general public as well as to researchers more broadly (Goleman, 1995). Subsequent headlines reflected bold claims about emotional intelligence (often referred to as "EQ" for emotional quotient) being more important than analytical intelligence (IQ) (e.g., Gibbs, 1995). The notion of "EQ" became a catchall for a number of interpersonal skills, abilities, and personality traits (Mayer, Salovey, & Caruso, 2008). The influence of the concept is now widespread. Harvard Business School published a bestselling compilation of articles on emotional intelligence (Goleman, Boyatzis, & McKee, 2015), corporate trainings on emotional intelligence abound, and hundreds of trade books have been published on the topic. Universities offer emotional intelligence training for staff and several even mandate it for first-year students – although programs vary from the catchall approach to a narrower view based on the version proposed in 1990.

Since its introduction, the study and practice of emotional intelligence has expanded and, given the range of constructs under what became a very broad umbrella term, some have attempted to categorize the various approaches into "ability" models and "mixed" models (Mayer, Salovey, & Caruso, 2000). Ability models view emotional intelligence as a construct related to other intelligences and consisting of a set of mental abilities. Mixed models view emotional intelligence as a blend of standard personality traits and various abilities. We explore these in depth in the "What Is Emotional Intelligence?" section of this chapter.

The two strains have generated a great deal of research. Some studies demonstrate discrete validity between emotional intelligence and neighboring concepts such as general intelligence and the Big Five personality traits, while other studies suggest that emotional intelligence, or more specifically EQ, is nothing more than a repackaging of the Big Five or other characteristics (Brackett & Mayer, 2003; Tett, Fox, & Wang, 2005). Although some of the initial claims made by the media – such as emotional intelligence being a far better predictor of success than IQ – have since been tempered (in Goleman's introduction to the tenth anniversary edition of

Emotional Intelligence, for example; Goleman, 2005), empirical evidence continues to point to the importance of some conceptions of emotional intelligence for school, work, social, and health outcomes (Brackett, Rivers, & Salovey, 2011).

Humans are spending an increasing amount of time in digital and virtual environments and a large sector of the US and global economy is devoted to designing machine learning and artificial intelligence systems to further revolutionize daily life (Bughin et al., 2017). As virtual worlds increasingly intersect and meld with physical ones, sometimes becoming indistinguishable for working, learning, and social interaction, how will our understanding of emotional intelligence be transformed? Is it an intelligence that will help us design environments that improve the human condition, relieve suffering, and promote equity? Or is emotional intelligence value-neutral – that is, can it be applied to achieve any outcome, whether deemed a social good or not? Examining the extant research on the skills that comprise emotional intelligence, the correlates of those skills, and how the skills develop has the potential to inform the design of emotionally intelligent systems to improve health and well-being, productivity, relationships, and quality of life in profound ways.

In this chapter, we review some of the most popular models of emotional intelligence and the measures associated with each. We then offer an overview of the behavioral and neural correlates of the ability conception of emotional intelligence before discussing how emotional intelligence develops, the degree to which it is malleable in adults, and an existing school-based program designed to promote emotional intelligence skills. We conclude with an exploration of possibilities for the research landscape in the next thirty years.

What Is Emotional Intelligence?

The Ability Model

The seminal model of emotional intelligence, proposed by Salovey and Mayer (1990), represented emotional intelligence as the ability to effectively process and act on affective information gathered about both the self and others. Perceiving and expressing emotions, regulating emotions, and utilizing emotional states to enhance thinking and motivation were posited as the specific mental processes that comprise one's emotional intelligence. The most widespread ability model of emotional intelligence today is based on an update to the original model and specifies four categories of related abilities: (1) perceiving emotions, (2) using emotions to facilitate thought, (3) understanding emotions, and (4) regulating emotions (Mayer & Salovey, 1997) (see Table 29.1). Recent revisions to this model are outlined in Mayer, Caruso, and Salovey (2016).

In this model, each of the four branches is composed of several interrelated skills – ranging from basic to advanced – which are acquired over the course of development. The branches themselves also ascend by degree of complexity: perceiving emotions is thought to involve a simpler set of abilities than the more deliberative, willful regulation of emotions.

The first branch, perceiving emotions, encompasses recognizing and expressing emotions, including skills such as identifying and distinguishing between the emotional states one experiences and accurately detecting the sincerity of another person's emotional expression. The second branch, using emotions to facilitate thought, includes skills for generating and exploiting emotional states, such as knowing which emotional state is best for orienting attention to relevant information and generating a specific emotional state to expedite a judgment or choice. The third branch, understanding emotions, includes comprehending the causes of and relationships between emotions. Skills in this branch include accurately labeling emotions and understanding complex emotions, such as differentiating feelings of envy and jealousy. The fourth branch, managing emotions, refers to effectively managing emotional states in the pursuit of specific goals. This branch includes skills such as remaining open to both positive and negative emotional states and selectively engaging with emotions depending on the value of that emotion to a given goal (Mayer & Salovey, 1997). The branches have remained conceptually intact since their introduction; however, Mayer, Caruso, and Salovey (2016) have divided some of the skills within each branch into two or more specific abilities and added new skills to the using and understanding branches.

Although the ability model may not explicitly endorse a basic emotions approach, it was based on the premise that certain aspects of emotion (e.g., facial expressions, physical sensations) are universal. The question of universality is now under substantial debate in the field, however (e.g., Ekman & Cordaro, 2011; Gendron et al., 2014). If one subscribes to a social-constructivist theory of emotions, an ability model of emotional intelligence may "look" a lot different and, instead, may focus on the knowledge of cultural display rules or differences across cultures in the definition of emotion. In one study of emotional intelligence across five different countries, and in another looking at differences between a sample from the United States and China, it appears that there is close to universality in the way people answer the emotion perception items; however, there is a good deal of difference in how people in the United States and China define an "effective" emotion management strategy (although there was more "sharedness" than specificity even in emotion management) (Shao, Doucet, & Caruso, 2015).

Measuring ability emotional intelligence. The Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) was developed to measure emotional intelligence as represented by the four-branch ability model. This measure, in its current form (MSCEIT v2.0), contains 141 items and is divided into four sections, each corresponding to one of the four branches of the model: perceiving emotions, using emotions to facilitate thought, understanding emotions, and regulating emotions (Mayer et al., 2003). Each of the four sections contains two types of tasks. In the emotion perception section, respondents determine (1) the emotions being expressed on provided pictures of faces and (2) the emotions conveyed in provided pieces of artwork. In the using emotions to facilitate thought section, respondents (1) produce a specific emotion and answer questions about the sensations that accompany this emotional state (e.g., how cold it is) and (2) assess which emotions would best assist

Table 29.1 *The four-branch model of emotional intelligence, with added areas of reasoning (from Mayer et al., 2016).*

The Four Branches	Types of Reasoning[a]
4. Managing emotions	• Effectively manage others' emotions to achieve a desired outcome[b] • Effectively manage one's own emotions to achieve a desired outcome[b] • Evaluate strategies to maintain, reduce, or intensify an emotional response[b] • Monitor emotional reactions to determine their reasonableness • Engage with emotions if they are helpful; disengage if not • Stay open to pleasant and unpleasant feelings, as needed, and to the information they convey
3. Understanding emotions	• Recognize cultural differences in the evaluation of emotions[c] • Understand how a person might feel in the future or under certain conditions (affective forecasting)[c] • Recognize likely transitions among emotions such as from anger to satisfaction • Understand complex and mixed emotions • Differentiate between moods and emotions[c] • Appraise the situations that are likely to elicit emotions[c] • Determine the antecedents, meanings, and consequences of emotions • Label emotions and recognize relations among them
2. Facilitating thought using emotions	• Select problems based on how one's ongoing emotional state might facilitate cognition • Leverage mood swings to generate different cognitive perspectives • Prioritize thinking by directing attention according to present feeling • Generate emotions as a means to relate to experiences of another person[c] • Generate emotions as an aid to judgment and memory
1. Perceiving emotion	• Identify deceptive or dishonest emotional expressions[b] • Discriminate accurate vs. inaccurate emotional expressions[b] • Understand how emotions are displayed depending on context and culture[c] • Express emotions accurately when desired • Perceive emotional content in the environment, visual arts, and music[b] • Perceive emotions in other people through their vocal cues, facial expression, language, and behavior[b] • Identify emotions in one's own physical states, feelings, and thoughts

Note.

[a] The bullet points are based on Mayer and Salovey (1997) except as indicated in superscripts b and c. Within a row, the bulleted items are ordered approximately from simplest to most complex, bottom to top. The four-branch model depicts the problem-solving areas of emotional intelligence and is not intended to correspond to the factor structure of the area.

[b] An ability from the original model was divided into two or more separate abilities.

[c] A new ability was added.

[d] Note that the Branch 2 abilities can be further divided into the areas of generating emotions to facilitate thought (the bottom two bulleted items) and tailoring thinking to emotion (the top three bulleted items).

Reprinted from "The ability model of emotional intelligence: Principles and updates," by J. D. Mayer, D. R. Caruso, & P. Salovey, 2016, *Emotion Review, 8*(4), p. 294. Copyright 2016 by SAGE Publications. Reprinted with permission.

with a given task or behavior. In the understanding emotions section, respondents (1) determine which emotions could combine to form a different given emotion and (2) identify which emotion would most likely stem from an intensification of a given emotion. Finally, in the emotion regulation section, respondents determine (1) which actions a character should take to generate a specific emotional state, based on a provided fictional story, and (2) the best way for someone to manage another person's emotional state (Mayer et al., 2003). A youth version of the MSCEIT (the MSCEIT-YRV) designed for those aged ten to eighteen years is also available (Mayer, Salovey, & Caruso, 2014). Its structure is similar to the adult version of the MSCEIT with distinct tasks for each of the four branches of the model.

Responses to items on the MSCEIT are scored along a continuum of "correctness" based on two scoring methods: consensus and expert. Each response receives a proportion of points based on the extent to which a representative sample (consensus scoring) or sample of emotions scholars (expert scoring) agreed with each response choice. For example, in consensus scoring, if an individual selected a response that 66 percent of general respondents also chose, the individual would receive 0.66 points for that response. Points are aggregated across all the items and respondents are given both a general-consensus score and an expert-consensus score. The general consensus versus expert consensus has been shown to correlate highly ($r = 0.91$), as have respondents' scores measured by these two methods ($r = 0.98$) (Mayer et al., 2003). More recently, the YRV replaces consensus scoring with "veridical" scoring whereby relevant literature on emotions is used to construct the scoring algorithm (Mayer et al., 2014).

In the years following the development of the MSCEIT v2.0, studies have examined whether it parses into four factors in accordance with the theoretical model. Several such analyses found that the MSCEIT v2.0 does not, after all, neatly apportion into four factors. Some argue that the MSCEIT is best represented by a single factor (Legree et al., 2014). Others have argued that the second branch – using emotion to facilitate thought – should be omitted from the model (Fan et al., 2010; MacCann et al., 2014; Maul, 2011; Palmer et al., 2005; Rossen, Kranzler, & Algina, 2008). A three-factor model, proposed by Joseph and Newman (2010), posits that emotion perception is causally antecedent to emotion understanding, which in turn causally precedes emotion regulation. This model integrates cognitive ability and elements of the Big Five and proposes that conscientiousness supports emotion perception, cognitive ability supports emotion understanding, and low neuroticism supports emotion regulation (Joseph & Newman, 2010). In a recent update to their model, Mayer, Caruso, and Salovey (2016) agreed that the accumulated data do not support the second branch (using emotions to facilitate thought) as a distinct factor. They argue, however, that it is indeed a means through which people solve emotion-related problems and should therefore remain a component of the model even if the construction of the MSCEIT fails to capture it as a discrete ability (Mayer et al., 2016). Another concern about the MSCEIT is how well the items that comprise the first branch assess emotion perception skills, as this branch has been shown to have low convergent validity with other ability-based measures of emotion perception, such as the Japanese and Caucasian Brief Affect Recognition Test (JACBART) and the Diagnostic Analysis of Nonverbal Accuracy Scales (DANVA-2) (Mayer, Roberts, & Barsade, 2008; Roberts et al., 2006).

Mixed Models

Several other conceptualizations of emotional intelligence were proposed in the literature following Salovey and Mayer's introduction of the ability model of emotional intelligence. These models broadened the catalogue of emotional intelligence components to include not just mental abilities associated with using emotions to guide thinking and thinking to guide emotions but also traits and attributes. Because these models generally include at least one of the core abilities identified in the four-branch model (e.g., emotion regulation) and nonability qualities (e.g., happiness or optimism), these conceptions of emotional intelligence often are referred to as "mixed" models (Mayer et al., 2000).

One popular model, the Bar-On Model of Emotional-Social Intelligence, explains individual performance as a consequence of a host of intrapersonal and interpersonal competencies (Bar-On, 1997, 2006). These competencies include several of the abilities outlined in Salovey and Mayer's four-branch model but also a range of dispositional features, such as optimism, self-regard, and impulse control. Another well-known model, proposed by Petrides and Furnham (2001, 2003), conceptualized "trait emotional intelligence," which is defined as "a constellation of emotion-related self-perceptions and dispositions, assessed through self-report" (Petrides & Furnham, 2003, p. 40). This construct includes dispositional features such as adaptability, assertiveness, and self-esteem (Petrides & Furnham, 2001).

Mixed-model measures. Bar-On's concept of Emotional-Social Intelligence is measured by the Emotional Quotient Inventory (EQ-i) 2.0. It consists of 133 Likert-style items that assess an individual's self-perceived competencies across five categories: self-perception, self-expression, interpersonal, decision-making, and stress management (Multi-Health Systems, 2011). Another mixed-model approach, the Petrides and Furnham (2003) model of trait emotional intelligence, gave rise to the Trait Emotional Intelligence Questionnaire (TEIQue). The TEIQue asks individuals how much they agree with 144 items that address fifteen categories of abilities and traits, including emotion perception (e.g., "I often find it difficult to recognize what emotion I'm feeling"), happiness (e.g., "Life is beautiful"), and self-esteem (e.g., "I believe I'm full of personal strengths") (Petrides & Furnham, 2003, p. 47). Other well-known measures of mixed or trait emotional intelligence include the Schutte Emotional Intelligence Scale (Schutte et al., 1998), which relies on self-report, and the Emotional Competence Inventory (Boyatzis, Goleman, & Rhee, 2000), which employs self-report and other-report techniques.

Which Emotional Intelligence Model Is Correct?

The debate over the different conceptions of emotional intelligence has at times been framed as a matter of which model – ability or mixed – is "correct"; however, it is not necessarily the case that only one of these constructs is valid (Petrides & Furnham, 2001). It is possible, for example, that the trait-based model described by Petrides and Furnham (2001, 2003) represents a wholly independent (or perhaps slightly overlapping) construct from the ability-based model described by Mayer and Salovey (1997). The dispute, rather, is in large part over which model should lay claim to the title of "emotional intelligence." Multiple distinct constructs using the same identifier adds confusion to the field (Cherniss, 2010; Mayer, Salovey, et al., 2008) and it is likely that one model does in fact better explain how emotional abilities are harnessed and deployed through a means characteristic of the term "intelligence." Ultimately, an intelligence is an aptitude, not a disposition, and should be comprised of interrelated abilities as assessed by performance measures (Mayer, Caruso, & Salovey, 1999). To assess intelligence, people are not asked to report how smart they believe they are (self-report), nor are their family or peers asked how smart they are (other-report); people are asked to perform tasks to demonstrate their intelligence (performance). Thus, we argue that the four-branch model of emotional intelligence most accurately represents an "emotional intelligence."

Some of the strongest evidence that the ability model of emotional intelligence can be considered an intelligence comes from recent work by MacCann and colleagues. Results from an in-depth study requiring up to eight hours of time from participants revealed that a three-factor model of ability emotional intelligence – as measured by the MSCEIT – meets the criteria for a second-stratum factor of intelligence in the Cattell-Horn-Carroll (CHC) model, alongside factors such as crystallized intelligence and fluid intelligence (MacCann et al., 2014). See Mayer (2018) for an in-depth explanation of emotional intelligence as a "broad intelligence" in the CHC model of intelligence, including how it relates to other people-centered intelligences. (For more information on people-centered intelligences, see Mayer, Panter, & Caruso, 2012 for an overview of personal intelligence; Conzelmann, Weis, & Süß, 2013 for an overview of social intelligence; and Wagner and Sternberg, 1985 for an overview of practical intelligence. See also Chapters 30 and 31 in this volume.)

The remainder of this chapter will only report on research employing ability models, unless otherwise specified. The previous edition of this handbook also provides more information on the debate over the different conceptualizations of emotional intelligence (Mayer et al., 2011). Recent developments on the mixed and trait models of emotional intelligence are available elsewhere (e.g., Petrides et al., 2016).

What Are the Correlates of Emotional Intelligence?

Behavioral Correlates

One of the reasons Daniel Goleman's account of the phenomenon garnered so much attention was its strong claims regarding the importance of emotional intelligence for success. Although there is no empirical evidence for claims such as "for star performance in all jobs, in every field, emotional competence is twice as important as purely cognitive abilities" (Goleman, 1998, p. 34), there is evidence that emotional intelligence impacts many aspects of one's life, including social functioning, educational performance, workplace effectiveness, and health and well-being.

Social functioning. The four-branch model posits that emotional intelligence predicts aspects of social effectiveness, including the abilities to accurately read others' emotional states and generate positive emotions in oneself and others. Thus, we would expect scores on the ability measure of emotional intelligence to correlate significantly with measures of the quality of interpersonal relationships. Controlling for factors such as general intelligence and the Big Five personality traits, scores on the MSCEIT are positively and significantly related to a host of social variables, including relationship quality (Brackett, Warner, & Bosco, 2005; Lopes, Salovey, & Straus, 2003; Lopes et al., 2004, 2005) and interpersonal sensitivity (Lopes et al., 2005), using both self reports and peer reports for outcome measures.

MSCEIT scores in college student samples also are negatively related to levels of maladaptive social tendencies: those scoring higher on the MSCEIT have fewer incidents of aggressive behavior, including indirect aggression (e.g., exclusion, spreading rumors) (García-Sancho, Salguero, & Fernández-Berrocal, 2016) and physical aggression (García-Sancho, Salguero, & Fernández-Berrocal, 2017). Further, those with higher scores on emotion perception abilities as measured by the MSCEIT are less likely to use styles of humor associated with hostility (e.g., teasing, sarcasm) (Yip & Martin, 2006). Additionally, adolescents who score higher on the MSCEIT-YRV tend to be rated by their teachers as having fewer behavioral problems in school, less social stress, and lower levels of clinical maladjustment (Rivers et al., 2012).

It is possible, however, that emotional intelligence is not only used for achieving social good. Emotional intelligence has been found to facilitate interpersonally deviant behavior when the emotionally intelligent individual also possesses Machiavellian traits (Côté et al., 2011). Emotional intelligence may make it easier for those with deviant motives to be deviant and may allow one to emotionally manipulate, exploit, and control others (Mayer, 2001). Others have suggested that, even when deviant motives are not present, it is possible that those who are high on emotional intelligence may use their skills strategically for personal gain, perhaps at the expense of others (Kilduff, Chiaburu, & Menges, 2010).

Academic performance. Emotional intelligence correlates with academic success. In a diverse sample of 273 fifth and sixth graders, scores on the MSCEIT-YRV correlated moderately to strongly with teacher-reported student work habits ($r = 0.38$) and grades in English language arts ($r = 0.65$) and math ($r = 0.51$) (Rivers et al., 2012). These results were supported by a longitudinal study conducted with 413 students aged eleven to twelve years, which showed that scores on the MSCEIT-YRV predicted academic success five years later (Qualter et al., 2012). This effect also holds for older students; a recent study found that, among a sample of undergraduates, scores on the MSCEIT were correlated significantly with grade point average (GPA), controlling for IQ (scores on Raven's Advanced Progressive Matrices) and personality (scores on the reduced version of the Eysenck Personality Questionnaire) (Lanciano & Curci, 2014). Other studies have found correlations between MSCEIT scores and grades in college students; however, these correlations become nonsignificant once verbal intelligence is controlled for

(Barchard, 2003; Brackett & Mayer, 2003). This is not necessarily surprising given that the MSCEIT is a fully language-based assessment and Barchard (2003) and Brackett and Mayer (2003) report correlations of $r = 0.44$ and $r = 0.32$, respectively, between scores on the MSCEIT and verbal intelligence.

Workplace effectiveness. There also is evidence to suggest that emotional intelligence and workplace success are linked. A meta-analysis of nine studies assessing the relationship between work performance and MSCEIT scores showed that emotional intelligence correlated with work performance ($r = 0.24$), even after controlling for personality and cognitive ability (O'Boyle et al., 2011). Another study, investigating the means through which emotional intelligence might improve job performance, found teamwork effectiveness to be a mediator between emotional intelligence and job performance in managerial positions (Farh, Seo, & Tesluk, 2012). Several other studies report a similar connection between emotional intelligence and job performance, identifying potential mediating mechanisms, including the ability to tolerate stress (Lopes et al., 2006), the experience of positive affect (Brackett et al., 2010), leadership emergence (Côté et al., 2010), and leadership effectiveness (Rosete & Ciarrochi, 2005).

Health and well-being. Empirical evidence suggests a relationship between emotional intelligence and well-being. A meta-analysis of eleven studies assessing the relationship between MSCEIT scores and measures of mental health reported a small but significant relationship between emotional intelligence and mental health ($r = 0.17$) (Martins, Ramalho, & Morin, 2010). More specifically, MSCEIT scores correlate negatively with depression (Davis & Humphrey, 2012; Fernández-Berrocal & Extremera, 2016; Hertel, Schütz, & Lammers, 2009), general and social anxiety (Jacobs et al., 2008; O'Connor & Little, 2003), borderline personality disorder (Gardner & Qualter, 2009; Hertel et al., 2009), anorexia nervosa (Hambrook, Brown, & Tchanturia, 2012), and psychopathy (Ermer et al., 2012).

Psychological health is not simply the absence of mental illness – it also encapsulates the presence of positive qualities of well-being (Seligman & Csikszentmihalyi, 2000). To that end, a recent meta-analysis examining the relationship between emotional intelligence and subjective well-being identified three studies assessing the connection between MSCEIT scores and subjective well-being and reported a modest, yet significant, and positive relationship ($r = 0.22$) (Sánchez-Álvarez, Extremera, & Fernández-Berrocal, 2016).

Neural Correlates

Since the first edition of this handbook (Sternberg & Kaufman, 2011), there has a been a substantial increase in neuroscientific research on emotional intelligence. Two lesion-mapping studies conducted on a large sample of Vietnam War veterans provide insight into the neural mechanisms that might underlie emotional intelligence abilities (Barbey, Colom, & Grafman, 2014; Operskalski et al., 2015). Using a sample of participants who suffered traumatic brain injuries, researchers applied voxel-based lesion-symptom mapping (VLSM) to compare specific brain areas in patients with damage in that area to those with no damage in that area. VLSM is a technique that allows for high statistical power while controlling for the general effects of having undergone a traumatic brain injury (Barbey et al., 2014). In their first study, Barbey and colleagues administered the MSCEIT to 152 brain-trauma patients and used VLSM to identify which damaged brain areas explained the particular deficits in emotional intelligence displayed by these patients. They found that emotional intelligence is linked to neural processes that fall along a social cognitive network, including the areas that support "the uniquely human ability to reason about the contents of mental states" (i.e., the left temporoparietal junction) and the areas that support "emotional empathy" and "shared attention and collaborative goals" (i.e., the left orbitofrontal cortex) (Barbey et al., 2014, p. 267). Barbey and colleagues also report that emotional intelligence abilities were associated with white matter systems connecting the frontal, temporal, and parietal lobes – in particular, the superior longitudinal/arcuate fasciculus, which provides evidence that emotional intelligence abilities are not fully housed in distinct swathes of the brain but depend on information being shared across cortices via this social cognitive network. Their analysis also suggested that emotional intelligence and general intelligence (as measured by the Wechsler Adult Intelligence Scale – Third Edition) draw on many of the same networks, including the dorsal and ventral perisylvian language systems (Barbey et al., 2014).

Employing similar methods to assess the four branches of emotional intelligence individually, a study by Operalski and colleagues (2015) revealed that emotion perception abilities and emotion regulation partially rely on distinct neural structures, despite engaging many of the same areas of the brain. In this study, emotion perception abilities – but not emotion regulation abilities – were related to activity in the ventral temporal lobe, the area associated with categorical conceptual knowledge and emotional expression processing (e.g., Harry et al., 2013). Emotion regulation abilities were linked to three areas: those associated with decision-making and emotional processing (i.e., right posterior orbitofrontal cortex), relational memory and emotional processing (i.e., posterior hippocampus), and attention, executive control, theory of mind, episodic memory, and language processing (i.e., portions of the parietal cortex) (Operskalski et al., 2015).

The results reported by Barbey and colleagues (2014) and Operskalski and colleagues (2015) need to be interpreted within the context of the VLSM methodology and its shortcomings. In particular, VLSM does not provide information about brain mechanisms that are undamaged in the sample of traumatic brain injury patients, which could explain why the other two branches of emotional intelligence (using emotions to facilitate thought and understanding emotions) did not significantly map onto the sample in the second study (Operskalski et al., 2015). Furthermore, it is important to remember that this is purely correlational data from a relatively homogeneous sample. Nevertheless, these findings represent an important step in our understanding of emotional intelligence and provide some initial holistic glimpses of the networks employed in generating the abilities specified by the four-branch model, supporting the division of emotional intelligence into branches of abilities.

Findings from electroencephalography (EEG) activation studies also suggest that those who score highly on the MSCEIT display lower levels of neural activation when solving emotion-related problems, suggesting higher levels of neural efficiency (Jaušovec & Jaušovec, 2005; Jaušovec, Jaušovec, & Gerlič, 2001). The neural efficiency hypothesis is further supported by functional magnetic resonance imaging (fMRI) data collected from individuals engaged in a social exchange reasoning task, which revealed a negative correlation between MSCEIT scores and brain activity (Reis et al., 2007). More recently, Pisner and colleagues (2017) found correlations between scores on the understanding emotions and regulating emotions sections of the MSCEIT and measures of white matter along several important neural tracts, even controlling for IQ. Their findings suggest that stronger emotional intelligence abilities might be due in part to the speed and efficiency of communication between areas of the brain key to affective perception and processing (Pisner et al., 2017). However, causal relationships cannot be determined from these data.

Summary

This section presented research findings on the behavioral and neural correlates of emotional intelligence, using primarily MSCEIT scores as the ability measure of emotional intelligence. MSCEIT scores predict effectiveness across social, academic, and workplace domains, as well as mental health. In terms of neural correlates, higher MSCEIT scores appear to be associated with greater neural efficiency in brain areas implicated in emotion processing. The studies conducted by Barbey and colleagues bring us closer to understanding which brain areas are involved in the emotional intelligence abilities assessed by the MSCEIT.

More research is needed to understand the causal effects of emotional intelligence to move our understanding beyond its correlates. Toward this end, researchers must develop an effective paradigm for experimentally manipulating emotional intelligence.

Development of Emotional Intelligence

In this section, we review one of the few existing models of emotional intelligence development. We also explore the evidence for the malleability of emotional intelligence and discuss some recent attempts to promote more effective emotional intelligence development in schools.

Developmental Models

One of the only development models of emotional intelligence, proposed by Zeidner and colleagues (2003), outlines three interrelated processes that contribute to the development of emotional intelligence abilities: temperament, rule-based skill acquisition, and self-aware emotion regulation.

First, the temperament – the largely biologically informed differences in the type and intensity of emotions experienced – of an infant is expected to influence or interact with the coping of emotions early in life. The reinforcement of emotion expression and regulation tendencies by caregivers likely shapes an infant's developing abilities for perceiving, understanding, using, and managing emotions.

Second, rule-based skill acquisition refers to a process whereby individual differences in emotional intelligence abilities stem from aptitudes for emotion expression and management skills. According to Izard (2001), this can be thought of as a capacity for emotional adaptiveness, which is at least partially tied into the development of verbal skills that allow for emotional expression and, eventually, labeling emotions. In other words, children who more accurately perceive and express emotions are then able to learn better regulation strategies – and these early skills likely compound in a cycle of reinforcement. Rule-based skill acquisition is thought to also hinge on the modeling and reinforcement the child receives from his or her social environment (Zeidner et al., 2003). This perspective of abilities resulting from a rule-based acquisition process aligns well with the four-branch model, whose hierarchical structure designates emotion perception as antecedent to higher-order abilities such as understanding and regulating emotions (Mayer & Salovey, 1997).

The third process of development specified by Zeidner and colleagues (2003) is self-aware emotion regulation, which concerns children's ability to reflect on their emotional states and is believed to occur later in the maturation process. Self-aware emotion regulation involves explicit emotional self-evaluation, as opposed to implicit adoption of forms of emotional expression and regulation that caregivers model. This process hinges on caregivers – in particular whether and how caregivers discuss emotion and its expression with children. As children grow older, broader sociocultural influences – peers, institutions, cultural representations, and so on – play a prominent role. Ultimately, the metacognitive component of self-aware emotion regulation allows children to develop a more conscious and deliberate emotional understanding (i.e., skills that fall under the third branch of the ability model), which then paves the road for the development of more sophisticated emotion regulation strategies (Zeidner et al., 2003).[1]

This model links these three processes, positing that temperament influences how children develop rules related to emotional expression, which, as children gain a greater capacity to reflect on their emotional state, graduates to the more self-directed process of self-aware emotion regulation (Zeidner et al., 2003). Temperament and rule-based acquisition are thought, however, to continue to distinctly impact development even after children develop the self-reflective abilities marked by self-aware emotion regulation. The model specifies further that these three identified processes are intertwined with genetic influences, caregivers, peer networks, and culture and together these processes and influences shape the trajectory of a child's emotional development.

This developmental model offers a framework to understand how individual differences in emotional intelligence might arise through environmental and

[1] See Zeidner and colleagues (2003) for a thorough review of the literature relevant to the three processes discussed in this section. Critiques of the model have also been published (e.g., see Arsenio, 2003; Fox, 2003).

socialization processes in childhood. Its emphasis on the interaction between the innate (i.e., temperament) and the learned (i.e., skill acquisition and self-reflection) and the implication of one's sociocultural environment in emotional intelligence development are critical. Although these factors surely come as little surprise to social and developmental psychologists, they are important to emphasize because of their implications for the value of environments that are intentionally designed to teach and support skills associated with emotional intelligence development.

One critical barrier to the research on emotional intelligence development is the absence of a validated scale to measure ability emotional intelligence in young children. The closest the field has seen is the MSCEIT-YRV but it is not designed for children under the age of ten and so it is still not possible to reliably assess the four-branch model of emotional intelligence in young children. Given that the processes discussed in this section manifest much earlier than the age of ten, the lack of valid measurement instruments remains a significant limitation for testing theories of emotional intelligence development and is a major reason for the dearth of empirical work in this area. Developing a measure for this age group is no small feat, however, as it must overcome the obstacle of limited verbal acuity. There is, of course, a wealth of research investigating emotional development – including many components of the four branches – but it is not immediately evident how it all coheres to form a strong theory that can offer straightforward and testable predictions about the development of a holistic emotional intelligence.

How Malleable Is Emotional Intelligence?

It is unknown the extent to which emotional intelligence is malleable in adults. There is a lack of rigorous experimental studies that use ability-based measures to answer this question. That said, the limited evidence we do have suggests that at least some of the four branches of emotional intelligence may be amenable to improvement through intervention.

In one study, participants showed improvements on the understanding emotions and managing emotions sections of the MSCEIT – but not the perceiving emotions and using emotions to facilitate thought sections – after attending eleven classes about emotional intelligence compared to a control group that took classes on career or business topics (Dacre Pool & Qualter, 2012). Another study randomly assigned a group of athletes to either a no-treatment control condition or an emotional intelligence training program that consisted of ten workshop sessions and found that the treatment group experienced significant increases in their total MSCEIT scores as well as their scores for each of the four branches (Crombie, Lombard, & Noakes, 2011). Additionally, Reuben and colleagues found a small increase in MSCEIT scores of MBA students attending a semester-long course on ability emotional intelligence compared to those in a control group (Reuben, Sapienza, & Zingales, 2009). Although promising, the results from these studies should be interpreted cautiously until replicated using random assignment and large sample sizes.

Additional studies have explored the malleability of individual branches of emotional intelligence. For example, mindfulness training (see Teper, Segal, & Inzlicht, 2013 for an overview) and even working memory training (Schweizer et al., 2013) have been shown to increase emotion regulation abilities. There is also evidence that people's abilities to perceive emotions in facial expressions can be increased through training (Kemeny et al., 2012; Matsumoto & Hwang, 2011). Furthermore, aiming to address the shortcomings of previous research assessing the malleability of emotional intelligence, Herpertz, Schütz, and Nezlek (2016) conducted a rigorous randomized controlled trial evaluating the efficacy of an emotion perception training and reported a significant increase in scores on the emotion perception section of the MSCEIT, which held when tested six months later. This finding, however, should be viewed in light of the questionable validity of the first branch of the MSCEIT (Mayer, Roberts, & Barsade, 2008).

There is also preliminary evidence that emotional intelligence may develop up until middle age. A recent study of 12,198 participants aged seventeen to seventy-six found a U-shaped relationship between age and scores on the MSCEIT, suggesting that emotional intelligence may increase through middle age before declining as one gets older (Cabello et al., 2016). Follow-up studies should aim to test how emotional intelligence changes over the course of a lifetime within – rather than between – participants to ensure that this analysis is not picking up any generational differences.

Taken in aggregate, these findings support the possibility that emotional intelligence is malleable, at least to some degree, even in adult populations. However, the lines of research discussed thus far deserve further study and there needs to be much more research investigating the tractability of the using emotions to facilitate thought and understanding emotions branches.

Efforts to Improve Emotional Intelligence

A slew of emotional intelligence training programs and courses have cropped up in the last few decades in light of the demonstrated associations between emotional intelligence and social, educational, workplace, and well-being benefits discussed in this chapter. Many of these programs are steeped in the mixed-model approach and many have not been evaluated rigorously for efficacy (as evidenced by the lack of research on the malleability of emotional intelligence); however, there are a select few rigorously designed programs that address many of the skills of the four-branch model. This section will focus on one such program, RULER, as an exemplar of a program informed by research on ability-based emotional intelligence.

RULER is a curriculum designed to teach five core social-emotional skills to school-aged children: recognizing, understanding, labeling, expressing, and regulating emotions (Torrente, Rivers, & Brackett, 2016). The curriculum is based on the four-branch model and seeks to teach skills aligned with the four core abilities. RULER specifies a range of age-appropriate activities and modules that teachers can use in their classrooms. A quasi-experimental study reported encouraging results from the implementation of RULER at the middle school level, including increases in student social-emotional competence, good work habits, and English-language arts grades (Brackett et al., 2012).

RULER is designed to promote the development of social-emotional skills not only in students but also in teachers and administrators. At schools that implement RULER, faculty and administrators participate in "train-the-trainer" workshops that educate them about the same social-emotional skills they will be promoting in the classroom (Torrente et al., 2016). A randomized controlled trial of RULER in sixty-two middle schools found teacher and classroom outcomes, including improvements in overall emotional climate of the classroom (as judged by both teachers and outside observers) (Rivers et al., 2013), and subsequent boosts in classroom effectiveness, including instructional support and classroom organization two years after adoption (as judged by outside observers) (Hagelskamp et al., 2013). These studies provide preliminary evidence that adopting the RULER curriculum and attending RULER training workshops may impact teachers' behaviors. To date, there have not been experimental studies testing the extent to which RULER impacts emotional intelligence scores of teachers and students.

Designing Environments to Promote Emotional Intelligence

The empirical evidence we have to date points to emotional intelligence as an important factor in critical components of daily life: social functioning, educational and workplace success, and mental health and well-being. Schools are designed to promote the acquisition of knowledge and the establishment and maintenance of positive social relationships; workplaces are designed to encourage efficiency and collaboration; and handheld "smart" devices are designed to facilitate communication, entertainment, and efficiency – why not also design these environments to foster a suite of emotional intelligence abilities that appear to have such positive and far-ranging effects? Especially as artificial intelligence and machine learning advances continue to expand our potential to "humanize" the virtual environments we increasingly inhabit, we have a ripe opportunity to design these environments to promote emotional intelligence skill building.

Fine-Tuning the Environment

We are shaped by the ecological and cultural strata we inhabit (Bronfenbrenner, 1977; Markus & Kitayama, 2010). Not only do our peers and family members influence our development and behavior but so do the institutions in which we participate, the governmental systems to which we belong, the media we consume, and the larger cultures within which we operate.

Programs founded on emotional intelligence theory and evidence could be adapted for university, corporate, or governmental contexts to successfully teach emotional intelligence skills and encourage emotionally adaptive behavior. But clearly not all environments would be suitable for a full-on educational program. Instead, what if we acknowledged the strengths of the holistic approach of RULER and also took a leaf out of the behavioral economics literature? Nudge theory, proposed by Thaler and Sunstein (2008), argues that systems will always influence people in *some* direction – whether grocery stores situate sugary cereals at eye level or on the top shelf, this placement necessarily nudges a change in cereal-purchasing behavior. Presumably our environments also have the potential to nudge our emotional learning and virtual environments may provide a means of nudging a large percentage of the population.

For example, what if, every time you open Facebook, above your profile photo the current prompt "What's on your mind?" was replaced with one that suggests deeper reflection of your own or others' emotional states? Could seeing this nuanced prompt a dozen or more times a day help Facebook users practice emotion recognition? The nudge literature documents many powerful effects from small tweaks and, given that, as of March 2018, Facebook boasted 1.45 billion daily active users (Facebook, 2018), emotional nudges in contexts such as social media seem an avenue worth investigating.

Nudges offer an intriguing path toward the intentional design of emotionally intelligent environments. By considering the influence of the subtle cues of our environments, nudges represent a broader approach to emotional intelligence development than traditional, individual-focused methods. Ultimately, however, nudges are passive; they may encourage the practice of emotional intelligence skills but may not effectively educate. So, while considering and testing the effects of nudges, we should simultaneously build explicit education and training into our physical and virtual environments.

Virtual Learning

When people use computers and smart devices, they often behave as if they were interacting directly with a sentient source (Sundar & Nass, 2000). A recent study demonstrated that, when interacting with a virtual character, participants' evaluations (e.g., perceived social presence; perceived rapport) and responses (e.g., number of words used; degree of self-disclosure) were nearly the same regardless of whether they believed the virtual characters were avatars being controlled by real people or embodied conversational agents fully controlled by computer algorithms (Von Der Pütten et al., 2010). Furthermore, research indicates that robots and virtual agents who engage people in social-emotional interactions are deemed more trustworthy and supportive, which could further enhance the influence these systems have on our learning and behavior (Fan et al., 2017; Lohani, Stokes, McCoy, Bailey, Joshi, & Rivers, 2016; Lohani, Stokes, McCoy, Bailey, & Rivers, 2016).

These findings have implications for the design of technologies. If people do indeed treat computer systems as fellow social agents and autonomous sources of information, it seems probable that the behavior of these virtual agents (avatars, teammates, video game characters, digital assistants) could impact human behavior and learning. As our representations of human-like systems become more realistic (think early GPS voice-based navigation compared to Siri or the cartoonish video games from the 1990s compared to the hyperrealistic games that are currently popular), we increasingly may have the opportunity to teach through these systems.

Granic and colleagues (2014) argue that video games may already be promoting adaptive emotion regulation strategies. They posit that, by posing new challenges and altering rule systems throughout gameplay – which generates stressful situations and require players to adapt quickly to shifting expectations – many video games inherently reward reappraisal and discourage rumination. Although this hypothesis has not been tested directly and there is limited rigorous research looking at the direct effects of popular video games on emotion regulation abilities (Villani et al., 2018), a recent randomized controlled trial of a video game designed to help children regulate their anxiety found that playing the game led to a reduction in anxiety symptoms comparable to a comparison group that received cognitive behavioral therapy (CBT), even at a six-month follow-up (Schoneveld, Lichtwarck-Aschoff, & Granic, 2018). Furthermore, a quasi-experimental study testing the effects of a video game designed to teach players emotional intelligence skills found increased MSCEIT scores in the video game condition compared to a wait-list control group, including increased scores on the emotion regulation section of the test (Cejudo & Latorre, 2015).

Video games represent a particularly exciting opportunity for social-emotional skill integration because they are highly engaging, social, and popular. More than half (65 percent) of US households have someone who plays at least three hours of video games a week (ESA, 2017) and teenagers are playing more than an hour of video games a day, on average (Rideout, 2015).

Research indicates that in-game behavior can inform behavior in nongame contexts. For example, Gentile and colleagues found participants were more likely to help others in a puzzle task after playing a video game with a prosocial bent compared to playing a violent or neutral video game (Gentile et al., 2009). Perhaps games could be designed to promote emotional intelligence skills, too. A game could teach players to read emotions from facial expressions, which could help them detect traitorous characters later on in the game. Or players could progressively "collect" emotion regulation strategies that they would need to apply throughout the game to assist characters struggling to maintain their composure in stressful, frightening, or enraging situations. There are abundant opportunities to weave social and emotional learning into video games and developing emotional intelligence skills might even align with the goals of both players and game creators given that so many of the most popular video games (e.g. *World of Warcraft, Call of Duty, Overwatch*) take place in social environments that hinge on effective communication and cooperation (Rivers, 2018).

Findings from the emotional intelligence literature might also help us create systems that are, in some capacity, emotionally intelligent themselves. One study drew on the emotional regulation literature to test the effects of verbal reappraisal on drivers' negative reactions to frustrating driving conditions (e.g., getting cut off by other drivers; being stuck in heavy traffic) (Harris & Nass, 2011). The researchers found that, by having a driving simulator issue verbal reappraisals of the frustrating driving conditions (e.g., "The driver must not have seen you; otherwise he would not have chosen to change lanes"), participants drove slower and received a better driving score from the simulator, which accounts for actions such as collisions, speeding and traffic light violations, and missed stop signs (Harris & Nass, 2011, pp. 750–751). Subsequent experiments are needed to investigate whether systems like these could increase people's emotion management skills outside of the driving context and how these systems should be designed such that people would choose to use them in their daily lives.

Toward the Future: Machine Emotional Intelligence?

Smartphone apps and interfaces have the potential to dynamically react to users' emotional states. A study employing emotion recognition technology on a smartphone platform found that the software was able to detect emotions with 96 percent accuracy among a set of facial images expressing six common emotions (Alshamsi, Kepuska, & Meng, 2017). As this technology improves, might it also have the capability to dynamically *educate* people? Perhaps smartphones could help users be more aware and reflective of their current emotional state by unobtrusively signaling to users when they detect a prespecified emotion (e.g., briefly tinting the screen with a warm glow when someone is feeling happy). Given the sharp rise in the number and type of "intelligent" systems that surround us on a daily basis (smartphones, smart pens, smart refrigerators), why not make these systems *emotionally* intelligent, too?

Conclusions

We opened this chapter by referencing some of the bold claims about the predictive value of emotional intelligence that began surfacing in the mid-1990s. We explored a handful of conceptions of emotional intelligence that arose shortly thereafter and narrowed in on the four-branch ability model in order to provide a more detailed portrait of the research on the behavioral and neural correlates of emotional intelligence, the development of emotional intelligence, and the degree to which emotional intelligence is malleable. Finally, we journeyed into various possibilities for tweaking our environments to promote emotional intelligence across the life span.

Numerous areas of research on emotional intelligence need deeper and more rigorous investigation. Our understanding of how emotional intelligence maps onto brain structures is still in its infancy. There is a dearth of discussion around how emotional intelligence develops throughout childhood, despite a wealth of research on narrower facets of emotional development (e.g., emotion regulation). We have only a handful of studies, and no rigorous randomized controlled trials, assessing how much we can move adult emotional intelligence scores through intervention. And yet, there has been enormous progress in the field over the course of three decades, with many exciting contributions coming since the previous edition of this handbook. In particular, we have a much clearer picture of what emotional intelligence predicts across a host of domains important for everyday functioning and well-being; and the clear portrait that these findings paint – that emotional intelligence is associated with better social, educational, occupational, and health outcomes – should alone motivate us to delve deeper into the development and plasticity of emotional intelligence.

So, what might the next thirty years of research on emotional intelligence look like? From a methodological standpoint, it seems critical that we develop a way to assess emotional intelligence in young children. This will help us map out the developmental trajectory of emotional intelligence and perhaps give us insight into the most effective times and places to intervene. Furthermore, the field would benefit from moving beyond correlational findings; but, to accomplish such a feat, we need methods for experimentally manipulating emotional intelligence in an ethical and timely manner. Advances in assessment, and scoring of such assessments, also is key to enhancing our understanding of emotional intelligence more broadly.

From an intervention standpoint, we need randomized controlled trials designed to enhance emotional intelligence. Would emotional intelligence nudges produce any meaningful effect sizes? Are video games an effective medium for learning and practicing emotional intelligence? Or should we look to other tried-and-true interventions for inspiration? There is an extensive body of work on implicit theories of change – the degree to which one believes a certain feature is malleable – in realms such as cognitive intelligence (e.g., Blackwell, Trzesniewski, & Dweck, 2007) and personality (e.g., Yeager et al., 2013). Does encouraging someone to adopt the mindset that emotional intelligence is malleable make that person more likely to experience emotional intelligence growth? A recent study supports this possibility, finding that those who report stronger endorsement of the belief that emotional intelligence is malleable also score higher on the MSCEIT (Cabello & Fernández-Berrocal, 2015).

From a learning standpoint, an obvious path is to investigate the effects of increasingly popular virtual environments like social media and video games on emotional intelligence development. How has this shift in the way we spend our time affected our abilities to perceive, utilize, understand, and regulate emotions? And also, what are the effects of emotional intelligence on how we interact with technology? Do emotionally intelligent individuals engage in technologies like social media in more adaptive ways, such that they reap the potential benefits of social media (e.g., increased social networks), while buffering themselves against some of the potential harmful effects (e.g., increased loneliness)?

And, finally, from an ethical standpoint, it is important to remember that emotional intelligence is neither inherently prosocial nor sinister – like IQ, it can be harnessed for many purposes, good or bad. Any growth in our knowledge about emotional intelligence over the coming decades should be accompanied by the development of ethical principles that can guide us as we apply the science of emotional intelligence to understanding and managing ourselves and others.

The next thirty years will, of course, witness a host of other major developments not discussed or forecasted in this chapter and we anticipate that there will be a growing emphasis on integrating emotional learning with technology as we increasingly rely on it for our social and emotional experiences.

References

Alshamsi, H., Kepuska, V., & Meng, H. (2017). Real time automated facial expression recognition app development on smart phones. In *Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2017 8th IEEE*

Annual (pp. 384–392). Vancouver: IEEE. https://doi.org/10.1109/IEMCON.2017.8117150

Arsenio, W. F. (2003). Emotional intelligence and the intelligence of emotions: A developmental perspective on mixed EI models. *Human Development*, *46*(2–3), 97–103. https://doi.org/10.1159/000068582

Barbey, A. K., Colom, R., & Grafman, J. (2014). Distributed neural system for emotional intelligence revealed by lesion mapping. *Social Cognitive and Affective Neuroscience*, *9*(3), 265–272. https://doi.org/10.1093/scan/nss124

Barchard, K. A. (2003). Does emotional intelligence assist in the prediction of academic success. *Educational and Psychological Measurement*, *63*(5), 840–858. https://doi.org/10.1177/0013164403251333

Bar-On, R. (1997). *The emotional intelligence inventory (EQ-I): Technical manual*. Toronto: Multi-Health Systems.

Bar-On, R. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicothema*, *18*, 13–25.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263. https://doi.org/10.1111/j.1467–8624.2007.00995.x

Boyatzis, R., Goleman, D., & Rhee, K. (2000). Clustering competence in emotional intelligence: Insights from the Emotional Competence Inventory. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 343–362). San Francisco: Jossey-Bass.

Brackett, M. A., & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin*, *29*(9), 1147–1158. https://doi.org/10.1177/0146167203254596

Brackett, M. A., Palomera, R., Mojsa-Kaja, J., Reyes, M. R., & Salovey, P. (2010). Emotion-regulation ability, burnout, and job satisfaction among British secondary-school teachers. *Psychology in the Schools*, *47*(4), 406–417. https://doi.org/10.1002/pits.20478

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2012). Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learning and Individual Differences*, *22*(2), 218–224. https://doi.org/10.1016/j.lindif.2010.10.002

Brackett, M. A., Rivers, S. E., & Salovey, P. (2011). Emotional intelligence: Implications for personal, social, academic, and workplace success. *Social and Personality Psychology Compass*, *5*(1), 88–103. https://doi.org/10.1111/j.1751–9004.2010.00334.x

Brackett, M. A., Warner, R. M., & Bosco, J. S. (2005). Emotional intelligence and relationship quality among couples. *Personal Relationships*, *12*(2), 197–212. https://doi.org/10.1111/j.1350–4126.2005.00111.x

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*(7), 513–531.

Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., . . . & Trench, M. (2017). Artificial intelligence: The next digital frontier? Discussion paper, June. McKinsey Global Institute. www.mckinsey.com/~/media/McKinsey/Industries/

Advanced Electronics/Our Insights/How artificial intelligence can deliver real value to companies/MGI-Artificial-Intelligence-Discussion-paper.ashx

Cabello, R., & Fernández-Berrocal, P. (2015). Implicit theories and ability emotional intelligence. *Frontiers in Psychology*, *6*, 1–8. https://doi.org/10.3389/fpsyg.2015.00700

Cabello, R., Sorrel, M. A., Fernández-Pinto, I., Extremera, N., & Fernández-Berrocal, P. (2016). Age and gender differences in ability emotional intelligence in adults: A cross-sectional study. *Developmental Psychology*, *52*(9), 1486–1492. https://doi.org/10.1037/dev0000191

Cejudo, J., & Latorre, S. (2015). Effects of the Spock videogame on improving emotional intelligence in adolescents. *Electronic Journal of Research in Educational Psychology*, *13*(2), 319–342. https://doi.org/10.14204/ejrep.36.15060

Cherniss, C. (2010). Emotional intelligence: Toward clarification of a concept. *Industrial and Organizational Psychology*, *3*, 110–126.

Conzelmann, K., Weis, S., & Süß, H.-M. (2013). New findings about social intelligence. *Journal of Individual Differences*, *34*(3), 119–137. https://doi.org/10.1027/1614–0001/a000106

Côté, S., DeCelles, K. A., McCarthy, J. M., van Kleef, G. A., & Hideg, I. (2011). The Jekyll and Hyde of emotional intelligence: Emotion-regulation knowledge facilitates both prosocial and interpersonally deviant behavior. *Psychological Science*, *22*(8), 1073–1080. https://doi.org/10.1177/0956797611416251

Côté, S., Lopes, P. N., Salovey, P., & Miners, C. T. H. (2010). Emotional intelligence and leadership emergence in small groups. *Leadership Quarterly*, *21*(3), 496–508. https://doi.org/10.1016/j.leaqua.2010.03.012

Crombie, D., Lombard, C., & Noakes, T. (2011). Increasing emotional intelligence in cricketers: An intervention study. *International Journal of Sports Science and Coaching*, *6*(1), 69–86. https://doi.org/10.1260/1747-9541.6.1.69

Dacre Pool, L., & Qualter, P. (2012). Improving emotional intelligence and emotional self-efficacy through a teaching intervention for university students. *Learning and Individual Differences*, *22*(3), 306–312. https://doi.org/10.1016/j.lindif.2012.01.010

Davis, S. K., & Humphrey, N. (2012). Emotional intelligence predicts adolescent mental health beyond personality and cognitive ability. *Personality and Individual Differences*, *52*(2), 144–149. https://doi.org/10.1016/j.paid.2011.09.016

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370. https://doi.org/10.1177/1754073911410740

Ermer, E., Kahn, R. E., Salovey, P., & Kiehl, K. A. (2012). Emotional intelligence in incarcerated men with psychopathic traits. *Journal of Personality and Social Psychology*, *103*(1), 194–204. https://doi.org/10.1037/a0027328

ESA (Entertainment Software Association). (2017). Essential facts about the computer and videogame industry. www.theesa.com/wp-content/uploads/2017/06/!EF2017_Design_FinalDigital.pdf

Facebook. (2018). Newsroom, May 18. https://newsroom.fb.com/company-info/

Fan, H., Jackson, T., Yang, X., Tang, W., & Zhang, J. (2010). The factor structure of the Mayer-Salovey-Caruso Emotional Intelligence Test V 2.0 (MSCEIT): A meta-analytic structural equation modeling approach. *Personality and Individual Differences*, *48*(7), 781–785. https://doi.org/10.1016/j.paid.2010.02.004

Fan, L., Scheutz, M., Lohani, M., McCoy, M., & Stokes, C. (2017). Do we need emotionally intelligent artificial agents? First results of human perceptions of emotional intelligence in humans compared to robots. In J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, I. Leite, & S. Kopp (Eds.), *Lecture notes in computer science* (pp. 129–141). Cham: Springer. https://doi.org/10.1007/978-3-319-67401-8_15

Farh, C. I. C. C., Seo, M.-G., & Tesluk, P. E. (2012). Emotional intelligence, teamwork effectiveness, and job performance: The moderating role of job context. *Journal of Applied Psychology*, *97*(4), 890–900. https://doi.org/10.1037/a0027377

Fernández-Berrocal, P., & Extremera, N. (2016). Ability emotional intelligence, depression, and well-being. *Emotion Review*, *8*(4), 311–315. https://doi.org/10.1177/1754073916650494

Fox, N. A. (2003). Not quite ready to invest. *Human Development*, *46*(2–3), 104–108. https://doi.org/10.1159/000068583

García-Sancho, E., Salguero, J. M., & Fernández-Berrocal, P. (2016). Angry rumination as a mediator of the relationship between ability emotional intelligence and various types of aggression. *Personality and Individual Differences*, 89, 143–147. https://doi.org/10.1016/j.paid.2015.10.007

García-Sancho, E., Salguero, J. M., & Fernández-Berrocal, P. (2017). Ability emotional intelligence and its relation to aggression across time and age groups. *Scandinavian Journal of Psychology*, *58*(1), 43–51. https://doi.org/10.1111/sjop.12331

Gardner, K., & Qualter, P. (2009). Emotional intelligence and borderline personality disorder. *Personality and Individual Differences*, *47*(2), 94–98. https://doi.org/10.1016/j.paid.2009.02.004

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion*, *14*(2), 251–262. https://doi.org/10.1037/a0036052

Gentile, D. A., Anderson, C. A., Yukawa, S., Ihori, N., Saleem, M., Ming, L. K., . . . & Sakamoto, A. (2009). The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal, and experimental studies. *Personality and Social Psychology Bulletin*, *35*(6), 752–763. https://doi.org/10.1177/0146167209333045

Gibbs, N. (1995). The EQ factor. *Time*, October 2.

Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. New York: Bantam Books.

Goleman, D. (1998). *Working with emotional intelligence*. New York: Bantam Books.

Goleman, D. (2005). Introduction. In *Emotional Intelligence: Why it Can Matter More than IQ* (10th anniversary ed., pp. ix–xviii).New York: Bantam Books.

Goleman, D., Boyatzis, R., & McKee, A. (2015). *HBR's 10 must reads on emotional intelligence*. Cambridge, MA: Harvard Business Review Press.

Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist*, *69*(1), 66–78. https://doi.org/10.1037/a0034857

Hagelskamp, C., Brackett, M. A., Rivers, S. E., & Salovey, P. (2013). Improving classroom quality with the RULER approach to social and emotional learning: Proximal and distal outcomes. *American Journal of Community Psychology*, *51*(3–4), 530–543. https://doi.org/10.1007/s10464-013–9570–x

Hambrook, D., Brown, G., & Tchanturia, K. (2012). Emotional intelligence in anorexia nervosa: Is anxiety a missing piece of the puzzle? *Psychiatry Research*, *200*(1), 12–19. https://doi.org/10.1016/j.psychres.2012.05.017

Harris, H., & Nass, C. (2011). Emotion regulation for frustrating driving contexts. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems – CHI '11* (pp. 749–752). New York: ACM. https://doi.org/10.1145/1978942.1979050

Harry, B., Williams, M. A., Davis, C., & Kim, J. (2013). Emotional expressions evoke a differential response in the fusiform face area. *Frontiers in Human Neuroscience*, *7*, 1–6. https://doi.org/10.3389/fnhum.2013.00692

Herpertz, S., Schütz, A., & Nezlek, J. (2016). Enhancing emotion perception, a fundamental component of emotional intelligence: Using multiple-group SEM to evaluate a training program. *Personality and Individual Differences*, *95*, 11–19. https://doi.org/10.1016/j.paid.2016.02.015

Hertel, J., Schütz, A., & Lammers, C.-H. (2009). Emotional intelligence and mental disorder. *Journal of Clinical Psychology*, *65*(9), 942–954. https://doi.org/10.1002/jclp

Izard, C. E. (2001). Emotional intelligence or adaptive emotions? *Emotion*, *1*(3), 249–257. https://doi.org/10.1037/1528–3542.1.3.249

Jacobs, M., Snow, J., Geraci, M., Vythilingam, M., Blair, R. J. R., Charney, D. S., . . . & Blair, K. S. (2008). Association between level of emotional intelligence and severity of anxiety in generalized social phobia. *Journal of Anxiety Disorders*, *22*(8), 1487–1495. https://doi.org/10.1016/j.janxdis.2008.03.003

Jaušovec, N., & Jaušovec, K. (2005). Differences in induced gamma and upper alpha oscillations in the human brain related to verbal/performance and emotional intelligence. *International Journal of Psychophysiology*, *56*(3), 223–235. https://doi.org/10.1016/j.ijpsycho.2004.12.005

Jaušovec, N., Jaušovec, K., & Gerlič, I. (2001). Differences in event-related and induced EEG patterns in the theta and alpha frequency bands related to human emotional intelligence. *Neuroscience Letters*, *311*(2), 93–96. https://doi.org/10.1016/S0304–3940 (01)02141–3

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology*, *95*(1), 54–78. https://doi.org/10.1037/a0017286

Kemeny, M. E., Foltz, C., Cavanagh, J. F., Cullen, M., Giese-Davis, J., Jennings, P., . . . Ekman, P. (2012). Contemplative/emotion training reduces negative emotional behavior and promotes prosocial responses. *Emotion*, *12*(2), 338–350. https://doi.org/10.1037/a0026118

Kilduff, M., Chiaburu, D. S., & Menges, J. I. (2010). Strategic use of emotional intelligence in organizational settings: Exploring the dark side. *Research in Organizational Behavior*, *30*, 129–152. https://doi.org/10.1016/j.riob.2010.10.002

Lanciano, T., & Curci, A. (2014). Incremental validity of emotional intelligence ability in predicting academic achievement. *The American Journal of Psychology*, *127*(4), 447–461. https://doi.org/10.5406/amerjpsyc.127.4.0447

Legree, P. J., Psotka, J., Robbins, J., Roberts, R. D., Putka, D. J., & Mullins, H. M. (2014). Profile similarity metrics as an alternate framework to score rating-based tests: MSCEIT reanalyses. *Intelligence*, *47*, 159–174. https://doi.org/10.1016/j.intell.2014.09.005

Lohani, M., Stokes, C., McCoy, M., Bailey, C. A., Joshi, A., & Rivers, S. E. (2016). Perceived role of physiological sensors impacts trust and reliance on robots. In *25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 513–518). New York: IEEE. https://doi.org/10.1109/ROMAN.2016.7745166

Lohani, M., Stokes, C., McCoy, M., Bailey, C. A., & Rivers, S. E. (2016). Social interaction moderates human-robot trust-reliance relationship and improves stress coping. In *ACM/IEEE International Conference on Human-Robot Interaction* (pp. 471–472). Christchurch: IEEE. https://doi.org/10.1109/HRI.2016.7451811

Lopes, P. N., Brackett, M. A., Nezlek, J. B., Schütz, A., Sellin, I., & Salovey, P. (2004). Emotional intelligence and social interaction. *Personality and Social Psychology Bulletin*, *30*(8), 1018–1034. https://doi.org/10.1177/0146167204264762

Lopes, P. N., Grewal, D., Kadis, J., Gall, M., & Salovey, P. (2006). Evidence that emotional intelligence is related to job performance and affect and attitudes at work. *Psicothema*, *18*, 132–138.

Lopes, P. N., Salovey, P., Côté, S., & Beers, M. (2005). Emotion regulation abilities and the quality of social interaction. *Emotion*, *5*(1), 113–118. https://doi.org/10.1037/1528–3542.5.1.113

Lopes, P. N., Salovey, P., & Straus, R. (2003). Emotional intelligence, personality, and the perceived quality of social relationships. *Personality and Individual Differences*, *35*, 641–658.

MacCann, C., Joseph, D. L., Newman, D. A., & Roberts, R. D. (2014). Emotional intelligence is a second-stratum factor of intelligence: Evidence from hierarchical and bifactor models. *Emotion*, *14*(2), 358–374. https://doi.org/10.1037/a0034755

Markus, H. R., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives on Psychological Science*, *5*(4), 420–430. https://doi.org/10.1177/1745691610375557

Martins, A., Ramalho, N., & Morin, E. (2010). A comprehensive meta-analysis of the relationship between emotional intelligence and health. *Personality and Individual Differences*, *49*(6), 554–564. https://doi.org/10.1016/j.paid.2010.05.029

Matsumoto, D., & Hwang, H. S. (2011). Evidence for training the ability to read microexpressions of emotion. *Motivation and Emotion*, *35*(2), 181–191. https://doi.org/10.1007/s11031-011–9212–2

Maul, A. (2011). The factor structure and cross-test convergence of the Mayer-Salovey-Caruso model of emotional intelligence. *Personality and Individual Differences*, *50*(4), 457–463. https://doi.org/10.1016/j.paid.2010.11.007

Mayer, J. D. (2001). Emotion, intelligence, and emotional intelligence. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 410–431). Mahwah, NJ: Lawrence Erlbaum Associates.

Mayer, J. D. (2018). Intelligences about things and intelligences about people. In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 270–286). Cambridge: Cambridge University Press.

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, *27*(4), 267–298. https://doi.org/ISSN:0160–2896

Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The ability model of emotional intelligence: Principles and updates. *Emotion Review*, *8*(4), 290–300. https://doi.org/10.1177/1754073916639667

Mayer, J. D., Caruso, D. R., Salovey, P., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, *3*(1), 97–105.

Mayer, J. D., Panter, A. T., & Caruso, D. R. (2012). Does personal intelligence exist? Evidence from a new ability-based measure. *Journal of Personality Assessment*, *94*(2), 124–140. https://doi.org/10.1080/00223891.2011.646108

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, *59*(1), 507–536. https://doi.org/10.1146/annurev.psych.59.103006.093646

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–31). New York: Basic Books.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2000). Models of emotional intelligence. In R. J. Sternberg (Ed.), *Handbook of Intelligence* (pp. 396–420). Cambridge: Cambridge University Press.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist*, *63*(6), 503–517. https://doi.org/10.1037/0003-066X.63.6.503

Mayer, J. D., Salovey, P., & Caruso, D. R. (2014). *Test manual for the Mayer-Salovey-Caruso Emotional Intelligence Test – Youth Research Version*. Toronto: Multi-Health Systems.

Mayer, J. D., Salovey, P., Caruso, D. R., & Cherkasskiy, L. (2011). Emotional intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence*. New York: Cambridge University Press.

Multi-Health Systems. (2011). EQ-i launch kit. Toronto. http://downloads.mhs.com/eqi/EQi-Launch-Kit.pdf

O'Boyle, E. H., Humphrey, R. H., Pollack, J. M., Hawver, T. H., & Story, P. A. (2011). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior*, *32*(5), 788–818. https://doi.org/10.1002/job.714

O'Connor, R. M., & Little, I. S. (2003). Revisiting the predictive validity of emotional intelligence: Self-report versus ability-based measures. *Personality and Individual Differences*, *35*(8), 1893–1902. https://doi.org/10.1016/S0191–8869(03)00038–2

Operskalski, J. T., Paul, E. J., Colom, R., Barbey, A. K., & Grafman, J. (2015). Lesion mapping the four-factor structure of emotional intelligence. *Frontiers in Human Neuroscience*, *9*, 649. https://doi.org/10.3389/fnhum.2015.00649

Palmer, B. R., Gignac, G., Manocha, R., & Stough, C. (2005). A psychometric evaluation of the Mayer-Salovey-Caruso Emotional Intelligence Test Version 2.0. *Intelligence*, *33*(3), 285–305. https://doi.org/10.1016/j.intell.2004.11.003

Petrides, K. V., & Furnham, A. (2001). Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. *European Journal of Personality*, *15*(6), 425–448. https://doi.org/10.1002/per.416

Petrides, K. V., & Furnham, A. (2003). Trait emotional intelligence: Behavioural validation in two studies of emotional recognition and reactivity to mood induction. *European Journal of Personality*, *17*, 39–57.

Petrides, K. V., Mikolajczak, M., Mavroveli, S., Sanchez-Ruiz, M.-J., Furnham, A., & Pérez-González, J.-C. (2016). Developments in trait emotional intelligence research. *Emotion Review*, *8*(4), 335–341. https://doi.org/10.1177/1754073916650493

Pisner, D. A., Smith, R., Alkozei, A., Klimova, A., & Killgore, W. D. S. (2017). Highways of the emotional intellect: White matter microstructural correlates of an ability-based measure of emotional intelligence. *Social Neuroscience*, *12*(3), 253–267. https://doi.org/10.1080/17470919.2016.1176600

Qualter, P., Gardner, K. J., Pope, D. J., Hutchinson, J. M., & Whiteley, H. E. (2012). Ability emotional intelligence, trait emotional intelligence, and academic success in British secondary schools: A 5 year longitudinal study. *Learning and Individual Differences*, *22*(1), 83–91. https://doi.org/10.1016/j.lindif.2011.11.007

Reis, D. L., Brackett, M. A., Shamosh, N. A., Kiehl, K. A., Salovey, P., & Gray, J. R. (2007). Emotional Intelligence predicts individual differences in social exchange reasoning. *NeuroImage*, *35*(3), 1385–1391. https://doi.org/10.1016/j.neuroimage.2006.12.045

Reuben, E., Sapienza, P., & Zingales, L. (2009). *Can we teach emotional intelligence?* Unpublished Manuscript, Columbia Business School, New York.

Rideout, V. (2015). *The common sense consensus: Media use by tweens and teens*. San Francisco: Common Sense. www.commonsensemedia.org/sites/default/files/uploads/research/census_researchreport.pdf

Rivers, S. E. (2018). Thriving through gameplay. *Well Played*, *7*(2), iv–xvi. http://press.etc.cmu.edu/index.php/product/well-played-vol-7-no-2/

Rivers, S. E., Brackett, M. A., Reyes, M. R., Elbertson, N. A., & Salovey, P. (2013). Improving the social and emotional climate of classrooms: A clustered randomized controlled trial testing the RULER approach. *Prevention Science*, *14*(1), 77–87. https://doi.org/10.1007/s11121-012–0305–2

Rivers, S. E., Brackett, M. A., Reyes, M. R., Mayer, J. D., Caruso, D. R., & Salovey, P. (2012). Measuring emotional intelligence in early adolescence with the MSCEIT-YV: Psychometric properties and relationship with academic performance and psychosocial functioning. *Journal of Psychoeducational Assessment*, *30*(4), 344–366. https://doi.org/10.1177/0734282912449443

Roberts, R. D., Schulze, R., O'Brien, K., MacCann, C., Reid, J., & Maul, A. (2006). Exploring the validity of the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) with established emotions measures. *Emotion*, *6*(4), 663–669. https://doi.org/10.1037/1528–3542.6.4.663

Rosete, D., & Ciarrochi, J. (2005). Emotional intelligence and its relationship to workplace performance outcomes of leadership effectiveness. *Leadership and Organization Development Journal*, *26*(5), 388–399. https://doi.org/10.1108/01437730510607871

Rossen, E., Kranzler, J. H., & Algina, J. (2008). Confirmatory factor analysis of the Mayer-Salovey-Caruso Emotional Intelligence Test V 2.0 (MSCEIT). *Personality and Individual Differences*, *44*(5), 1258–1269. https://doi.org/10.1016/j.paid.2007.11.020

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, *9*(3), 185–211. https://doi.org/10.1016/S0962–1849(05)80058–7

Sánchez-Álvarez, N., Extremera, N., & Fernández-Berrocal, P. (2016). The relation between emotional intelligence and subjective well-being: A meta-analytic investigation. *Journal of Positive Psychology*, *11*(3), 276–285. https://doi.org/10.1080/17439760.2015.1058968

Schoneveld, E. A., Lichtwarck-Aschoff, A., & Granic, I. (2018). Preventing childhood anxiety disorders: Is an applied game as effective as a cognitive behavioral therapy-based program? *Prevention Science*, *19*(2), 220–232. https://doi.org/10.1007/s11121-017–0843–8

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, *25*, 167–177.

Schweizer, S., Grahn, J., Hampshire, A., Mobbs, D., & Dalgleish, T. (2013). Training the emotional brain: Improving affective control through emotional working memory training. *The Journal of Neuroscience*, *33*(12), 5301–5311. https://doi.org/10.1523/JNEUROSCI.2593–12.2013

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, *55*(1), 5–14. https://doi.org/10.1037/0003-066X.55.1.5

Shao, B., Doucet, L., & Caruso, D. R. (2015). Universality versus cultural specificity of three emotion domains: Some evidence based on the cascading model of emotional intelligence. *Journal of Cross-Cultural Psychology*, *46*(2), 229–251. https://doi.org/10.1177/0022022114557479

Sternberg, R. J., & Kaufman, S. B. (Eds.). (2011). *The Cambridge handbook of intelligence*. Cambridge: Cambridge University Press.

Sundar, S. S., & Nass, C. (2000). Source orientation in human-computer interaction. *Communication Research*, *27*(6), 683–703. https://doi.org/10.1177/009365000027006001

Teper, R., Segal, Z. V., & Inzlicht, M. (2013). Inside the mindful mind: How mindfulness enhances emotion regulation through improvements in executive control. *Current Directions in Psychological Science*, *22*(6), 449–454. https://doi.org/10.1177/0963721413495869

Tett, R. P., Fox, K. E., & Wang, A. (2005). Development and validation of a self-report measure of emotional intelligence as a multidimensional trait domain. *Personality and Social Psychology Bulletin*, *31*(7), 859–888. https://doi.org/10.1177/0146167204272860

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Torrente, C., Rivers, S. E., & Brackett, M. A. (2016). Teaching emotional intelligence in schools: An evidence-based approach. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *Psychosocial skills and school systems in the 21st century: Theory, research, and practice* (pp. 325–346). Cham: Springer. https://doi.org/10.1007/978-3-319-28606-8_13

Villani, D., Carissoli, C., Triberti, S., Marchetti, A., Gilli, G., & Riva, G. (2018). Videogames for emotion regulation: A systematic review. *Games for Health Journal*, *7*(2). https://doi.org/10.1089/g4h.2017.0108

Von Der Pütten, A. M., Krämer, N. C., Gratch, J., & Kang, S. H. (2010). "It doesn't matter what you are!": Explaining social effects of agents and avatars. *Computers in Human Behavior*, *26*(6), 1641–1650. https://doi.org/10.1016/j.chb.2010.06.012

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits. The role of tacit knowledge. *Journal of Personality and Social Psychology*, *49*(2), 436–458. https://doi.org/10.1037/0022–3514.49.2.436

Yeager, D. S., Miu, A. S., Powers, J., & Dweck, C. S. (2013). Implicit theories of personality and attributions of hostile intent: A meta-analysis, an experiment, and a longitudinal intervention. *Child Development*, *84*(5), 1651–1667. https://doi.org/10.1111/cdev.12062

Yip, J. A., & Martin, R. A. (2006). Sense of humor, emotional intelligence, and social competence. *Journal of Research in Personality*, *40*(6), 1202–1208. https://doi.org/10.1016/j.jrp.2005.08.005

Zeidner, M., Matthews, G., Roberts, R. D., & MacCann, C. (2003). Development of emotional intelligence: Towards a multi-level investment model. *Human Development*, *46*(23), 69–96. https://doi.org/10.1159/000068580

30 Practical Intelligence

Jennifer Hedlund

Richard Wagner concluded his chapter on practical intelligence in the previous edition of the *Cambridge Handbook of Intelligence* with the prophetic question as to whether there would be a separate chapter on the topic in the next edition. He suggested that many of the key ideas associated with practical intelligence were converging with other theoretical perspectives, such as life-span development, expertise, and embodied cognition (Wagner, 2011). Although efforts to understand practical intelligence have occasionally intersected with other bodies of knowledge, and the growth of practical intelligence research has somewhat slowed in the period since the last edition, practical intelligence continues to have a distinct presence in the literature. Perhaps the most prolific area of research related to practical intelligence has been the study of tacit knowledge. Although researchers continue to seek further understanding of the content of tacit knowledge and its relation to successful performance, there has been a growing emphasis on identifying ways to facilitate tacit knowledge acquisition.

This chapter will provide an overview of research on practical intelligence, with particular attention to the work of Sternberg and his colleagues, who have been most associated with the concept, and illustrate how research on practical intelligence has evolved over time. A substantial portion of this chapter will focus on the role of tacit knowledge in understanding practical intelligence. The chapter concludes with a discussion of potential ways to facilitate tacit knowledge acquisition and develop practical intelligence.

Defining Practical Intelligence

The concept of practical intelligence appeared in the literature as early as the 1940s in relation to the use of situational judgment tests (SJTs) to assess managerial potential (McDaniel & Whetzel, 2005). In-basket tests, used to assess managerial potential since the 1950s, have also been associated with the concept of practical intelligence (Frederiksen, 1966, 1986). Specific attention to the concept of practical intelligence, however, emerged in the 1980s in relation to broader conceptualizations of intelligence and the competencies needed to perform everyday tasks. Robert Sternberg brought the concept of practical intelligence to the forefront with his triarchic theory of intelligence (Sternberg, 1985, 1988) and subsequent theory of successful intelligence (Sternberg, 1997).

Practical intelligence is defined as the ability that individuals use to find a more optimal fit between themselves and the demands of the environment through adapting, shaping, or selecting a new environment in the pursuit of personally valued goals (Sternberg, 1985, 1997). It has been characterized as "street smarts" or "common sense" and can be contrasted with analytical intelligence or "book smarts." Baum, Bird, and Singh (2011) described practical intelligence as an ability complex that "overlaps concepts related to expertise, decision making, and judgment" (p. 398). Yalon-Chamovitz and Greenspan (2005) referred to practical intelligence as the "cognitive underpinning of everyday function" (p. 220). Practical intelligence is best understood in relation to the types of problems individuals encounter in everyday life.

Understanding Practical Problems

Everyone encounters problems for which solutions are neither readily available nor readily derivable from acquired knowledge. These problems occur in the workplace, in school, at home, or really anywhere. Sternberg and his colleagues (Sternberg, 1985, 1997; Wagner & Sternberg, 1986), building on a distinction made by Neisser (1976), differentiated practical problems from academic problems. Academic problems tend to be well-defined, formulated by others, complete in the information they provide, removed from ordinary experience, of little or no intrinsic interest, and have only one correct answer and one method of obtaining the correct answer.

Practical problems, in contrast, tend to be poorly defined, unformulated or in need of reformulation, lacking in information necessary for solution, of personal interest, related to everyday experience, and to present multiple "correct" solutions, each with liabilities as well as assets, along with multiple methods for picking a problem solution. People who are adept at solving one kind of problem may not be adept at solving problems of the other kind (Sternberg, 1985, 1997; Wagner & Sternberg, 1986). Initially, the examination of practical intelligence issued from a concern that the intelligence of adults functioning largely outside the academic environment was evaluated primarily by traditional tests of intelligence originally developed to predict academic success (for a review, see Berg, 2000).

Research on Practical Problem-Solving

Research conducted in a wide range of settings and with diverse populations has demonstrated the distinction between academic and practical problem-solving (for more extensive reviews, see Ceci & Roazzi, 1994; Rogoff & Lave, 1984; Sternberg & Wagner, 1986).

In research with schoolchildren, several investigators have found that performance on academic tasks does not necessarily correspond to performance on real-world tasks, even when the same processes are involved (e.g., mathematical reasoning). Researchers studying children in Brazil who worked as street vendors found that they performed significantly better on mathematical problems presented in the context of vending than on those problems presented in an academic context, and were able to complete monetary transactions without the aid of pencil and paper or calculators (Carraher, Carraher, & Schliemann, 1985; Nunes, Schliemann, & Carraher, 1993). In contrast, Perret-Clermont (1980) found that schoolchildren who showed no difficulty in solving arithmetic problems on a paper-and-pencil test could not solve the same problems in an everyday context (e.g., counting bunches of flowers).

Grigorenko and colleagues (2004) developed and administered a measure of practical intelligence, called the Yup'ik Scale of Practical Intelligence (YSPI), to 261 adolescents from a rural and a semi-urban Yup'ik Alaskan community. Children from the semi-urban community scored better on a measure of crystallized intelligence while those from the rural setting scored better on the measure of practical intelligence. Additionally, in the rural community sample, the YSPI was the best predictor of peer and adult ratings on several traits, including thinking skills, respect for elders, hunting skills, and household skills.

Practical abilities also are relevant to effective performance in school. Mandelman, Barbot, and Grigorenko (2016) administered the Aurora Battery, designed to assess analytical, practical, and creative intelligence, to a sample of 145 middle school students. Measures of practical intelligence, which included tasks such as tracing the best carpool route between friends' houses and dividing monetary amounts among friends, were the best predictors of grade point average (GPA).

Research with adults has produced similar results. In a series of studies by Ceci and his colleagues (Ceci & Liker, 1986, 1988; Ceci & Ruiz, 1992; Ceci & Roazzi, 1994), racetrack handicappers were observed to use complex algorithms to predict post time odds. The successful use of these algorithms was found to be unrelated to IQ scores or use of the same algorithms in a different context (e.g., stock market prediction). In research with assemblers at a milk processing plant, Scribner (1984, 1986) found that the use of complex strategies to fill orders in a manner that minimized the number of moves required to complete the order was unrelated to the assemblers' IQ scores, arithmetic test scores, or grades in school. Similarly, grocery shoppers' accuracy at identifying the best values, by performing calculations of the cost per unit, was unrelated to their scores on a mathematics test (Lave, Murtaugh, & de la Roche, 1984; Murtaugh, 1985).

Research also demonstrates the potential benefit of practical intelligence to daily life activities. Grigorenko and Sternberg (2001) examined the relative influence of analytical, creative, and practical intelligence on adaptive functioning in a sample of 293 men and 452 women from a large industrial city in Central Russia. Practical intelligence was the most consistent and strongest predictor of both physical and mental health. Specifically, individuals with higher practical intelligence scores reported better physical health, lower anxiety and depression, and higher self-efficacy than individuals with lower practical intelligence scores.

Yalon-Chamovitz and Greenspan (2005) studied the relevance of practical intelligence in individuals with intellectual disabilities. They developed a video-based measure of practical intelligence that involved scenes of instrumental activities of daily living such as cooking and doing laundry. They embedded three practical problems or errors in the scenarios that might occur while completing the task in each video. Participants were scored based on when they noticed a problem, their explanation of the problem, and their recommend solution. In addition, a guardian or case manager rated each participant on their ability to perform activities of daily living . The authors suggested that the video measure of practical intelligence provided a more direct assessment of cognition than caregiver ratings, which are based solely on outward behavior. They found that scores on the video assessment correlated significantly with ratings of activities of daily living. Further, practical intelligence scores were highly correlated with experience, suggesting that practical abilities can be developed by increasing opportunities to participate in everyday activities.

The above studies indicate that demonstrated abilities do not necessarily correspond between everyday tasks (e.g., price-comparison shopping) and traditional academic tasks (e.g., math achievement tests). Few of these researchers would dispute the claim that intelligence as conventionally defined and measured predicts performance both in and outside of school. However, there is wide recognition that other aspects of intelligence may be equally, if not more important, than *g* to the performance of tasks both within and outside academic settings. The greatest support for the relevance of practical intelligence to everyday performance comes from the extensive body of research on tacit knowledge.

Practical Intelligence and Tacit Knowledge

In solving practical problems, individuals draw on a broad base of knowledge, some of which is acquired through formal training and some of which is derived from personal experience. Much of the knowledge associated with successful problem-solving can be characterized as tacit because it may not be openly expressed or stated. Although people's actions may reflect their knowledge, they may find it difficult to articulate what they know. Research on expert knowledge indicates that experts draw on a well-developed repertoire of knowledge in responding to problems in their respective domains (Scribner, 1986), that the knowledge tends to be procedural in nature and to operate outside of focal awareness (Chi, Glaser, & Farr, 1988), and that it reflects the structure of the situation more closely than it does the structure of formal, disciplinary knowledge (Groen & Patel, 1988).

The term "tacit knowledge" has roots in works on the philosophy of science (Polanyi, 1966), ecological psychology (Neisser, 1976), and organizational behavior (Schön, 1983) and has been used to characterize the knowledge gained from everyday experience that has an implicit, unarticulated quality. Such notions about the tacit quality of the knowledge associated with everyday problem-solving also are reflected in the common language of the workplace as people attribute successful performance to "learning by doing" and to "professional intuition" or "instinct."

Sternberg and his colleagues (Sternberg, 1997; Sternberg et al., 2000) view tacit knowledge as an aspect of practical intelligence that enables individuals to adapt to, select, and shape real-world environments. It is knowledge that reflects the practical ability to learn from experience and to apply that knowledge in pursuit of personally valued goals. Baum, Bird, and Singh (2011) suggest that tacit knowledge represents "knowing" while practical intelligence represents both "knowing and doing." Prior to discussing research on tacit knowledge, it is important to understand how it has been conceptualized and operationalized in the literature.

Conceptualizing Tacit Knowledge

Tacit knowledge has been defined by Sternberg and his colleagues (Sternberg, 1997; Sternberg et al., 1995, 2000) according to three main features that correspond to the conditions under which tacit knowledge is acquired, its structural representation, and the conditions of its use. First, tacit knowledge generally is acquired on one's own with little support from other people or resources. According to Sternberg (1988), the acquisition of tacit knowledge is facilitated by the cognitive processes of selective encoding (sorting relevant from irrelevant information in the environment), selective combination (integrating information into a meaningful interpretation of the situation), and selective comparison (relating new information to existing knowledge). When these processes are not well supported, as often is the case in learning from everyday experiences, the likelihood increases that some individuals will fail to acquire the knowledge.

Second, tacit knowledge is procedural in nature. It is knowledge about how to act in particular situations or classes of situations. Drawing on Anderson's (1983) distinction between procedural and declarative knowledge, tacit knowledge can be viewed as a subset of procedural knowledge that is drawn from personal experience and that guides action without being easily articulated. The third characteristic feature of tacit knowledge is that it has practical value to the individual. Knowledge that is experience-based and action-oriented will likely be more instrumental to achieving one's goals than will be knowledge that is based on someone else's experience or that does not specify action.

Measuring Tacit Knowledge

Because people often find it difficult to articulate their tacit knowledge, researchers rely on observable indicators of its existence. Tacit knowledge typically is measured in the responses individuals provide to practical situations or problems. The format of most tacit knowledge inventories (TKIs) closely resembles SJTs (Chan & Schmitt, 1998; Legree, 1995; Motowidlo, Dunnette, & Carter, 1990). These types of tests generally are used to measure interpersonal and problem-solving skills (Hanson & Ramos, 1996; Motowidlo et al., 1990) or behavioral intentions (Weekley & Jones, 1997). In a SJT or TKI, each question presents a problem relevant to the domain of interest (e.g., a manager intervening in a dispute between two subordinates) followed by a set of options (i.e., strategies) for solving the problem (e.g., meet with the two subordinates individually to find out their perspective on the problem; hold a meeting with both subordinates and have them air their grievances). Respondents are asked either to choose the best and worst alternatives from among a few options or to rate on a Likert scale the quality or appropriateness of several potential responses to the situation. Their responses are then scored relative to an expert or consensus mean.

There has been some debate about whether TKIs are measures of practical intelligence. McDaniel and Whetzel (2005) criticized the use of TKIs as measures of practical intelligence since SJTs generally have been found to be multidimensional in nature. Specifically, they argue that "no factor analysis of any situational judgment test has provided evidence of a general factor of practical intelligence, whether correlated with *g* or not" (McDaniel & Whetzel, 2005, p. 519). SJTs tend to exhibit strong correlations with measures of *g* and personality, indicating that they may assess multiple factors related to successful performance. They further argue that the incremental validity of TKIs and other SJTs over *g* in predicting job performance is to be expected given that they measure "non-cognitive job-related constructs" (p. 523).

Stemler and Sternberg (2006) suggest that practical intelligence offers a potential explanation for the consistent incremental validity of SJTs in predicting performance and may address the lack of a shared theoretical framework surrounding the development and use of SJTs. They argue that practical intelligence is best assessed by measuring both the cognitive (i.e., the knowledge, both explicit and tacit, that underlie actions) and the behavioral (i.e., the actions themselves) elements. However, behavioral assessments are more challenging in terms of time and resources while cognitive assessments (e.g., TKIs) can be administered more efficiently.

To address some of the criticism of TKIs, Cianciolo and colleagues (2006) tested the underlying factor structure of three different TKIs. The College Life Questionnaire presented everyday scenarios that undergraduate students might encounter such as dealing with roommates or paying tuition bills. The Common Sense Questionnaire consisted of common situations that entry- to mid-level employees might deal with on the job. Lastly, the Everyday Situational Judgment Inventory consisted of live-action situations that a typical young American might face. In two studies with college students, they established that each TKI was represented by a single underlying factor structure. Additionally, the covariance among the latent tacit knowledge and practical problem-solving factors was accounted for by a single factor, which was consistent with practical intelligence. They did observe some overlap in measures of practical and general intelligence, with practical intelligence exhibiting a factor loading of 0.48 on *g*, but argue that the evidence suggests that practical intelligence is not the same as general intelligence.

Research on Tacit Knowledge

Tacit knowledge has been studied in domains as diverse as sales, primary education, college admissions, military leadership, information technology, and policing and has been related to a variety of performance indicators, including supervisor ratings, grades in school, economic success, and innovation. Researchers also have examined the relationship between tacit knowledge and other relevant constructs such as experience, general intelligence, personality, and learning styles. These relationships are reviewed in more detail below.

Tacit Knowledge and Performance

Tacit knowledge tests typically correlate in the range of 0.2 to 0.5 with various performance criteria. Research with managers has found tacit knowledge to correlate significantly with salary, whether or not the manager worked for a company at the top of the Fortune 500 list, and ratings on the ability to generate new business (Wagner, 1987; Wagner & Sternberg, 1985). In a study with business executives attending a leadership development program, Wagner and Sternberg (1990) found that tacit knowledge scores explained 32 percent of the variance in performance beyond scores on a traditional IQ test and explained variance beyond measures of personality and cognitive style.

Colonia-Willner (1998, 1999) administered the Tacit Knowledge Inventory for Managers (TKIM; Wagner & Sternberg, 1991) to 200 bank managers and found that tacit knowledge scores significantly predicted managerial skill, whereas psychometric and verbal reasoning did not. Baczyńska (2015) administered the TKIM along with measures of analytical and emotional intelligence to ninety-eight line managers in Poland who were participating in a one-day assessment center. Tacit knowledge was the only significant predictor of each of five ratings of managerial competency, accounting for 21 percent to 38 percent of the total variance in those ratings.

In research with military leaders, Hedlund and colleagues (2003) developed TKIs for three levels of military leadership: platoon leaders, company commanders, and battalion commanders. The Tacit Knowledge for Military Leadership (TKML) inventories were administered to 368 platoon leaders, 163 company commanders, and 31 battalion commanders along with the TKIM and a measure of verbal ability. Scores on the TKML correlated with either peer or supervisor ratings of leadership effectiveness at all three levels. TKML scores also accounted for small but significant incremental validity in leadership effectiveness beyond verbal ability and managerial tacit knowledge.

In research with salespeople, Wagner and colleagues (1999) found that tacit knowledge correlated with sales volume and sales awards received. Additionally, Sujan, Sujan, and Bettman (1991) found that more effective salespeople used more domain-specific and problem-oriented strategies while less effective salespeople used more global and relationship-oriented strategies. Jisr and Maamari (2017) found that tacit knowledge correlated significantly with innovation performance in a sample of 331 service industry professionals from twenty different companies in Lebanon.

Taylor and colleagues (2013) created the Police Officer Tacit Knowledge Inventory (POTKI) to measure the practical abilities of police officers. They administered the POTKI along with the Common Sense Questionnaire to twenty-two novice and forty-eight experienced police officers. Respondents were scored based on their distance from the mean expert ratings across response options, as well as their ratings on items designated as "better" or "worse" choices from among the options. They found that tacit knowledge scores correlated significantly with common sense scores and greater agreement with the "better" options were associated with higher supervisor ratings.

In a subsequent study, Taylor, Van Der Heijden, and Genuchi (2017) collected additional data from police applicants. In general, they found more agreement among experts on tacit knowledge items designated as "better" options than items considered "worse." The response patterns also exhibited significant differences between expert and novice police officers with questions addressing intrapersonal knowledge differentiating among experts and novices more effectively than questions about interpersonal tacit knowledge.

Cianciolo and colleagues (2006) administered the Common Sense Questionnaire to samples in the United States and Spain. Scores on the Common Sense Questionnaire correlated significantly with supervisor ratings in the United States but not in Spain, suggesting some cultural differences in the relevance of tacit knowledge to performance. However, there were high correlations between Spanish and US ratings of the individual items, suggesting similar preferences for response options across samples.

Tacit knowledge has been shown to relate to performance in several studies with educators, ranging from elementary school to college. In research with academic psychologists, tacit knowledge scores correlated with citation rate, number of publications, and quality of department (Wagner, 1987; Wagner & Sternberg, 1985). At the elementary school level, Grigorenko, Sternberg, and Strauss (2006) found that higher tacit knowledge scores were associated with higher principal ratings. In a US sample, teachers with higher tacit knowledge scores rated themselves as less effective, while teachers in Israeli with higher tacit knowledge scores rated themselves as more effective. These differences may reflect differences in self-efficacy between the US and Israeli samples. Grigorenko and colleagues (2006), however, found that ratings on the tacit knowledge items were highly correlated between the US and Israeli samples (r = 0.59), suggesting that the knowledge assessed by the TKI is generalizable across cultures.

Although tacit knowledge research typically has focused on problems outside the classroom, there is also evidence that tacit knowledge has relevance to academic performance. Scores on a TKI for college students correlate with indices of academic performance and adjustment to college (Sternberg, Wagner, & Okagaki, 1993). Sternberg and The Rainbow Project Collaborators (2006) found moderate correlations between several measures of tacit knowledge and both high school and college GPA. However, the measures of practical intelligence did not account for additional variance in grades once measures of analytical and creative were included. Insch, McIntyre, and Dawley (2008) found that scores on all six dimensions of an Academic Tacit Knowledge Scale exhibited significant relationships with GPA in a sample of undergraduate business students. Fox and Spector (2000) found that practical intelligence significantly predicted evaluations of undergraduate students' performance on a simulated interview. Razali and Trevelyan (2012) found that the practical intelligence of engineering students correlated significantly with their performance on a fault diagnosis test. Together these studies consistently show that individuals with higher tacit knowledge perform significantly better in their respective performance domains.

Tacit Knowledge and Experience

The common phrase "experience is the best teacher" reflects the view that experience provides opportunities to develop important knowledge and skills related to performance. Research comparing novices and experts has consistently found differences in the amount and pattern of tacit knowledge as a function of expertise. Findings regarding the relationship between the amount of time on the job and tacit knowledge have been less consistent.

Wagner and Sternberg (1985) found a significant correlation between tacit knowledge and a manager's level within the company. Similarly, Wagner (1987) found differences in tacit knowledge scores among business managers, business graduate students, and general undergraduates, with the managers exhibiting the highest scores. These differences in tacit knowledge also were observed when comparing psychology professors, psychology graduate students, and undergraduates. In research with salespeople, Wagner and colleagues (1999) found that scores on a TKI for salespeople correlated significantly with number of years of sales experience.

A study by Baum, Bird, and Singh (2011) found that both venture and industry experience were significantly related to practical intelligence in a sample of 283 entrepreneurs in the printing industry. Joseph and colleagues (2010) compared sixty-eight IT professionals and fifty-four IT undergraduates on the SoftSkills for IT (SSIT). Experienced IT professionals generated significantly more responses, took significantly less time to respond, and provided significantly higher quality responses than the novices. Additionally, Taylor, Van Der Heijden, and Genuchi (2017) found significant differences between expert and novice police officers in their response patterns on a TKI.

Several studies have found that the relationship between tacit knowledge and performance varies based on employee rank. Hedlund and colleagues (2003) found no significant relationships between months in the current position and tacit knowledge scores in a sample of military leaders. However, the relationship between tacit knowledge and performance was strongest at the highest level of leadership. Similarly, Taylor, Psotka, and Legree (2015) found the strongest relationships between leadership style and tacit knowledge scores at the highest level of military command, suggesting that leaders at the highest level have developed the broadest range of responses to a variety of situations. Additionally, Tan and Libby (1997) found that the level of tacit knowledge for auditing distinguished top and bottom performers at the higher rank but not at the lower ranks of auditors employed by a major accounting firm in Singapore.

Some studies have focused specifically on better understanding the relationship between experience and tacit knowledge. Armstrong and Mahmud (2008) studied 356 Malaysian public sector managers and found that novices had significantly lower tacit knowledge scores than successful managers but there were no significant differences between novices and typical managers. Individuals primarily exposed to managerial functions had significantly higher tacit knowledge than those who were primarily responsible for other tasks (e.g., engineering, accounting). No significant relationship, however, was found between length of experience and tacit knowledge scores.

Finally, Elliott and colleagues (2011) studied 501 student teachers in England and compared their responses on the Tacit Knowledge Inventory – High School to a sample of 163 experienced teachers. They found no significant differences between experienced and novice teachers on ratings of "good" tacit knowledge items but experienced teachers were significantly more likely to recognize "bad" responses. Additionally, student teachers improved significantly in their ability to identify "bad" responses but there was no similar improvement in the identification of "good" responses.

Tacit Knowledge and General Cognitive Ability

Although there is some debate as to whether TKIs measure a factor that is distinct from general intelligence (Gottfredson, 2003), most of the research indicates that the correlations between tacit knowledge and conventional intelligence tests are trivial to moderate at best. Scores on TKIs exhibit nonsignificant correlations with a test of verbal reasoning in undergraduate samples (Wagner, 1987; Wagner & Sternberg, 1985), an IQ test for a sample of business executives (Wagner & Sternberg, 1990), and a test of general intelligence in a sample of salespeople (Wagner et al., 1999). In research with military leaders, tacit knowledge scores exhibited nonsignificant correlations with a measure of verbal ability in samples of platoon leaders and battalion commanders but a small, significant correlation in a sample of company commanders (Hedlund et al., 2003). In a study of Air Force recruits, Eddy (1988) found that scores on a TKI for managers were unrelated to scores on the Armed Services Vocational Aptitude Battery (ASVAB).

Some researchers have found moderate correlations in the 0.2 to 0.3 range between tacit knowledge scores and measure of general intelligence (Cianciolo et al., 2006; Fox & Spector, 2000; Sternberg & The Rainbow Project Collaborators, 2006). Sternberg and colleagues (2001), however, obtained negative correlations between scores on a TKI for natural herbal medicines and tests of general and crystallized intelligence in a sample of rural Kenyan children.

Although there is some evidence that measures of practical and general intelligence share some common variance, there also is evidence that TKIs provide incremental validity beyond measures of *g* in explaining individual differences in performance (Hedlund et al., 2003; Wagner & Sternberg, 1990).

Tacit Knowledge and Other Constructs

In addition to experience and general cognitive ability, researchers have explored the relationship of tacit knowledge to several other factors, including personality, learning orientation, gender, and ethnicity.

In regard to personality, Wagner and Sternberg (1990) found that tacit knowledge scores, with two exceptions, exhibited nonsignificant correlations with several personality dimensions in a sample of business executives. Additionally, tacit knowledge scores consistently accounted for a significant increment in variance beyond the personality measures.

In regard to learning orientation, Armstrong and Mahmud (2008) found that individuals with an accommodating learning style and who worked in a managerial context had the highest tacit knowledge scores. Additionally, individuals who fell in the upper quartile of all four learning orientations (accommodating, assimilating, convergent, divergent) had significantly higher tacit knowledge scores than those who fell in the lowest quartile of participants. Baum and colleagues' (2011) study of entrepreneurs in the printing industry examined the influence of four learning orientations (concrete experience, active experimentation, abstract conceptualization, reflective observation) on practical intelligence. They found that entrepreneurs with a concrete experience or active experimentation orientation exhibited higher practical intelligence.

Finally, traditional intelligence tests often are found to exhibit group differences in scores as a function of gender and race (for reviews, see Loehlin, 2000; Neisser et al., 1996). TKIs, because they are not restricted to abilities developed in school, may be less susceptible to these differences. In Eddy's (1988) study of Air Force recruits, comparable levels of performance on the TKI were found among majority and minority group members and among males and females. The same was not true for scores on the ASVAB. Sternberg and The Rainbow Project Collaborators (2006) found that ethnic group differences on practical intelligence measures were significant but of a smaller magnitude than those observed for the SAT. Additionally, Hedlund and colleagues (2006), whose work is discussed in more detail in the following section, found that practical intelligence measures exhibited less disparity across gender and racial/ethnic groups than did the Graduate Management Admissions Test (GMAT).

Alternative Measures of Practical Intelligence

TKIs have emerged as the predominant method for studying practical intelligence in part because they allow individuals to be assessed across numerous problem situations and they are relatively easy to score. A limitation of TKIs is that they assume that individuals have had the opportunities to acquire domain-specific knowledge. In many cases, organizations must evaluate individuals who have not necessarily had the same opportunities. Alternative methods that measure the potential to adapt and to learn on the job may provide broader applicability.

One approach to assessing practical intelligence focuses on the skills involved in solving practical problems. According to Sternberg (1985, 1997), individuals who effectively solve practical problems are able to recognize that a problem exists, define the problem clearly, allocate appropriate resources to the problem, formulate strategies for solving the problem, monitor their solutions, and evaluate the outcomes of those solutions. Furthermore, in order to understand the problem in the first place, individuals need to be able to filter relevant information from irrelevant information, relate new information to existing knowledge, and compile information into a meaningful picture.

An individual's potential to acquire tacit knowledge can be evaluated by measuring how well the individual defines the problem, decides what information to attend to and how to interpret the information, generates and evaluates alternative possible solutions, and chooses and monitor a course of action (Hedlund & Sternberg, 2001; Matthew & Sternberg, 2009). These processes can be measured using a format similar to case studies or in-basket tests. Individuals are presented with sufficient information to solve each problem so as not to rely heavily on existing knowledge and they are asked to respond to a set of question prompts directly targeting the use of problem-solving and knowledge-acquisition skills. Respondents can be evaluated based on how well they exhibit each of the problem-solving skills as well as the quality of their solution.

This alternative approach to measuring practical intelligence was explored within the context of business school admissions. Hedlund and colleagues (2006) developed two approaches to measuring practical intelligence, one knowledge-based and the other skill-based. The situational judgment problems (SJPs) were akin to TKIs and assessed students' ability to recognize more and less effective responses to managerial situations. The case scenario problems (CSPs) presented a fictitious business case, which consisted of a brief overview of the problem, the respondent's role, a history of the organization, and various documents such as organizational charts, departmental memos, email correspondence, financial tables, and/or product descriptions. The scenarios were followed by a series of open-ended questions aimed at assessing problem identification, solution generation, information processing, and outcome monitoring.

Hedlund and colleagues (2006) administered the SJPs and CSPs to two samples of incoming MBA students (total $N = 792$). Scores on the SJPs and CSPs related significantly to first year and final GPA and exhibited modest but significant correlations with participation in extracurricular activities and leadership positions. There were no significant correlations between either question format and scores on the GMAT and scores on both the SJPs and CSPs explained significant variance in grades beyond GMAT scores and undergraduate GPA.

In comparing the two question formats, Hedlund and colleagues (2006) found that the CSPs exhibited slightly better predictive and incremental validities than the SJPs with regard to academic performance. The SJPs, on the other hand, produced less racial/ethnic disparity in scores than the CSPs. The students preferred the format of the SJPs and evaluated their own performance to be higher on the SJPs but they tended to view the CSPs as more relevant to job performance and more potentially useful in an admissions process. In general, TKIs have the advantage of being easier to develop, administer, and score but measures of practical problem-solving skills may provide broader insight into the skills that underlie knowledge acquisition. They also provide a potential avenue for facilitating the development of practical abilities.

Developing Practical Intelligence

Although the majority of research to date has focused on assessing individual differences in practical intelligence, there is emerging interest in identifying ways to develop practical intelligence. These efforts may involve directly teaching the "lessons learned" of more experienced and successful practitioners or helping individuals develop the skills to learn more effectively from their own experiences (Cianciolo, Antonakis, & Sternberg, 2004; Hedlund & Sternberg, 2001; Wagner, 1997).

Sharing Tacit Knowledge

One of the products of research on tacit knowledge is a body of knowledge that can be incorporated into training and development initiatives in order to share the lessons of experience with others. The uncovered tacit knowledge may be shared directly with others (e.g., reading a story about someone's experience) or it may be used to help guide individuals to the types of situations that are conducive to acquiring relevant tacit knowledge. Prescriptions for designing effective training programs suggest that training should build on trainees' prior knowledge, use relevant and concrete examples, help trainees interpret their experiences, provide opportunities to apply general principles, and provide feedback (see, e.g., Campbell, 1988; Howell & Cooke, 1989).

The sharing on tacit knowledge can take the form of "rules of thumb" about how to respond in various situations or case studies that allow learners to assess the situation, evaluate the course of action taken, and assess the consequences of the action. The situations from tacit knowledge inventories can be developed into behavioral role-playing scenarios or simulations, which have been shown to be effective methods for developing practical competencies (Burke & Day, 1986; Keys & Wolfe, 1990; Latham, 1988; Thornton & Cleveland, 1990).

Facilitating Knowledge Acquisition

Even when efforts are made to provide opportunities to acquire knowledge, it is clear that some individuals are more skilled at learning from experience than others. By understanding the processes that underlie the successful acquisition of tacit knowledge, individuals can be taught to be more sensitive to the lessons of experience. Individuals, for example, can be given strategies to help them focus on the knowledge-acquisition components of selective encoding, selective combination, and selective comparison (Sternberg, 1988, 1997). Teaching these strategies could entail providing examples in which the relevant information is highlighted, showing charts or figures that illustrate how the information is combined, and explaining how the new information is related to prior knowledge. Individuals could also be given question prompts to practice using in solving new and unfamiliar problems.

Several studies demonstrate the potential to facilitate knowledge acquisition. Sternberg, Wagner, and Okagaki (1993) assigned participants to one of five conditions that varied in regard to the cues provided to help with tacit knowledge acquisition. For example, in the selective encoding condition, relevant information was highlighted and a relevant rule of thumb provided. Participants in the control group (with no cues) performed the worst in terms of their accuracy in identifying relevant information. Participants assigned to the selective encoding and selective combination conditions showed the most gain in tacit knowledge scores. These findings suggest that prompting individuals to focus on certain information can enhance the acquisition of tacit knowledge.

Matthew and Sternberg (2009) studied the effectiveness of three training interventions, or critical thinking exercises, that emphasized different aspects of the condition-action structure of tacit knowledge. The condition-focused method was aimed at helping learners focus on problem identification and goal formulation. The action-focused method focused the learner on the link between action and outcomes, including alternative responses. The condition and action-focused method encouraged reflection on both the condition and the action.

In the first study with Army officers, participants in the combined condition and action intervention showed significant improvement in tacit knowledge scores. In the second study, undergraduate students completed the College Student Questionnaire and two college case studies designed for the experiment. There were no significant effects of reflection condition on tacit knowledge scores. Matthew and Sternberg (2009) suggested that the interventions may have been too brief to have had a substantial impact on practical problem-solving skills. However, the findings with the military sample, although modest, provide encouragement that such skills can be developed through intervention.

Razali and Trevelyan (2012) examined the influence of laboratory classes on the development of practical intelligence among engineering students. Students completed a practical intelligence test that consisted of various engineering problems before and after the laboratory classes. The experimental group showed significant improvement in practical intelligence scores and significantly higher posttest practical intelligence than the control group.

In addition to experimental interventions, researchers have explored the influence of human resource (HR) practices on knowledge acquisition and knowledge sharing (Chuang, Jackson, & Jiang, 2016). Their specific focus was on knowledge-intensive teams in the IT industry, whose work involves collaboration among team members to locate, share, create, and apply knowledge. The authors conceptualized knowledge on a continuum from mostly explicit (i.e., easily codified and recorded) to mostly tacit (i.e., more complex and subjective).

Chuang, Jackson, and Jiang (2016) surveyed 172 team leaders and 826 members from thirty-four IT firms in Taiwan. They found that HR practices had the most influence on knowledge acquisition when the knowledge was highly explicit (or less tacit). In other words, HR practices were more effective in facilitating knowledge acquisition of more explicit knowledge but less effective with the acquisition of more subtle, less readily observable tacit knowledge. Supportive HR practices were equally effective at promoting knowledge sharing regardless of whether the knowledge was more or less tacit. They suggest that future research should "explore other potential means for facilitating the acquisition of tacit knowledge since the power of HRM systems appears to be constrained by knowledge tacitness" (p. 545).

The Future of Practical Intelligence Research

Wagner's (2011) chapter in the previous edition of the *Cambridge Handbook of Intelligence* suggested that the convergence of several theoretical perspectives might diminish the need for a distinct focus on practical intelligence. Although some research has raised questions about the distinctiveness of practical intelligence from general intelligence (Cianciolo et al., 2006; Gottfredson, 2003), the body of evidence continues to grow in support of the concept of practical intelligence and the value it adds to understanding performance in a wide variety of contexts. The most prolific area of research on practical intelligence has been the work on tacit knowledge, which has demonstrated its relevance to performance in employment, education, and everyday life.

Among researchers who have studied tacit knowledge, there is a general consensus that efforts should be aimed at facilitating the acquisition and dissemination of such knowledge. Two promising directions for future research reflect these views. The first direction builds on work that has already begun to identify the tacit knowledge that distinguishes novice from expert performers. Researchers should explore the most effective ways to share the lessons of experience in order to facilitate more effective and efficient development of expertise. The dissemination of tacit knowledge would be particularly advantageous in settings where there is high turnover and limited time for reflection (e.g., McQueen & Janson, 2016) or where the stakes associated with allowing novices to perform tasks are high (e.g., Taylor, Van Der Heijden, & Genuchi, 2017). One way of sharing knowledge is through communities of practice, where individuals come together to discuss problems and exchange lessons learned (Wagner, 2011). In the military, tacit knowledge scenarios have been shared through websites and an experimental forum developed to facilitate in-depth discussion and reflection on the problems depicted in the scenarios (Cianciolo et al., 2004). Technological advancements not only increase opportunities to disseminate knowledge but also support dynamic processing of that knowledge. Kahn and Khader (2014), for example, propose that e-learning tools can facilitate the sharing of tacit knowledge with the right person at the right time. They describe a dynamic query-handling system that transfers a query from a novice and matches that query with the appropriate expert.

The second promising area for further research builds on efforts to improve knowledge acquisition through interventions aimed at enhancing problem-solving skills. Research has shown that tacit knowledge acquisition can be facilitated by providing cues to help learners process information more effectively (Sternberg, Wagner, & Okagaki, 1993) or by encouraging learners to reflect on their problem-solving processes (Matthew & Sternberg, 2009). Reflection has long been recognized as vital to experience-based learning but more research is needed to understand what types of reflection methods best facilitate the acquisition of knowledge that typically lies outside of focal awareness. Researchers might also explore how different learning orientations might influence the effectiveness of different reflection techniques. Identifying ways to help individuals more effectively and efficiently learn from experience might prove to be the greatest contribution of the research on practical intelligence, and secure a place for a chapter on practical intelligence in the next edition of this handbook.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Armstrong, S. J., & Mahmud, A. (2008). Experiential learning and the acquisition of managerial tacit knowledge. *Academy of Management Learning and Education, 7*, 189–208. https://doi.org/10.5465/AMLE.2008.32712617

Baczyńska, A. (2015). Analytical, practical and emotional intelligence and line manager competencies. *Management and Business Administration. Central Europe*, *23*, 34–54. https://doi.org/10.7206/mba.ce.2084-356.156

Baum, J. R., Bird, B. J., & Singh, S. (2011). The practical intelligence of entrepreneurs: Antecedents and a link with new venture growth. *Personnel Psychology*, *64*(2), 397–425. https://doi.org/10.1111/j.1744-6570.2011.01214.x

Berg, C. A. (2000). The development of adult intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 117–137). New York: Cambridge University Press.

Burke, M. J., & Day, R. R. (1986). A cumulative study of the effectiveness of managerial training. *Journal of Applied Psychology*, *71*, 232–246.

Campbell, J. P. (1988). Training design for performance improvement. In J. P. Campbell & R. J. Campbell (Eds.), *Productivity in organizations: New perspectives from industrial and organizational psychology* (pp. 177–216). San Francisco: Jossey-Bass.

Carraher, T. N., Carraher, D., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology*, *3*, 21–29.

Ceci, S. J., & Liker, J. (1986). Academic and nonacademic intelligence: An experimental separation. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 119–142). New York: Cambridge University Press.

Ceci, S. J., & Liker, J. (1988). Stalking the IQ-expertise relationship: When the critics go fishing. *Journal of Experimental Psychology: General*, 117, 96–100.

Ceci, S. J., & Roazzi, A. (1994). The effects of context on cognition: Postcards from Brazil. In R. J. Sternberg & R. K. Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence* (pp. 74–101). New York: Cambridge University Press.

Ceci, S.J., & Ruiz, A. (1992). The role of general ability in cognitive complexity: A case study of expertise. In R. Hoffman (Ed.), *The psychology of expertise: Cognitive research and empirical AI* (pp. 218–230). New York: Springer-Verlag.

Chan, D., & Schmitt, N. (1998). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology*, *82*, 143–159.

Chuang, C., Jackson, S. E., & Jiang, Y. (2016). Can knowledge-intensive teamwork be managed? Examining the roles of HRM systems, leadership, and tacit knowledge. *Journal of Management, 42*, 524–554. https://doi.org/10.1177/0149206313478189

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.) (1988). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cianciolo, A. T., Antonakis, J., & Sternberg, R. J. (2004). Practical intelligence and leadership: Using experience as a 'mentor'. In D. V. Day, S. J. Zaccaro, S. M. Halpin, D. V. Day, S. J. Zaccaro, S. M. Halpin (Eds.), *Leader development for transforming organizations: Growing leaders for tomorrow* (pp. 211–236). Mahwah, NJ: Lawrence Erlbaum Associates.

Cianciolo, A. T., Grigorenko, E. L., Jarvin, L., Gil, G., Drebot, M. E., & Sternberg, R. J. (2006). Practical intelligence and tacit knowledge: Advancements in the measurement of developing expertise. *Learning and Individual Differences, 16*, 235–253. https://doi.org/10.1016/j.lindif.2006.04.002

Colonia-Willner, R. (1998). Practical intelligence at work: Relationship between aging and cognitive efficiency among managers in a bank environment. *Psychology and Aging*, *13*, 45–57.

Colonia-Willner, R. (1999). Investing in practical intelligence: Ageing and cognitive efficiency among executives. *International Journal of Behavioral Development*, 23, 591–614.

Eddy, A. S. (1988). The relationship between the Tacit Knowledge Inventory for Managers and the Armed Services Vocational Aptitude Battery. Unpublished master's thesis, St. Mary's University, San Antonio, TX.

Elliott, J. G., Stemler, S. E., Sternberg, R. J., Grigorenko, E. L., & Hoffman, N. (2011). The socially skilled teacher and the development of tacit knowledge. *British Educational Research Journal, 37*, 83–103.

Fox, S., & Spector, P. E. (2000). Relations of emotional intelligence, practical intelligence, general intelligence, and trait affectivity with interview outcomes: It's not all just "G." *Journal of Organizational Behavior*, *21*, 203–220.

Frederiksen, N. (1966). Validation of a simulation technique. *Organizational Behavior and Human Performance*, *1*, 87–109.

Frederiksen, N. (1986). Toward a broader conception of human intelligence. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 84–116). New York: Cambridge University Press.

Gottfredson, L. S. (2003). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence*, *31*, 343–397.

Grigorenko, E. L., & Sternberg, R. J. (2001). Analytical, creative, and practical intelligence as predictors of self-reported adaptive functioning: A case study in Russia. *Intelligence, 29*, 57–73. https://doi.org/10.1016/S0160-2896(00)00043-X

Grigorenko, E. L., Meier, E., Lipka, J., Mohatt, G., Yanez, E., & Sternberg, R. J. (2004). Academic and practical intelligence: A case study of the Yup'ik in Alaska. *Learning and Individual Differences, 14*(4), 183–207.

Grigorenko, E. L., Sternberg, R. J., & Strauss, S. (2006). Practical intelligence and elementary-school teacher effectiveness in the United States and Israel: Measuring the predictive power of tacit knowledge. *Thinking Skills and Creativity*, *1*, 14–33.

Groen, G. J., & Patel, V. L. (1988). The relationship between comprehension and reasoning in medical expertise. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise* (pp. 287–310). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hanson, M. A., & Ramos, R. A. (1996). Situational judgment tests. In R. S. Barrett (Ed.), *Fair employment strategies in human resource management* (pp. 119–124). Westport, CT: Greenwood Publishing.

Hedlund, J., Forsythe, G. B., Horvath, J. A., Williams, W. M., Snook, S., & Sternberg, R. J. (2003). Identifying and assessing tacit knowledge: Understanding the practical intelligence of military leaders. *Leadership Quarterly*, *14*, 117–140.

Hedlund, J., & Sternberg, R. J. (2001). Practical intelligence: Implications for personnel research. In G. R. Ferris (Ed.), *Research in personnel and human resource management*, Vol. 19. Stamford, CT: JAI Press.

Hedlund, J., Wilt, J. M., Mebel, K. L., Ashford, S. J., & Sternberg, J. (2006). Assessing practical intelligence in business school admissions: A supplement to graduate management admissions test. *Learning and Individual Differences*, *16*(2), 101–217.

Howell, W. C., & Cooke, N. J. (1989). Training the human information processor: A review of cognitive models. In I. L. Goldstein (Ed.), *Training and development in organizations* (pp. 121–182). San Francisco: Jossey-Bass.

Insch, G. S., McIntyre, N., & Dawley, D. (2008). Tacit knowledge: A refinement and empirical test of the Academic Tacit Knowledge Scale. *Journal of Psychology*, *142*, 561–579.

Jisr, R. E., & Maamari, B. E. (2017). Effectuation: Exploring a third dimension to tacit knowledge. *Knowledge and Process Management*, *24*, 72–78. https://doi.org/10.1002/kpm.1536

Joseph, D., Ang, S., Chang, R. H., & Slaughter, S. A. (2010). Practical intelligence in IT: Assessing soft skills of IT professionals. *Communications of the ACM*, *53*, 149–154. https://doi.org/10.1145/1646353.1646391

Keys, B., & Wolfe, J. (1990). The role of management games and simulations in education and research. *Journal of Management*, *16*, 307–336.

Khan, A. A., & Khader, S. A. (2014). An approach for externalization of expert tacit knowledge using a query management system in an e-learning environment. *International Review of Research in Open and Distance Learning*, *15*, 257–274.

Latham, G. P. (1988). Human resource training and development. *Annual Review of Psychology*, *39*, 545–582.

Lave, J., Murtaugh, M., & de la Roche, O. (1984). The dialectic of arithmetic in grocery shopping. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 67–94). Cambridge, MA: Harvard University Press.

Legree, P. J. (1995). Evidence for an oblique social intelligence factor established with a Likert-based testing procedure. *Intelligence*, *21*, 247–266.

Loehlin, J. C. (2000). Group differences in intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 176–193). New York: Cambridge University Press.

Mandelman, S. D., Barbot, B., & Grigorenko, E. L. (2016). Predicting academic performance and trajectories from a measure of successful intelligence. *Learning and Individual Differences, 51*, 387–393. https://doi.org/10.1016/j.lindif.2015.02.003

Matthew, C. T., & Sternberg, R. J. (2009). Developing experience-based (tacit) knowledge through reflection. *Learning & Individual Differences, 19*, 530–540.

McDaniel, M. A., & Whetzel, D. L. (2005). Situational judgment test research: Informing the debate on practical intelligence theory. *Intelligence*, *33*, 515–525.

McQueen, R. J., & Janson, A. (2016). Accelerating tacit knowledge building of client-facing consultants: Can organizations better support these learning processes? *Learning Organization*, *23*, 202–217.

Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology*, *75*, 640–647.

Murtaugh, M. (1985). The practice of arithmetic by American grocery shoppers. *Anthropology and Education Quarterly*, *16*, 186–192.

Neisser, U. (1976). *Cognition and reality.* San Francisco: Freeman.

Neisser, U., Boodo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Nunes, T., Schliemann, A. D., & Carraher, D. W. (1993). *Street mathematics and school mathematics*. New York: Cambridge University Press.

Perret-Clermont, A.N. (1980). *Social interaction and cognitive development in children*. London: Academic Press.

Polanyi, M. (1966). *The tacit dimensions*. Garden City, NY: Doubleday.

Razali, Z. B., & Trevelyan, J. P. (2012). An evaluation of students' practical intelligence and ability to diagnose equipment faults. *Procedia – Social and Behavioral Sciences*, *56*, 42–51.

Rogoff, B., & Lave, J. (Eds.). (1984). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Scribner, S. (1984). Studying working intelligence. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 9–40). Cambridge, MA: Harvard University Press.

Scribner, S. (1986). Thinking in action: Some characteristics of practical thought. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 13–30). New York: Cambridge University Press.

Stemler, S. E., & Sternberg, R. J. (2006). Using situational judgment tests to measure practical intelligence. In J. A. Weekley, R. E. Ployhart, J. A. Weekley, & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 107–131). Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Penguin Books.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume Books.

Sternberg, R. J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology*, *24*, 259–375.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., et al. (2000). *Practical intelligence in everyday life*. New York: Cambridge University Press.

Sternberg, R. J., Nokes, K., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., & Grigorenko, E. L. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence*, *29*, 401–418.

Sternberg, R. J., & The Rainbow Project Collaborators (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence, 34*, 321–350.

Sternberg, R. J., & Wagner, R. K. (Eds.). (1986). *Practical intelligence: Nature and origins of competence in the everyday world*. New York: Cambridge University Press.

Sternberg, R. J., & Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: The nature and role of tacit knowledge in work and at school. In H. Reese & J. Puckett (Eds.), *Advances in lifespan development* (pp. 205–227). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. *American Psychologist*, *50*, 912–927.

Sujan, H., Sujan, M., & Bettman, J. R. (1991). The practical know-how of selling: differences in knowledge content between more-effective and less-effective performers. *Marketing Letters*, *2*, 367.

Tan, H., & Libby, R. (1997). Tacit managerial versus technical knowledge as determinants of audit expertise in the field. *Journal of Accounting Research*, *35*, 97–113.

Taylor, T., Elison-Bowers, P., Werth, E., Bell, E., Carbajal, J., Lamm, K., & Velazquez, E. (2013). A Police Officer's Tacit Knowledge Inventory (POTKI): Establishing construct validity and exploring applications. *Police Practice and Research, 14*, 478–490. https://doi.org/10.1080/15614263.2013.802847

Taylor, T. Z., Psotka, J., & Legree, P. (2015). Relationships among applications of tacit knowledge and transformational/transactional leader styles: An exploratory comparison of the MLQ and TKML. *Leadership and Organization Development Journal, 36*, 120–136. https://doi.org/10.1108/LODJ-01-2013-0008

Taylor, T. Z., Van der Heijden, B. I., & Genuchi, M. C. (2017). The Police Officer Tacit Knowledge Inventory (POTKI): Towards determining underlying structure and applicability as a recruit screening tool. *Applied Cognitive Psychology, 31*, 236–246. https://doi.org/10.1002/acp.3321

Thornton, G. C., & Cleveland, J. N. (1990). Developing managerial talent through simulation. *American Psychologist*, *45*, 190–199.

Wagner, R. K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology*, *52*, 1236–1247.

Wagner, R. K. (2011). Practical intelligence. In R. J. Sternberg, S. B. Kaufman, R. J. Sternberg, & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 550–563). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511977244.028

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology*, *49*, 436–458.

Wagner, R. K., & Sternberg, R. J. (1986). Tacit knowledge and intelligence in the everyday world. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 51–83). New York: Cambridge University Press.

Wagner, R. K., & Sternberg, R. J. (1990). Street smarts. In K. E. Clark & M. B. Clark (Eds.), *Measures of leadership* (pp. 493–504). West Orange, NJ: Leadership Library of America.

Wagner, R. K., & Sternberg, R. J. (1991). *Tacit Knowledge Inventory for Managers*. San Antonio, TX: The Psychological Corporation.

Wagner, R. K., Sujan, H., Sujan, M., Rashotte, C. A., & Sternberg, R. J. (1999). Tacit knowledge in sales. In R. J. Sternberg & J. A. Horvath (Eds.), *Tacit knowledge in professional practice* (pp. 155–182). Mahwah, NJ: Lawrence Erlbaum Associates.

Weekley, J. A., & Jones, C. (1997). Video-based situational testing. *Personnel Psychology*, *50*, 25–49.

Yalon-Chamovitz, S., & Greenspan, S. (2005). Ability to identify, explain and solve problems in everyday tasks: Preliminary validation of a direct video measure of practical intelligence. *Research in Developmental Disabilities: A Multidisciplinary Journal*, *26*, 219–230.

31 Social Intelligence

John F. Kihlstrom and Nancy Cantor

The term *social intelligence* was first used by Dewey (1909) and Lull (1911) but the modern concept has its origins in Thorndike's (1920) division of intelligence into three facets pertaining to the ability to understand and manage ideas (abstract intelligence), concrete objects (mechanical intelligence), and people (social intelligence). In Thorndike's classic formulation: "By social intelligence is meant the ability to understand and manage men and women, boys and girls – to act wisely in human relations" (p. 228). Similarly, Moss and Hunt (1927) defined social intelligence as the "ability to get along with others" (p. 108). Vernon (1933) provided the most wide-ranging definition of social intelligence as the "ability to get along with people in general, social technique or ease in society, knowledge of social matters, susceptibility to stimuli from other members of a group, as well as insight into the temporary moods or underlying personality traits of strangers" (p. 44).

By contrast, Wechsler (1939) gave scant attention to social intelligence in the development of the Wechsler Adult Intelligence Scale (WAIS). He did acknowledge that the Picture Arrangement subtest of the WAIS might serve as a measure of social intelligence because it assesses the individual's ability to comprehend social situations (Campbell & McCord, 1996). In Wechsler's (1958) view, however, "social intelligence is just general intelligence applied to social situations" (p. 75). This dismissal was repeated in Matarazzo's (1972, p. 209) fifth and final edition of Wechsler's monograph, in which *social intelligence* dropped out as an index term.

Measuring Social Intelligence

Defining social intelligence seems easy enough, especially by analogy to abstract intelligence. When it came to *measuring* social intelligence, however, Thorndike (1920) noted somewhat ruefully that "convenient tests of social intelligence are hard to devise . . . Social intelligence shows itself abundantly in the nursery, on the playground, in barracks and factories and salesroom, but it eludes the formal standardized conditions of the testing laboratory. It requires human beings to respond to, time to adapt its responses, and face, voice, gesture, and mien as tools" (p. 231). Nevertheless, true to the goals of the psychometric tradition, researchers quickly translated the abstract definitions of social intelligence into standardized laboratory instruments for measuring individual differences in social intelligence (Landy, 2006; Taylor, 1990; Walker & Foley, 1973).

The George Washington Social Intelligence Test

The first of these was the George Washington Social Intelligence Test (GWSIT; Hunt, 1928; Moss, 1931; Moss & Hunt, 1927). Like the WAIS (which it preceded), the GWSIT was composed of a number of subtests, which could be combined to yield an aggregate score. Hunt (1928) originally validated the GWSIT through its correlations with adult occupational status, the number of extracurricular activities pursued by college students, and supervisor ratings of employees' ability to get along with people. There was some controversy about whether social intelligence should be correlated with personality measures of sociability or extraversion.

However, the GWSIT came under immediate criticism for its relatively high correlation with abstract intelligence. Thorndike and Stein (1937) concluded that the GWSIT "is so heavily loaded with ability to work with words and ideas, that differences in social intelligence tend to be swamped by differences in abstract intelligence" (p. 282). The inability to discriminate between social intelligence and IQ, coupled with difficulties in selecting external criteria against which the scale could be validated, led to declining interest in the GWSIT and, indeed, in the whole concept of social intelligence as a distinct intellectual entity. Spearman's (1927) *g* afforded no special place for social intelligence, of course; nor was social intelligence included, or even implied, in Thurstone's list of primary mental abilities.

Social Intelligence in Guilford's Structure of Intellect

Work on social intelligence fell off sharply until the 1960s, when interest was revived within the context of Guilford's Structure of Intellect model of intelligence. Guilford postulated a system of at least 120 separate intellectual abilities, based on all possible combinations of five categories of *operations* (cognition, memory, divergent production, convergent production, and evaluation), four categories of *content* (figural, symbolic, semantic, and behavioral), and six categories of *products* (units, classes, relations, systems, transformations, and implications). Within this system, social intelligence was represented by behavioral contents. Of the thirty facets of social intelligence predicted by the Structure of Intellect model (five operations × six products), however, actual tests were devised for only six cognitive abilities (Hoepfner & O'Sullivan, 1969) and six divergent production abilities (Hendricks, Guilford, & Hoepfner, 1969).

In constructing tests of behavioral cognition, O'Sullivan, Guilford, and deMille (1965) assumed that "expressive behavior, more particularly facial expressions, vocal inflections, postures, and gestures, are the cues from which intentional states are inferred" (p. 6). Their study yielded six factors clearly interpretable as cognition of behavior, which were not contaminated by nonsocial semantic and spatial abilities. However, later studies found substantial correlations between IQ and scores on the individual Guilford subtests as well as various composite social intelligence scores (Riggio, Messamer, & Throckmorton, 1991; Shanley, Walker, & Foley, 1971). Still, Shanley and colleagues (1971) conceded that the correlations obtained were not strong enough to warrant Wechsler's assertion that social intelligence is nothing more than general intelligence applied in the social domain.

Hendricks and colleagues (1969) attempted to develop tests for coping with other people, not just understanding their behavior – what they referred to as "basic solution-finding skills in interpersonal relations" (p. 3). Because successful coping involves the creative generation of many and diverse behavioral ideas, these investigators labeled these divergent-thinking abilities *creative social intelligence*. Scoring divergent productions proved considerably harder than scoring cognitions, as there are by definition no best answers and responses must be evaluated by independent judges for quality as well as quantity. Factor analysis yielded six factors clearly interpretable as divergent production in the behavioral domain, which were essentially independent of both divergent semantic production and (convergent) cognition in the behavioral domain (see also Chen & Michael, 1993; Romney and Pyryt, 1999; Snyder & Michael, 1983). In neither domain is there much evidence for the ability of any of these tests to predict external criteria of social intelligence.

Tests of the remaining three Structure of Intellect domains had not been developed by the time the Guilford program came to a close. Hendricks and colleagues (1969) noted that "these constitute by far the greatest number of unknowns in the [Structure of Intellect] model" (p. 6). However, O'Sullivan and colleagues (1965) did sketch out how these abilities were defined. *Convergent production* in the behavioral domain was defined as "doing the right thing at the right time" (p. 5) and presumably might be tested by a knowledge of etiquette. *Behavioral memory* was defined as the ability to remember the social characteristics of people (e.g., names, faces, and personality traits), while *behavioral evaluation* was defined as the ability to judge the appropriateness of behavior.

The Magdeburg Test of Social Intelligence

Given the difficulties in constructing and validating performance-based tests of social intelligence, as illustrated by the Guilford program, it is not surprising that many investigators have turned to self-report inventories such as the Tromso Social Intelligence Scale (Grieve, 2013; Silvera, Martinussen, & Dahl, 2001) and the Trait Social Intelligence Questionnaire (Petrides, Mason, & Sevdalis, 2011).

A renewed attempt to develop a performance-based assessment yielded the Magdeburg Test of Social Intelligence (MTSI; Conzelmann, Weis, & Süss, 2013), based on a model of social intelligence proposed by Weis and Süss (2007). The MTSI is an extensive battery of tests consisting of a variety of verbal, pictorial, audio, and video materials assessing various aspects of *social perception* (the ability to quickly perceive social information in complex settings), *social memory* (the ability to store and recall social information), and *social understanding* (the ability to understand social stimuli presented in a situational context). Unfortunately, exploratory factor analysis showed that the various measures of social perception did not converge on a single construct. The measures of social memory and social understanding, however, did show substantial convergent and discriminant validity, supporting the hypothesis that social intelligence is multidimensional in nature. None of these dimensions correlated with any of the "Big Five" personality traits. Two other aspects of social intelligence hypothesized by the Weis-Süss model, *social flexibility* (the ability to produce many and diverse solutions in a social situation) and *social knowledge* (the individual's fund of knowledge about the social world), are not assessed by the current version of the MTSI.

Convergent and Discriminant Validity in Social Intelligence

Following the Guilford studies, a number of investigators continued the attempt to measure social intelligence and determine its relation to general abstract intelligence. Most of these studies explicitly employed the logic of the multitrait-multimethod matrix, employing multiple measures of social and nonsocial intelligence and examining the convergent validity of alternative measures within each domain and discriminant validity across domains (e.g., Sechrest & Jackson, 1961; Lee et al., 2000; Weis & Süss, 2007).

Marlowe (1986) and his colleagues assembled a large battery of personality measures tapping various aspects of social intelligence, including interest and concern for other people, social performance skills, empathic ability, emotional expressiveness and sensitivity to others' emotional expressions, social anxiety, and lack of social self-efficacy and self-esteem. These scales were essentially unrelated to verbal and abstract intelligence but this apparent independence of social and general intelligence may be at least partially an artifact of method variance: Marlowe's measures of social intelligence were all self-report scales, whereas his measures of verbal and abstract intelligence were the usual sorts of objective performance tests.

Keeping the methods constant, Conzelmann and colleagues (2013) examined the correlations between the MTSI subscales and "academic" intelligence (the Berlin Intelligence Structure Test). Both social perception and social memory were correlated with academic intelligence, perhaps owing to the complexity of the MTSI tasks. Although social understanding proved to be unrelated to measures of academic reasoning, the distinction between social intelligence and intelligence in general remains problematic.

The Prototype of Social Intelligence

Although social intelligence has proved difficult for psychometricians to operationalize, it does appear to play a major role in people's naïve, intuitive concepts of intelligence. Sternberg and his colleagues asked subjects to list the behaviors that they considered characteristic of intelligence, academic intelligence, everyday intelligence, and unintelligence; other subjects then rated each of 250 of these in terms of how "characteristic" each was of the ideal person possessing each of the three forms of intelligence (Sternberg et al., 1981). Factor analysis of ratings provided by laypeople yielded a factor of "social competence." Prototypical behaviors reflecting social competence were these: accepts others for what they are; admits mistakes; displays interest in the world at large; is on time for appointments;

has social conscience; thinks before speaking and doing; displays curiosity; does not make snap judgments; makes fair judgments; assesses well the relevance of information to a problem at hand; is sensitive to other people's needs and desires; is frank and honest with self and others; and displays interest in the immediate environment.

Interestingly, a separate dimension of social competence did not consistently emerge in ratings made by a separate group of experts on intelligence. Rather, the experts focused on verbal intelligence and problem-solving ability, with social competence expressly emerging only in the ratings of the ideal "practically intelligent" person. Perhaps these experts shared Wechsler's dismissive view of social intelligence.

Similar studies were conducted by Kosmitzki and John (1993) and by Schneider, Ackerman, and Kanfer (1996), with similar results. In the Schneider and colleagues study, factor analysis revealed seven dimensions of social competence that were essentially uncorrelated with quantitative and verbal/reasoning ability. On the basis of these findings, Schneider and colleagues concluded that "it is time to lay to rest any residual notions that social competence is a monolithic entity, or that it is just general intelligence applied to social situations" (p. 479). As with Marlowe's (1986) study, however, the reliance on self-report measures of social intelligence compromises this conclusion, which remains to be confirmed using objective performance measures of the various dimensions in the social domain.

Social intelligence played little role in Sternberg's (1977) early componential view of human intelligence, which was intended to focus on reasoning and problem-solving skills as represented by traditional intelligence tests. However, social intelligence is explicitly represented in Sternberg's more recent *triarchic* view of intelligence (Sternberg, 1988), according to which intelligence is composed of analytical, creative, and practical abilities. Practical intelligence is defined in terms of problem-solving in everyday contexts and explicitly includes social intelligence (Sternberg & Wagner, 1986) – though it also includes such nonsocial skills as arithmetic and route-planning abilities. According to Sternberg, each type of intelligence reflects the operation of three different kinds of component processes: performance components, which solve problems in various domains; executive metacomponents, which plan and evaluate problem-solving; and knowledge-acquisition components, by which the first two components are learned through experience. "Successful" intelligence marshals all three kinds of abilities in pursuing goals and solving problems encountered along the way (Sternberg, 2018). For Sternberg, these abilities, and thus their underlying components, may well be somewhat independent of each other; but the actual relation among various intellectual abilities is an open, empirical question. Answering this question, of course, requires that we have psychometrically adequate instruments for assessing social intelligence – which brings us back to our starting point: How is social intelligence to be measured?

The Development of Social Intelligence

While psychometric research has focused on adults, there is also a longstanding interest in social intelligence among developmental psychologists (Greenspan & Love, 1997) – particularly those concerned with the assessment, treatment, and growth of children (and adults) with developmental disorders such as intellectual disability and autism.

Moral Reasoning

One stimulus for revived interest in social intelligence was the upsurge of interest in moral reasoning following the publication of Kohlberg's Piagetian theory of moral reasoning (e.g., Kohlberg, 1963). As Turiel (2006) notes, Piaget himself had viewed moral reasoning within the wider context of the child's knowledge and judgment of social relationships. So, just as Thorndike raised the question of how social intelligence related to academic intelligence, the Piaget-Kohlberg tradition raises the question of how age differences in moral reasoning are related to social reasoning in general. One view is that moral reasoning, while obviously related to social reasoning and to reasoning in general, constitutes a separate cognitive domain that might follow its own unique principles, developmental trajectory, and the like.

According to *social-cognitive domain theory* (Turiel, Killen, & Helwig, 1987; Smetana, 2006), morality is only one of several aspects of the social world about which children and adults acquire knowledge and engage in reasoning, judgment, and decision-making. The "conventional" domain of social knowledge has to do with norms of social behavior that vary from one context to another. The "personal" domain has to do with our understanding of individual persons as psychological entities, including the attributions that we make for our own and others' behaviors, and our ability to infer meaning in social situations. The "moral" domain concerns universally applicable and obligatory concepts of harm, welfare, fairness, and rights.

Most of the focus in social-cognitive domain theory has been on the moral domain and on children's developing the ability to understand moral concepts and render judgments of right and wrong. As a developmental theory, social-cognitive domain theory assumes that social-cognitive abilities are heterogeneous – that children's (and adults') abilities to reason about the social world and the trajectory of their development may well differ from one domain to another. But, for present purposes, social-cognitive domain theory offers an alternative description of the domains in which children and adults apply distinctively social intelligence.

Culture and Social Intelligence

While acknowledging that different aspects of social intelligence may have different developmental trajectories, the Piagetian tradition generally assumes that there is some objectively valid standard of morality (or, more broadly, social propriety) that individuals can identify through the application of rigorous, logical thought. On the other hand, increasing appreciation of cultural differences in mind and behavior suggests that there might not be such a single, universal standard (Shweder et al., 1998). In a discussion of the implications of multiculturalism for social intelligence, Shweder (2017) has suggested that "a highly developed social intelligence is one that is able to understand and sympathize with the unfamiliar and even ego-alien perspectives and attachments of the members of different cultural communities without shedding the attitudes, judgments, and feelings that give definition to one's own distinctive but culturally contoured and refined sense of self" (p. 321). At the very least, those who wish to construct assessments of social intelligence must attend to their own cultural biases; they may even wish to take up the challenge of devising "culture-fair" tests – or to consider the proposition that standards for intelligent social behavior may vary so much from one culture to the next as to require culture-specific tests.

Intellectual Disability

Social intelligence has always played a role in the assessment of intellectual disability (formerly known as mental retardation). This diagnosis requires not only evidence of subnormal intellectual functioning but also demonstrated evidence of "Deficits in adaptive functioning" that "limit functioning in one or more activities of daily life, such as communication, social participation, and independent living, across multiple environments, such as home, school, work, and community" (American Psychiatric Association, 2013). In other words, the diagnosis of intellectual disability involves deficits in social as well as academic intelligence. Furthermore, the wording of the diagnostic criteria implies that social and academic intelligence are not highly correlated – it requires positive evidence of *both* forms of impairment, meaning that the presence of one cannot be inferred from the presence of the other.

While the conventional diagnostic criterion for intellectual disability places primary emphasis on IQ and intellectual functioning, Greenspan and Love (1997) argued that it should emphasize social and practical intelligence instead. They proposed a hierarchical model of social intelligence consisting of three components: *social sensitivity*, reflected in role-taking and social inference; *social insight*, including social comprehension, psychological insight, and moral judgment; and *social communication*, subsuming referential communication and social problem-solving. Social intelligence, in turn, is only one component of *adaptive intelligence* (the others being *conceptual intelligence* and *practical intelligence*), which in turn joins *physical competence* and *socioemotional adaptation* (temperament and character) as the major dimensions of personal competence broadly construed. Greenspan and Love did not propose specific tests for any of these components of social intelligence but implied that they could be derived from experimental procedures used to study social cognition in general.

All this is well and good but, while the criterion for impaired intellectual functioning is clearly operationalized by an IQ threshold, there is as yet no standard by which impaired social functioning – impaired *social intelligence* – can be determined. The Vineland Social Maturity Scale (Doll, 1947) was an important step in this direction: This instrument yields aggregate scores of *social age* (analogous to mental age) and *social quotient* (by analogy to the intelligence quotient, calculated as social age divided by chronological age). The Vineland has been recently revised (Sparrow, Balla, & Cicchetti, 1984) but its adequacy as a pure measure of social intelligence is compromised by the fact that linguistic functions, motor skills, occupational skills, self-care, and self-direction are assessed as well as social cognition.

As an alternative, Taylor (1990) proposed a semi-structured Social Intelligence Interview covering such domains as social memory, moral development, recognition of and response to social cues, and social judgment. Unfortunately, such an interview, being idiographically constructed to take account of the individual's particular social environment, cannot easily yield numerical scores by which individuals can be compared and ranked. More important than ranking individuals, from Taylor's point of view, is identifying areas of high and low functioning within various environments experienced by the individual and determining the goodness of fit between the individual and the environments in which he or she lives.

The Autism Spectrum

Another group of developmental disabilities, autistic spectrum disorders, also invokes the concept of social intelligence. Kanner's (1943) classic description of autism portrays children who do not seem to be capable of engaging in normal social behavior or of maintaining normal social relationships, and the diagnostic criteria specified in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM; American Psychiatric Association, 2013) emphasize deficits in social relations: impairments in nonverbal behavior, failures to develop peer relationships, lack of spontaneous sharing and other aspects of social reciprocity; impairments in communication, including an inability to initiate or sustain conversations or social imitative play; and stereotyped patterns of behavior, including inflexibility in various behavioral routines. All of these features suggest that at least some cases of autism are characterized not just by social withdrawal and language impairment but by a specific impairment in the abilities that underlie effective social interaction.

Specifically, it has been proposed that autistic children and adults lack a "theory of mind" (ToM) by which they can attribute mental states to other people and reflect on their own mental lives (Bruner & Feldman, 1993; Tager-Flusberg, 2007). However, it is now recognized that autism lies on a spectrum, with Kanner's syndrome as an extreme case. This brings the problem of assessing social intelligence in intellectual disability and the autism spectrum directly in contact with a literature on the assessment of social intelligence in different cultures. Perhaps some autistic individuals lack some degree of social intelligence. On the other hand, perhaps their social intelligence is merely qualitatively different (Gernsbacher, 2015; Jaarsma & Welin, 2011). The fundamental questions endure: Is social cognition a separate faculty from nonsocial cognition? Is social intelligence anything different from general intelligence applied to the social domain? How does diversity in social intelligence relate to diversity in general intelligence?

Primate Social Intelligence

While the ontogenetic view of development focuses on the acquisition of social intelligence by individual children, the phylogenetic view asks questions about the evolution of social intelligence – and, in particular, about the social intelligence of our closest primate relatives. Most of this research has focused on whether any nonhuman species possess ToM, which might be taken as the most elementary aspect of social intelligence. Even though they fail nonverbal versions of the false belief test, laboratory studies confirm that chimpanzees, at least, possess the ability to understand the goals and intentions of others (Call & Tomasello, 2008). Studies of chimpanzees in more natural environments, as well as an appreciation of the complexities of primate societies, however, suggest a more expansive view of primate social intelligence, including the ability to understand the behavior of others in "human" terms of belief and desire (deWaal, 2016; deWaal & Ferrari, 2012; Seyfarth & Cheyney, 2015; Whiten & van de Waal, 2017).

Artificial Social Intelligence

Advances in artificial intelligence, robotics, and human-computer interaction, including Siri and other computer-based "virtual assistants," have led computer scientists to consider how incorporating various aspects of social cognition might enable machines to interact with humans more effectively (Bainbridge et al., 1994; Breazeal, 2002; Broadbent, 2017; Dautenhahn, 2007; Lepore, 2018). One possible approach would be to program various aspects of social intelligence directly, after the manner of Isaac Asimov's "Three Laws of Robotics." Another approach would be to employ powerful machine-learning algorithms to enable robots to acquire social intelligence from interactions with humans. Either approach will require establishing some consensus about what social intelligence is.

The Fall and Rise of Social Intelligence

Reviewing the literature published up to 1983, Landy (2006) characterized the search for social intelligence as "long, frustrating, and fruitless." Certainly it has been long and frustrating. Decade by decade, Landy traces a record of "disappointing empirical results and substantial theoretical criticism" (p. 82). This record did not, however, diminish the enthusiasm of both basic and applied social psychologists for the concept of social intelligence. Landy's review essentially stopped at 1983 and for good reason – for, very soon, events were to give social intelligence a new lease of life.

The Theory of Multiple Intelligences

The milestone event here was the theory of *multiple intelligences* proposed by Gardner (1983, 1993, 1999, 2006; see also Davis et al., 2011). Gardner proposed that intelligence is not a unitary cognitive ability but that there were at least seven (in the original formulation; later expanded to eight) quite different kinds of intelligence, each hypothetically dissociable from the others. While most of these proposed intelligences (linguistic, logical-mathematical, spatial, musical, bodily-kinesthetic, and naturalistic) are "cognitive" abilities somewhat reminiscent of Thurstone's primary mental abilities, two are explicitly personal and social in nature.

Intrapersonal intelligence is the ability to gain access to one's own internal emotional life and *interpersonal intelligence* is the ability to notice and make distinctions among other individuals. Gardner (1997) has also considered whether there is a specifically *moral* form of intelligence, which would count as another form of social intelligence.

Although Gardner's multiple intelligences are individual-differences constructs, in which some people or some diagnostic groups are assumed to have more of these abilities than others, Gardner does not rely entirely on the traditional psychometric procedures for documenting individual differences. Rather, his preferred method is a somewhat impressionistic analysis based on a convergence of signs provided by seven additional lines of evidence – *isolation by brain damage, exceptional cases, identifiable core operations, experimental tasks, distinctive developmental histories*, and a *unique symbol system* by which the ability in question can be manipulated and transmitted by a culture. For social intelligence, this symbol system is, at least in part, the language of traits – the thousands of terms that we use to describe each other's mental states but that do not apply to nonsentient objects. Gardner did not offer any new tests of social intelligence, nor did he provide compelling evidence that his multiple intelligences were really qualitatively different from each other. Still, claims for a neuropsychological dissociation between interpersonal intelligence and other forms of intelligence offer new life to the notion that social intelligence can be distinguished from linguistic, logical-mathematical, and spatial intelligence.

Emotional Intelligence

The idea of social intelligence also received a boost from arguments in favor of individual differences in *emotional* intelligence, defined as "the ability to monitor one's own and others' feelings, to discriminate among them, and to use this information to guide one's thinking and action" (Salovey & Mayer, 1990, p. 189; see also Mayer, Roberts, & Barsade, 2008; Mayer, Salovey, & Caruso, 2008). Emotional intelligence subsumes four component abilities: the ability to perceive emotions in oneself and others; to use emotions in the service of thinking and problem-solving; to understand emotions and the relations among them; and to manage emotions in oneself and others. Emotion is frequently evoked in a social context and many social interactions are laced with emotion. So emotional intelligence and social intelligence do share a sort of family resemblance and it would not be surprising to find that they are correlated. For example, the ability to decode nonverbal expressions of emotion (Rosenthal et al., 1979) is an important aspect of the ability to "read" social situations in general (e.g., Barnes & Sternberg 1989). On the other hand, emotional intelligence and social intelligence are not the same thing: There is nothing particularly social about snake phobia and there are many aspects of social cognition where emotion plays little or no role.

The idea of emotional intelligence quickly caught on in both academic and applied psychology (e.g., Goleman, 1995). Whereas Thorndike (1920) postulated social intelligence as the third member of a triad of intelligences, along with mechanical and abstract intelligence, it seems possible that, as suggested by Mayer, "Emotional intelligence could be . . . the replacement member of the triumvirate where social intelligence failed" (quoted in Goleman, 2006, p. 330). The explosion of interest in emotional intelligence probably has much to do with what might be called the "affective counterrevolution" in psychology – the feeling that, since the cognitive revolution of the 1950s and 1960s, psychology had gone overboard in emphasizing epistemology and needed to pay more attention to feelings and desires. Certainly there is little reason to think that emotional intelligence is a clearer concept than social intelligence or any easier to measure (Murphy, 2006). Whatever the reason, the upsurge of interest in emotional intelligence seems to have carried other "hot" or "personal" intelligences along with it, so that we can look forward to a revival of research interest in this topic (Mayer, Caruso, & Salovey, 2016).

Social Intelligence and Social Neuroscience

All the more so, perhaps, now that Goleman (2006) has done for social intelligence what he did earlier for emotional intelligence. Because rewarding social relationships are the key to happiness and health, and the key to rewarding social relationships is social intelligence, Goleman argued that we need new tools for the assessment of individual differences in social intelligence as well as educational programs that will enable people to learn how to increase their social intelligence in order to become happier and healthier. Whereas Gardner had postulated a single social intelligence, or perhaps two (counting *intra*personal as well as *inter*personal), Goleman argues for a highly differentiated set of social intelligences, grouped under two major headings. *Social awareness* includes the ability to perceive other people's internal mental states, to understand their feelings and thoughts, and to comprehend the demands of complex social situations. It includes modules dedicated to primal empathy, empathic accuracy, attunement, and social cognition. *Social facility*, or relationship management, "builds on social awareness to allow smooth, effective interactions" (p. 84) and includes interaction synchrony, self-presentation, influence, and concern for others.

Goleman (2006) provocatively characterizes previous work on social intelligence as a "scientific backwater" (p. 330) in need of total rethinking. Taking a cue from Gardner, who relied more on neuropsychology than on psychometrics, as well as the doctrine of modularity as it has developed in contemporary cognitive and social neuroscience, Goleman hypothesizes that social intelligence is mediated by an extensive network of neural modules, each dedicated to a particular aspect of social interaction. But, more than that, Goleman asserts that "new neuroscientific findings have the potential to reinvigorate the social and behavioral sciences," just as "the basic assumptions of economics . . . have been challenged by the emerging 'neuroeconomics,' which studies the brain during decision-making" (p. 324). On the other hand, it is a matter of historical fact that the real revolution in economics – the advances that garnered the Nobel Prizes – flowed from observational field studies (Simon, 1955) and paper-and-pencil questionnaires (Tversky & Kahneman, 1974). An argument can be made that, in personality and social psychology as in other areas of the field, psychological theory leads advances in neuroscience, not the other way around (Kihlstrom, 2010). Nevertheless, neuropsychological and brain-imaging research has already identified a number of brain modules or circuits that appear to be specialized for social cognition (Fiske & Prentice, 2011; Lieberman, 2007). Individual differences in the functioning of these areas may well prove to be related to individual differences in various aspects of social intelligence (Jimenez et al., 2013).

The Knowledge View of Social Intelligence

Intelligence, as defined in standard dictionaries, has two rather different meanings. In its most familiar meaning, intelligence has to do with the individual's ability to learn and reason. It is this meaning that underlies common psychometric notions such as *intelligence testing*, the *intelligence quotient*, and the like. As originally coined by E. L. Thorndike (1920) and pursued in the studies reviewed so far, *social intelligence* referred to the person's ability to understand and manage other people and to engage in adaptive social interactions. In its less common meaning, intelligence has to do with a body of information and knowledge. This second meaning is implicated in the titles of certain government organizations, such as the Central Intelligence Agency in the United States and its British counterparts MI5 and MI6. Both meanings are invoked by the concept of social intelligence. But, from Thorndike and Guilford to Gardner and Goleman and beyond, social intelligence research and theory have been predicated almost exclusively on what might be called the *ability view*.

Cantor and Kihlstrom offered an alternative *knowledge view* of social intelligence that refers simply to the individual's fund of knowledge about the social world (Cantor & Kihlstrom, 1987, 1989; Kihlstrom & Cantor, 1989, 2000, 2011). In contrast to the ability view of social intelligence, the knowledge view does not conceptualize social intelligence as a trait, or group of traits, on which individuals can be compared and ranked on a dimension from low to high. Rather, the knowledge view begins with the assumption that social behavior is *intelligent* – that it is mediated by what the person knows and believes to be the case and by cognitive processes of perception, memory, reasoning, and problem-solving, rather than being mediated by innate reflexes, conditioned responses, evolved genetic programs, and the like. Accordingly, the social intelligence view construes individual differences in social behavior – the public manifestations of personality – to be the product of individual differences in the knowledge that individuals bring to bear on their social interactions. Differences in social knowledge cause differences in social behavior but it does not make sense to construct measures of social IQ. The important variable is not *how much* social intelligence the person has but rather *what* social intelligence he or she possesses – what the individual knows about himself or herself, other people, the situations in which people encounter each other, and the behaviors they exchange when they are in them.

The Evolution of Cognitive Views of Personality

The social intelligence view of personality has its origins in the social-cognitive tradition of personality theory, in which construal and reasoning processes are central to issues of social adaptation. Thus, Kelly (1955) characterized people as naïve scientists generating hypotheses about future interpersonal events based on a set of personal constructs concerning self, others, and the world at large. These constructs were idiographic with respect to both content and organization. Individuals might be ranked in terms of the complexity of their personal construct systems but the important issue for Kelly was knowing *what* the individual's personal constructs were. Beyond complexity, the idiosyncratic nature of personal construct systems precluded much nomothetic comparison.

The initial formulation of social learning theory held that personality was largely learned behavior and that understanding personality required understanding the social conditions under which it was acquired (Miller & Dollard, 1941). Quite quickly, however, social learning slipped from its behaviorist roots and acquired a distinctly cognitive flavor (Bandura & Walters, 1963; Rotter, 1954). Bandura (1973) argued for the acquisition of social knowledge through precept and example rather than the direct experience of rewards and punishment and, later (Bandura, 1986), distinguished between the outcome expectancies emphasized by Rotter and individuals' "self-efficacy" expectancies concerning their ability to carry out the actions required to control the events in a situation. Although Rotter (1966) proposed a measure of generalized locus of control, Bandura argued that the important consideration is not whether an individual is relatively high or low in self-perceptions of social competence, or even actual social competence, but rather whether the person *believes* that he or she is competent to perform a particular behavior in some particular situation.

The immediate predecessor to the social intelligence view of personality is Mischel's (1968, 1973) cognitive social learning reconceptualization of personality. Although sometimes couched in behaviorist language, an emphasis on the *subjective meaning* of the situation marked even Mischel's early (1968) theory as cognitive in nature. Since that time, Mischel has broadened his conceptualization of personality to include a wide variety of different constructs, some derived from the earlier work of Kelly, Rotter, and Bandura and others imported from the study of human cognitive processes. From Mischel's (1973) point of view, the most important product of social learning is the individual's repertoire of *cognitive and behavioral construction competencies* – the ability to engage in a wide variety of skilled, adaptive behaviors, including both overt action and covert mental activities. These construction competencies are as close as Mischel gets to the ability view of social (or, for that matter, *non*social) intelligence.

On the other hand, the importance of perception and interpretation of events in Mischel's system calls for a second set of person variables, having to do with *encoding strategies* governing selective attention and Kellian personal *constructs* that filter people's perceptions, memories, and expectations. Following Rotter and Bandura, Mischel also stresses the role of stimulus-outcome, behavior-outcome, and self-efficacy *expectancies*. Also in line with Rotter's theory, Mischel notes that behavior will be governed by the *subjective values* associated with various outcomes. A final set of relevant variables consists of *self-regulatory systems and plans*, self-imposed goals and consequences that govern behavior in the absence (or in spite) of social monitors and external constraints. These variables are more in line with the knowledge view of social intelligence.

Social Intelligence as Social Knowledge

Following Winograd (1975) and Anderson (1976), Cantor and Kihlstrom (1987) classified social intelligence into two broad categories: *declarative social knowledge*, consisting of abstract concepts and specific memories, and *procedural social knowledge*, consisting of the rules, skills, and strategies by which the person manipulates and transforms declarative knowledge and translates knowledge into action. Following Tulving (1983), the individual's fund of declarative social knowledge, in turn, can be broken down further into context-free *semantic* social knowledge about the social world in general and *episodic* social memory for the particular events and experiences that make up the person's autobiographical record. Similarly, procedural knowledge can be subclassified in terms of cognitive and motoric social skills. These concepts, personal memories, interpretive rules, and action plans are the cognitive structures of personality. Together, they constitute the expertise that guides an individual's approach to solving the problems of social life.

The cognitive architecture of social intelligence will be familiar from the literature on social cognition (Carlston, 2013; Fiske & Macrae, 2012; Fiske & Taylor, 2007) – a literature that, interestingly, had its beginnings in early psychometric efforts to measure individual differences in social intelligence. For example, Vernon (1933) argued that one of the characteristics of a socially intelligent person was that he or she was a good judge of personality – a proposition that naturally led to inquiries into how people form impressions of personality. Research on person perception, in turn, led to an inquiry into the intuitive or implicit theories of personality that provide the cognitive basis for impression formation. Specifically, Cronbach (1955) argued that one's implicit theory of personality consisted of his or her knowledge of "the generalized Other" (p. 179) – of the important dimensions of personality and estimates of the mean and variance of each dimension within the population as well as estimates of the covariances among them. This intuitive knowledge might be widely shared and acquired as a consequence of socialization and acculturation processes; but he also assumed that there would be individual and cultural differences in this knowledge, leading to individual and group differences in social behavior.

Following Kelly (1955) and Mischel (1973), Cantor and Kihlstrom (1987) accorded *social concepts* a central status as cognitive structures of personality. If the purpose of perception is action, and if every act of perception is an act of categorization (Bruner, 1957), the particular categories that organize people's perception of the social world assume paramount importance in a cognitive analysis of personality. Some of these concepts concern the world of other people and the places we encounter them: knowledge of personality types, social groups, and social situations. Other concepts concern the *intra*personal world: the kinds of people we are, both in general and in particular classes of situations, and our theories of how we got that way. Some of these conceptual relations may be universal and others may be highly consensual within the individual's culture; but, as Kelly (1955) argued, some may be quite idiosyncratic. Regardless of whether they are shared with others, the individual's conceptual knowledge about the social world forms a major portion of his or her declarative social knowledge.

Another important set of declarative social knowledge structures represents the individual's autobiographical memory (Kihlstrom, 2009; McAdams & Manczak, 2015). In the context of social intelligence, autobiographical memory includes a narrative of the person's own actions and experiences but it also includes what he or she has learned through direct and vicarious experience about the actions and experiences of specific other people and the events that have transpired in particular situations. Every piece of conscious autobiographical memory is linked to a mental representation of the self as the agent or patient of some action, or the stimulus or experiencer of some state (Kihlstrom, Beer, & Klein, 2002).

On the procedural side, a substantial portion of the social intelligence repertoire consists of interpretive rules for making sense of social experience: for inducing social categories and deducing category membership; making attributions of causality; inferring other people's behavioral dispositions and emotional states; forming judgments of likability and responsibility; resolving cognitive dissonance; encoding and retrieving memories of our own and other people's behavior; predicting future events; and testing hypotheses about our social judgments. Some of these procedures are algorithmic in nature, while others may entail heuristic shortcuts (Nisbett & Ross, 1980). Some are enacted deliberately, while others may be evoked automatically, without much attention and cognitive effort on our part (Bargh, 1997; but see also Kihlstrom, 2008). They are all part of our repertoire of procedural social knowledge.

Social Intelligence in Life Tasks

From the knowledge view of social intelligence, the assessment of social intelligence has quite a different character than it does from the ability view. From a psychometric point of view, the questions posed have answers that are right or wrong: Is someone smart or not? Are smart people also friendly? Is it proper to giggle at a funeral? In this way, it is possible, at least in principle, to evaluate the accuracy of the person's social knowledge and the effectiveness of his or her social behaviors. However, the knowledge view, like the social intelligence approach to personality in general, abjures such rankings of people (Cantor, 2003). Rather than asking how socially intelligent people *are*, compared to some norm, the social intelligence view of personality asks what social intelligence people *have*, which they use to guide their interpersonal behavior. In fact, the social intelligence approach to personality is less interested in assessing the individual's repertoire of social intelligence than in seeking to understand the general cognitive structures and processes out of which individuality is constructed, how these develop over the life course of the individual, and how they play a role in ongoing social interactions. For this reason, Cantor and Kihlstrom (1987, 1989; Kihlstrom & Cantor, 1989) have not proposed any individual-differences measures by which the person's social intelligence can be assessed.

Although the social intelligence view of personality diverges from the psychometric approach to social intelligence on the matter of assessment, it agrees with some contemporary theorists that intelligence is context-specific (e.g., Sternberg, 1988). Social intelligence is specifically geared to solving the problems of social life and, in particular, managing the *life tasks, current concerns* (Klinger & Cox, 2011), or *personal projects* (Little, Salmela-Aro, & Phillips, 2007) that people select for themselves or that other people impose on them from outside. Social intelligence cannot be evaluated in the abstract but only with respect to the domains and contexts in which it is exhibited and the life tasks it is designed to serve. And, even in this case, "adequacy" cannot be judged from the viewpoint of the external observer but must come from the point of view of the particular person whose life tasks are in play.

Life tasks provide an integrative unit of analysis for studying the interaction between the person and the situation (e.g., Cantor, 1990, 2000, 2003; Cantor & Fleeson, 1994; Cantor & Harlow, 1994; Cantor et al., 2002; Snyder & Cantor, 1998). They may be explicit or implicit, abstract or circumscribed, universal or unique, enduring or stage-specific, rare or commonplace, poorly defined or well defined. Whatever their features, they give meaning to the individual's life and serve to organize his or her daily activities. Defined from the subjective point of view of the individual, they are the tasks that the person perceives himself or herself as "working on and devoting energy to solving during a specified period in life" (Cantor & Kihlstrom, 1987, p. 168). Life tasks are articulated by the individual as self-relevant, time-consuming, and meaningful. They provide a kind of organizing scheme for the individual's activities and they are embedded in the individual's ongoing daily life. They are responsive to the demands, structure, and constraints of the social environment in which the person lives. While often willingly undertaken, life tasks can also be imposed from outside and the ways in which they are approached may be constrained by sociocultural factors. Unlike the stage-structured views of Erikson and his popularizers, the social intelligence view of personality does not propose that everyone at a particular age is engaged in the same sorts of life tasks. Instead, periods of transition, when the person is entering into epochs in the life cycle, are precisely those times when individual differences in life tasks become most apparent.

The intelligent nature of life-task pursuit is illustrated by the strategies deployed in its service. People often begin to comprehend the problem at hand by simulating a set of plausible outcomes, relating them to previous experiences stored in autobiographical memory. They formulate specific plans for action and monitor their progress toward their goals, taking special note of obstacles and determining whether the actual outcome meets their original expectations. Much of the cognitive activity in life-task problem-solving involves forming causal attributions about outcomes and surveying autobiographical memory for hints about how things might have gone differently. When plans go awry or some unforeseen event frustrates progress, the person will map out a new path toward the goal or even choose a new goal compatible with a superordinate life task. Intelligence frees us from reflex, taxis, and instinct in social life as in nonsocial domains.

Development of Social Intelligence Revisited

From the knowledge view, with its emphasis on specific declarative and procedural social knowledge, the development of social intelligence is a matter of social learning rather than genetic endowment or mental maturation. Post-Piagetian views of cognitive development emphasize the child's construction and refinement, through experience, of various intuitive theories concerning the mind, physics, and biology (e.g., Gopnik, 2003, 2011). To this list we can add theories about self, others, and the social world – intuitive theories of personality, self, and society that capture our understanding of the way people interact and shape our interactions with other people.

Quo Vadis?

It is possible that the concept of social intelligence has outlived its usefulness and will be supplanted by emotional intelligence or some other variant on personal intelligence. Alternatively, it is possible that neuroscientific analyses will give new life to the study of social intelligence, as they promise to do in other areas of psychology. On the other hand, perhaps we should abandon the "ability" model of social intelligence completely, along with its psychometric emphasis on developing instruments for the measuring of individual differences in social competencies of various sorts – tests intended to rank people and on which some people must score high and others must score low. Instead of focusing on *how people compare*, perhaps we should focus on *what people know* and how they bring their social intelligence to bear on their interactions with other people, on the tasks life has set for them, and on the tasks they have set for themselves. In this way, we would honor the primary idea of the cognitive view of social interaction, which is that interpersonal behavior is intelligent, based on what the individual knows and believes – no matter how smart or stupid it may appear to other people.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Anderson, J. R. (1976). Language, memory, and thought. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bainbridge, W. S., Brent, E. E., Carley, K. M., Heise, D. R., Macy, M. W., Markovsky, B., & Skvoretz, J. (1994). Artificial social intelligence. *Annual Review of Sociology*, *20*, 407–436. https://doi.org/10.1146/annurev.so.20.080194.002203

Bandura, A. (1973). *Aggression: A social learning analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., & Walters, R. H. (1963). *Social learning and personality development*. New York: Holt, Rinehart and Winston.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.), *Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Lawrence Erlbaum Associates.

Barnes, M. L., & Sternberg, R. J. (1989). Social intelligence and the decoding of nonverbal cues. *Intelligence*, *13*, 263–287. https://doi.org/10.1016/0160-2896(89)90022-6

Breazeal, C. (2002). *Designing sociable robots*. Cambridge, MA: MIT Press.

Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual Review of Psychology*, *68*, 627–652. https://doi.org/10.1146/annurev-psych-010416-043958

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*, 123–152. http://dx.doi.org/10.1037/h0043805

Bruner, J. S., & Feldman, C. (1993). Theories of mind and the problem of autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 267–291). Oxford: Oxford University Press.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*(5), 187–192.

Campbell, J. M., & McCord, D. M. (1996). The WAIS-R Comprehension and Picture Arrangement subtests as measures of social intelligence: Testing traditional interpretations. *Journal of Psychoeducational Assessment*, *14*, 240–249. http://dx.doi.org/10.1177/073428299601400305

Cantor, N. (1990). From thought to behavior: "Having" and "doing" in the study of personality and cognition. *American Psychologist*, *45*, 735–750. http://dx.doi.org/10.1037/0003-066X.45.6.735

Cantor, N. (2000). Life task problem solving: Situational affordances and personal needs. In E. T. Higgins & A. Kruglanski (Eds.), *Motivational science: Social and personality perspectives* (pp. 100–110). New York: Psychology Press.

Cantor, N. (2003). Constructive cognition, personal goals, and the social embedding of personality. In L. G. Aspinwall & U. M. Staudinger (Eds.), *A psychology of human strengths: Fundamental directions and future directions for a positive psychology* (pp. 49–60). Washington, DC: American Psychological Association.

Cantor, N., & Fleeson, W. (1994). Social intelligence and intelligent goal pursuit: A cognitive slice of motivation. In W. D. Spaulding (Ed.), *Integrative views of motivation, cognition, and emotion: Nebraska Symposium on Motivation* (Vol. 41, pp. 125–180). Lincoln: University of Nebraska Press.

Cantor, N., & Harlow, R. (1994). Social intelligence and personality: Flexible life-task pursuit. In R. J. Sternberg & P. Ruzgis (Eds.), *Personality and intelligence* (pp. 137–168). Cambridge: Cambridge University Press.

Cantor, N., Kemmelmeier, M., Basten, J., & Prentice, D. A. (2002). Life-task pursuit in social groups: Balancing self-exploration and social integration. *Self and Identity*, *1*, 177–184. http://dx.doi.org/10.1080/152988602317319366

Cantor, N., & Kihlstrom, J. F. (1987). *Personality and social intelligence*. Englewood Cliffs, NJ: Prentice-Hall.

Cantor, N., & Kihlstrom, J. F. (1989). Social intelligence and cognitive assessments of personality. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 2, pp. 1–59). Hillsdale, NJ: Lawrence Erlbaum Associates.

Carlston, D. E. (Ed.). (2013). *The Oxford handbook of social cognition*. New York: Oxford University Press.

Chen, S. A., & Michael, W. B. (1993). First-order and higher-order factors of creative social intelligence within Guilford's structure-of-intellect model: A reanalysis of a Guilford data base. *Educational and Psychological Measurement*, 53, 619–641. http://dx.doi.org/10.1177/0013164493053003004

Conzelmann, K., Weis, S., & Süss, H. (2013). New findings about social intelligence: Development and application of the Magdeburg Test of Social Intelligence (MTSI). *Journal of Individual Differences*, *34*(3), 119–137. http://dx.doi.org/10.1027/1614-0001/a000106

Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin*, *52*, 177–193. http://dx.doi.org/10.1037/h0044919

Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society London: B. Biological Sciences*, *362*(1480), 679–704. http://dx.doi.org/10.1098/rstb.2006.2004

Davis, K., Christodoulou, J. A., Seider, S., & Gardner, H. (2011). The theory of multiple intelligences. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 485–503). New York: Cambridge University Press.

deWaal, F. B. M. (2016). Apes know what others believe: Understanding false beliefs is not unique to humans. *Science*, *354*(6308), 39–40.

deWaal, F. B. M., & Ferrari, P. F. (Eds.). (2012). *The primate mind: Built to connect with other minds*. Cambridge, MA: Harvard University Press.

Dewey, J. (1909). *Moral principles in education*. New York: Houghton Mifflin.

Doll, E. A. (1947). *Social maturity scale*. Circle Pines, MN: American Guidance Service.

Fiske, S. T., & Macrae, C. N. (Eds.). (2012). *SAGE handbook of social cognition*. Los Angeles: Sage.

Fiske, S. T., & Prentice, D. (Eds.). (2011). *Social neuroscience: Toward understanding the underpinnings of the social mind*. New York: Oxford University Press.

Fiske, S. T., & Taylor, S. E. (2007). *Social cognition: From brains to culture*. New York: McGraw-Hill.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.

Gardner, H. (1997). Is there a moral intelligence? In M. Runco (Ed.), *The creativity research handbook*. Cresskill, NJ: Hampton Press.

Gardner, H. (1999). Intelligence reframed: Multiple intelligences for the 21st century. New York: Basic Books.

Gardner, H. (2006). *Multiple intelligences: New horizons*. New York: Basic Books.

Gernsbacher, M. A. (2015). Diverse brains. *General Psychologist*, *49*(2), 29–37.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.

Goleman, D. (2006). Social intelligence: The new science of human relationships. New York: Bantam Books.

Gopnik, A. (2003). The theory theory as an alternative to the innateness hypothesis. In L. Antony & N. L. Hornstein (Eds.), *Chomsky and his critics*. New York: Basil Blackwell.

Gopnik, A. (2011). The Theory Theory 2.0: Probabilistic models and cognitive development. *Child Development Perspectives*, *5*(3), 161–163. http://dx.doi.org/10.1111/j.1750-8606.2011.00179.x

Greenspan, S., & Love, P. F. (1997). Social intelligence and developmental disorder: mental retardation, learning disabilities, and autism. In W. E. MacLean (Ed.), *Ellis' handbook of mental deficiency: Psychological theory and research* (3rd ed., pp. 311–342). Mahwah, NJ: Lawrence Erlbaum Associates.

Grieve, R. (2013). Can social intelligence be measured? Psychometric properties of the Tromsø Social Intelligence Scale – English Version. *Irish Journal of Psychology*, *34*(1), 1–12. https://doi.org/10.1080/03033910.2012.737758

Hendricks, M., Guilford, J. P., & Hoepfner, R. (1969). Measuring creative social intelligence. Reports from the Psychological Laboratory, University of Southern California, No. 42.

Hoepfner, R., & O'Sullivan, M. (1969). Social intelligence and IQ. *Educational and Psychological Measurement*, *28*, 339–344. http://dx.doi.org/10.1177/001316446802800211

Hunt, T. (1928). The measurement of social intelligence. *Journal of Applied Psychology*, 12, 317–334. http://dx.doi.org/10.1037/h0075832.

Jaarsma, P., & Welin, S. (2011). Autism as a natural human variation: Reflections on the claims of the neurodiversity movement. *Health Care Analysis*, *20*(1), 20–30. http://dx.doi.org/10.1007/s10728-011-0169-9

Jimenez, A. M., Gee, D. G., Cannon, T. D., & Lieberman, M. D. (2013). The social cognitive brain: A review of key individual difference parameters with relevance to schizophrenia. In D. L. Roberts & D. L. Penn (Eds.), *Social cognition in schizophrenia: From evidence to treatment* (pp. 93–119). New York: Oxford University Press.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*, 217–250. https://search.proquest.com/docview/615132040?accountid=14496

Kelly, G. (1955). *The psychology of personal constructs*. New York: W. W. Norton.

Kihlstrom, J. F. (2008). The automaticity juggernaut. In J. Baer, J. C. Kaufman, & R. F. Baumeister (Eds.), *Are we free? Psychology and free will* (pp. 155–180). New York: Oxford University Press.

Kihlstrom, J. F. (2009). "So that we might have roses in December": The functions of autobiographical memory. *Applied Cognitive Psychology*, *23*, 1179–1192. http://dx.doi.org/10.1002/acp.1618

Kihlstrom, J. F. (2010). Social neuroscience: The footprints of Phineas Gage. *Social Cognition*, *28*(6), 757–783. http://dx.doi.org/10.1521/soco.2010.28.6.757

Kihlstrom, J. F., Beer, J. S., & Klein, S. B. (2002). Self and identity as memory. In M. R. Leary & J. Tangney (Eds.), *Handbook of self and identity* (pp. 68–90). New York: Guilford Press.

Kihlstrom, J. F., & Cantor, N. (1989). Social intelligence and personality: There's room for growth. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 2, pp. 197–214). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kihlstrom, J. F., & Cantor, N. (2000). Social intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 359–379). New York: Cambridge University Press.

Kihlstrom, J. F., & Cantor, N. (2011). Social intelligence. In R. J. Sternberg (Ed.), *Cambridge handbook of intelligence* (pp. 564–581). New York: Cambridge University Press.

Klinger, E., & Cox, W. M. (2011). Motivation and the goal theory of current concerns. In W. M. Cox & E. Klinger (Eds.), *Handbook of motivational counseling: Goal-based approaches to assessment and intervention with addiction and other problems* (pp. 1–27). Chichester: Wiley.

Kohlberg, L. (1963). The development of children's orientations toward a moral order: I. Sequence in the development of moral thought. *Vita Humana*, 6, 11–33. https://search.proquest.com/docview/615428030?accountid=14496

Kosmitzki, C., & John, O. P. (1993). The implicit use of explicit conceptions of social intelligence. *Personality and Individual Differences*, *15*, 11–23. http://dx.doi.org/10.1016/0191-8869(93)90037-4

Landy, F. J. (2006). The long, frustrating and fruitless search for social intelligence: A cautionary tale. In K. R. Murphy (Ed.), *A critique of emotional intelligence: What are the problems and how can they be fixed?* (pp. 81–123). Mahwah, NJ: Lawrence Erlbaum Associates.

Lee, J.-E., Wong, C. T., Day, J. D., Maxwell, S., & Thorpe, S. (2000). Social and academic intelligences: A multitrait-multimethod study of their crystallized and fluid characteristics. *Personality and Individual Differences*, *29*, 539–553. http://dx.doi.org/10.1016/S0191-8869(99)00213-5

Lepore, J. (2018). It's still alive: Two hundred years of "Frankenstein." *New Yorker*, February 12–19.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259–289.

Little, B. R., Salmela-Aro, K., & Phillips, S. D. (Eds.). (2007). *Personal project pursuit: Goals, action, and human flourishing*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lull, H. G. (1911). Moral instruction through social intelligence. *American Journal of Sociology*, *17*, 47–60. http://dx.doi.org/10.1086/211944

Marlowe, H. A. (1986). Social intelligence: Evidence for multidimensionality and construct independence. *Journal of Educational Psychology*, *78*, 52–58. http://dx.doi.org/10.1037/0022-0663.78.1.52

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore, MD: Williams & Wilkins.

Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The ability model of emotional intelligence: Principles and updates. *Emotion Review*, *8*(4), 290–300. https://doi.org/10.1177%2F1754073916639667

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, *59*, 507–536. http://dx.doi.org/10.1146/annurev.psych.59.103006.093646

Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist*, *63*, 503–517. http://dx.doi.org/10.1037/0003-066X.63.6.503

McAdams, D. P., & Manczak, E. (2015). Personality and the life story. In M. Mikulincer, P. R. Shaver, M. L. Cooper & R. J. Larsen (Eds.), *APA handbook of personality and social psychology* (Vol. 4, *Personality processes and individual differences,* pp. 425–446). Washington, DC: American Psychological Association.

Miller, N. E., & Dollard, J. H. (1941). *Social learning and imitation*. New Haven, CT: Yale University Press.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, *80*, 252–283. http://dx.doi.org/10.1037/h0035002

Moss, F. A. (1931). Preliminary report of a study of social intelligence and executive ability. *Public Personnel Studies*, *9*, 2–9. https://search.proquest.com/docview/614975962?accountid=14496

Moss, F. A., & Hunt, T. (1927). Are you socially intelligent? *Scientific American*, *137*, 108–110. http://dx.doi.org/10.1038/scientificamerican0827-108

Murphy, K. R. (Ed.). (2006). *A critique of emotional intelligence: What are the problems and how can they be fixed?* Mahwah, NJ: Lawrence Erlbaum Associates.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings in social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

O'Sullivan, M., Guilford, J. P., & deMille, R. (1965). The measurement of social intelligence. Reports from the Psychological Laboratory, University of Southern California, No. 34.

Petrides, K. V., Mason, M., & Sevdalis, N. (2011). Preliminary validation of the construct of trait social intelligence. *Personality and Individual Differences*, *50*, 874–877. http://dx.doi.org/10.1016/j.paid.2010.12.029

Riggio, R. E., Messamer, J., & Throckmorton, B. (1991). Social and academic intelligence: Conceptually distinct but overlapping constructs. *Personality and Individual Differences*, *12*, 695–702. http://dx.doi.org/10.1016/0191-8869(91)90225-Z

Romney, D. M., & Pyryt, M. C. (1999). Guilford's concept of social intelligence revisited. *High Ability Studies*, *10*, 137–199. http://dx.doi.org/10.1080/1359813990100202

Rosenthal, R., Hall, J. A., DiMatteo, R., Rogers, P. L., & Archer, D. (1979). *Sensitivity to nonverbal communication: The PONS Test*. Baltimore, MD: Johns Hopkins University Press.

Rotter, J. B. (1954). *Social learning and clinical psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, *80*(1), 609. http://dx.doi.org/10.1037/h0092976

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality*, *9*, 185–211. http://dx.doi.org/10.2190/DUGG-P24E-52WK-6CDG

Schneider, R. J., Ackerman, P. L., & Kanfer, R. (1996). To "act wisely in human relations": Exploring the dimensions of social competence. *Personality and Individual Differences*, *21*, 469–482. http://dx.doi.org/10.1016/0191-8869(96)00084-0

Sechrest, L., & Jackson, D. N. (1961). Social intelligence and the accuracy of interpersonal predictions. *Journal of Personality*, *29*, 167–182. http://dx.doi.org/10.1111/j.1467-6494.1961.tb01653.x

Seyfarth, R. B., & Cheyney, D. L. (2015). Social cognition. *Animal Behaviour*, *103*, 191–202.

Shanley, L. A., Walker, R. E., & Foley, J. M. (1971). Social intelligence: A concept in search of data. *Psychological Reports*, *29*, 1123–1132. http://dx.doi.org/10.2466/pr0.1971.29.3f.1123

Shweder, R. (2017). Social intelligence in a multicultural world: What is it? Who needs it? How does it develop? In N. Budwig & E. Turiel (Eds.), *New perspectives on human development* (pp. 313–329). New York: Cambridge University Press.

Shweder, R., Goodnow, J., Hatano, G., LeVine, R., Markus, H., & Miller, P. (1998). The cultural psychology of development: One mind, many mentalities. In W. Damon (Ed.), *Handbook of child psychology* (pp. 716–792). New York: Wiley.

Silvera, D. H., Martinussen, M., & Dahl, T. I. (2001). The Tromsø social intelligence scale: A self-report measure of social intelligence. *Scandinavian Journal of Psychology*, *42*, 313–319. http://dx.doi.org/10.1111/1467-9450.00242

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118. http://dx.doi.org/10.2307/1884852

Smetana, J. G. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 119–153). Mahwah, NJ: Lawrence Erlbaum Associates.

Snyder, M., & Cantor, N. (1998). Understanding personality and social behavior: A functionalist strategy. In D. T. Gilbert & S. T. Fiske (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 635–679). Boston: McGraw-Hill.

Snyder, S. D., & Michael, W. B. (1983). The relationship between performance on standardized tests in mathematics and reading to two measures of social intelligence and one of academic self-esteem of primary school children. *Educational and Psychological Measurement*, *43*, 1141–1148. http://dx.doi.org/10.1177/001316448304300424

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scale*. Circle Pines, MN: American Guidance Service.

Spearman, C. (1927). *The abilities of man*. New York: Macmillan.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1988). *The triarchic mind: A new theory of intelligence*. New York: Viking.

Sternberg, R. J. (2018). Successful intelligence in theory, research, and practice. In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 308–321). New York: Cambridge University Press.

Sternberg, R. J. Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, *41*, 37–55. http://dx.doi.org/10.1037/0022-3514.41.1.37

Sternberg, R. J., & Wagner, R. (Eds.). (1986). *Practical intelligence: Nature and origins of competence in the everyday world*. Cambridge: Cambridge University Press.

Tager-Flusberg, H. (2007). Evaluating the theory-of-mind theory of autism. *Current Directions in Psychological Science*, *16*, 311–315. http://dx.doi.org/10.1111/j.1467-8721.2007.00527.x

Taylor, E. H. (1990). The assessment of social intelligence. *Psychotherapy*, *27*, 445–457. http://dx.doi.org/10.1037/0033-3204.27.3.445

Thorndike, E. L. (1920). Intelligence and its uses. *Harper's Magazine*, *140*, 227–235. https://search.proquest.com/docview/614886569?accountid=14496

Thorndike, R. L., & Stein, S. (1937). An evaluation of the attempts to measure social intelligence. *Psychological Bulletin*, *34*, 275–285. http://dx.doi.org/10.1037/h0053850

Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.

Turiel, E. (2006). The development of morality. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social emotional, and personality development*, Vol.3 (6th ed., pp. 789–857). Hoboken, NJ: Wiley.

Turiel, E., Killen, M., & Helwig, C. (1987). Morality: Its structure, functions, and vagaries. In J. Kagan & M. Lamb (Eds.), *The emergence of morality in young children*. Chicago, IL: University of Chicago Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. http://dx.doi.org/10.1126/science.185.4157.1124

Vernon, P. E. (1933). Some characteristics of the good judge of personality. *Journal of Social Psychology*, *4*, 42–57. http://dx.doi.org/10.1080/00224545.1933.9921556

Walker, R. E., & Foley, J. M. (1973). Social intelligence: Its history and measurement. *Psychological Reports*, *33*, 839–864. http://dx.doi.org/10.2466/pr0.1973.33.3.839

Wechsler, D. (1939). *The measurement and appraisal of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore, MD: Williams & Wilkins.

Weis, S., & Suss, H.-M. (2007). Reviving the search for social intelligence: A multitrait-multimethod study of its structure and construct validity. *Personality and Individual Differences*, *42*(1), 3–14. http://dx.doi.org/10.1016/j.paid.2006.04.027

Whiten, A., & van de Waal, E. (2017). Social learning, culture and the "socio-cultural brain" of human and non-human primates. *Neuroscience and Biobehavioral Reviews*, *82*, 58–75. http://dx.doi.org/10.1016/j.neubiorev.2016.12.018

Winograd, T. (1975). Frame representations and the procedural-declarative controversy. In D. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 185–210). New York: Academic Press.

32 Collective Intelligence

Thomas W. Malone and Anita Williams Woolley

The other kinds of intelligence described in this book all happen inside a single human brain. But this single-brain intelligence is not the only kind of intelligence on our planet. It would not be surprising to say that animals have intelligence and many people would say that computers also have a kind of (artificial) intelligence. Yet it is surprising to many people to realize that there is another important kind of intelligence on our planet: the *collective intelligence* that arises in *groups* of individuals. For instance, one could say that teams, families, companies, countries, economies, and scientific communities all have different kinds of collective intelligence. And, even though collective stupidity is just as possible as collective intelligence, these collectively intelligent groups are often the most intelligent entities on our planet.

The phenomenon of collective intelligence includes analogies at the group level of many of the aspects of individual intelligence that are described in other chapters of this volume. This way of viewing the world provides an opportunity for applying psychological concepts at a different level and bringing concepts from other fields into psychology.

What Is Collective Intelligence?

As with many important – but evocative – terms, there have been almost as many definitions of collective intelligence as there have been writers who have described it (for a representative list, see Malone & Bernstein, 2015, p. 10). For instance, Hiltz and Turoff (1978) defined collective intelligence as “a collective decision capability [that is] at least as good as or better than any single member of the group” (p. 44). Smith (1994) defined it as “a group of human beings [carrying] out a task as if the group, itself, were a coherent, intelligent organism working with one mind, rather than a collection of independent agents” (p. 1). And Levy (1997) defined it as “a form of *universally distributed intelligence*, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills” (p. 13).

Perhaps the simplest – and broadest – definition of collective intelligence comes from Malone and Bernstein (2015):

> *Collective intelligence – groups of individuals acting collectively in ways that seem intelligent.* (p. 3, emphasis in original)

As described by Malone (2018), this definition includes almost all human groups, such as hierarchies, markets, democracies, communities, and ecosystems. This is the definition on which we will focus here and several aspects of this definition are worth noting:

(1) The definition does not try to define "intelligence" itself. Thus, this definition is compatible with all of the other definitions of intelligence in this volume.
(2) By using the word "acting," the definition requires intelligence to be manifested in some kind of behavior. By this definition, for instance, a Wikipedia article would not, itself, be considered intelligent but the group of people who created it would be.
(3) The definition requires that, in order to analyze something as collective intelligence, one must specify a *group* of *individuals* that are involved. In some cases, this may be straightforward, such as noting the individual humans in a company, but, in other cases, it may be useful to draw these boundaries in unusual ways. For instance, one could analyze the collective intelligence of a whole economy by noting that the economy is a collection of many different organizations and people.
(4) The definition requires that the individuals act *collectively*, that is, that there be some connection among their activities. For instance, two unrelated people in two different cities, each making coffee on the same morning, are probably not an example of collective intelligence. But two servers working together to fill all the customer orders in a single coffee shop would be. It is important to note, however, that, even though the individuals' actions need to be connected, the individuals do *not* need to cooperate with each other or have the same goals. For instance, different actors in a market buy and sell things to others in the same market, and thus their actions are connected, but they may each have very different individual goals.
(5) Finally, by using the word "seem," the definition emphasizes that what is considered intelligent depends on the perspective of the observer. For instance, to evaluate whether an entity is acting intelligently, an observer needs to make assumptions about what the entity's goals are. At the individual level, when students take intelligence tests, we assume they are trying to give the answers the test designers consider correct. But if the goal of a student taking the test is to annoy the test-giver, then the test score will not be a good measure of that student's intelligence!

At the level of a group, it is usually even more important for an observer to attribute goals to the group. For instance, when analyzing the collective intelligence of an economy, it is often useful to evaluate how effectively the economy allocates societal resources, even if none of the individuals in the economy has that goal.

History of Studying Collective Intelligence

The earliest scholarly article we have found with the phrase "collective intelligence" in the title was by David Wechsler, the psychologist who developed some of the most widely used IQ tests (Wechsler, 1971). This article argues that collective intelligence is more than just collective behavior in that it involves cross-fertilization resulting in something that could not have been produced by individuals. About this same time, computer scientist Doug Engelbart was doing pioneering work on "augmenting human intellect" with computers, including computational support for team cooperation (Engelbart, 1962, p. 105; Engelbart & English, 1968). Later, Engelbart used the phrase "collective IQ" to describe this work and its broader implications (e.g., Engelbart, 1995).

In 1978, Roxanne Hiltz and Murray Turoff used the term "collective intelligence" to describe the goal of the computerized conferencing systems they pioneered (Hiltz & Turoff, 1978). In the 1980s and 1990s, the term collective intelligence began to be used more and more to describe phenomena from insect behavior (e.g., Franks, 1989) to groups of mobile robots (Mataric, 1993) to human groups (e.g., Atlee & Por, 2000; Isaacs, 1999) to electronically mediated human collaboration (e.g., Heylighen, 1999; Levy, 1997; Smith, 1994). This period also saw early studies using the closely related term "group intelligence" (Williams & Sternberg, 1988).

In the years 2000–2010, the term "collective intelligence" became even more widely used in publications from computer science to spirituality to business (e.g., Hamilton, 2004; Howe, 2008; Szuba, 2001). Of particular importance to the spread of the concept was a bestselling book on *The Wisdom of Crowds* (Surowiecki, 2004) and other books for a general audience featuring the concept of collective intelligence (e.g., Ridley, 2010; Tapscott & Williams, 2006).

This period also saw the first academic conferences on collective intelligence (Bastiaens, Baumol, & Kramer, 2010; Kowalczyk, 2009; Malone & von Ahn, 2012) and the first academic research centers focusing specifically on this topic (Canada Research Chair in Collective Intelligence, University of Ottawa, started in 2002; Center for Collective Intelligence, Massachusetts Institute of Technology (MIT), started in 2006).

In more recent years, an annual conference series on collective intelligence was established (ACM, 2018), a handbook of collective intelligence was published (Malone & Bernstein, 2015), and additional books describing the field for general audiences appeared (Malone, 2018; Mulgan, 2017).

How Does Collective Intelligence Relate to Other Fields?

As we will see in the rest of this chapter, the interdisciplinary field of collective intelligence offers an opportunity to apply concepts from psychology, such as memory, learning, and perception, at the level of groups rather than to individual humans. The field also draws on concepts from other fields, such as economics, sociology, political science, and organization theory, that study groups (see Malone & Bernstein, 2015).

The overlap occurs when there is a focus on overall collective behavior that can be regarded as more or less intelligent. For instance, analyzing how individual people's attitudes are determined or how they make economic choices would not – alone – be central to collective intelligence. But analyzing how different regulatory mechanisms in markets lead to more or less intelligent behavior by the markets as a whole would be central to collective intelligence.

The field of collective intelligence also overlaps with other fields, such as computer science (which uses terms such as crowdsourcing, human computation, and computer-supported cooperative work to study groups of people and computers; Bigham, Bernstein, & Adar, 2015), network science (which studies how aspects of network structure affect the performance of networks; Easley & Kleinberg, 2010), and biology (which studies how groups of animals, such as bees and ants, interact to produce overall behavior that is adaptive for the group; Gordon, 2015).

Elements of Collective Intelligence

One useful way of analyzing collective intelligence is to consider the key elements needed for any group to act intelligently. As Figure 32.1 suggests, we can analyze any intelligent group using four questions: *What* is being done? *Who* is doing it? *Why* are they doing it? And *how* are they doing it? The *how* question can, in turn, be divided into two kinds of processes, *cognitive processes* and *coordination processes*. We will use these categories as an organizing framework for the rest of this chapter.

What

Since the definition of collective intelligence emphasized earlier in this chapter (see the section "What Is Collective Intelligence") leaves intelligence itself undefined, any discussion of collective intelligence requires some – explicit or implicit – definition of intelligence. Even though, as amply illustrated in other

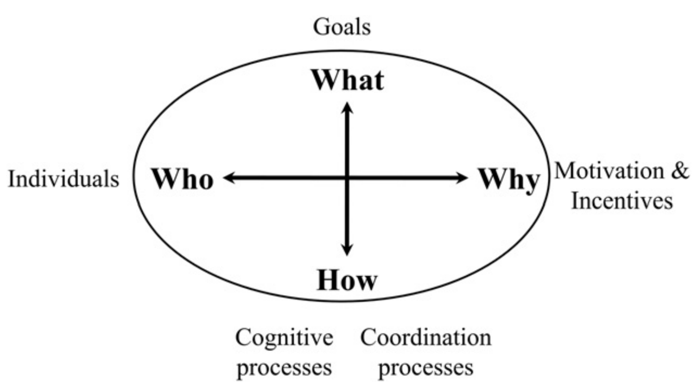

Figure 32.1 *Elements of collectively intelligent systems (after Malone, Laubacher, & Dellarocas, 2010; for a related framework, see Galbraith, 2002).*

chapters of this volume, there are many possible ways to define intelligence, Malone (2018) articulates a useful distinction that is often made implicitly between two broad kinds of intelligence:

- Specialized intelligence – the ability to achieve specific goals effectively in a given environment.
- General intelligence – the ability to achieve a wide range of different goals effectively in different environments. (p. 20)

Using these definitions, specialized intelligence would include, for example, Howard Gardner's definition of intelligence as "the ability to solve problems, or to create products, that are valued within one or more cultural settings" (Gardner, 1983); and general intelligence would include the view of intelligence first identified by Spearman (1904). Spearman called the factor he identified "*g*," for "general intelligence," and this factor is what most modern intelligence tests measure (e.g., Deary, 2000).

These two definitions, in turn, help distinguish two kinds of collective intelligence – *specialized collective intelligence* and *general collective intelligence* – and this provides a useful way of summarizing various results about collective intelligence.

Specialized Collective Intelligence

Much of the literature on teams, organizations, and other groups can be viewed as studies of specialized collective intelligence because it involves groups performing specific tasks. For example, taxonomies of tasks widely used in research and elsewhere have been developed by Steiner (1966, 1972) and McGrath (1984). In these taxonomies, tasks are characterized by the nature of the processes group members must engage in to carry them out effectively (Larson, 2009; McGrath, 1984). For instance, McGrath's task circumplex (1984) identifies four task categories that reflect different sets of team interaction processes: (1) *generate* tasks that require idea generation and divergent thinking; (2) *choose* tasks or decision-making tasks that require selecting among specified alternatives; (3) *negotiate* tasks involving resolution of conflicts of interest or viewpoints; and (4) *execute* tasks involving a high level of coordination, physical movement, or dexterity, to produce a correct or optimal solution.

The type of task a group is faced with has important implications for many other facets of the group to be discussed in the remainder of this chapter, including group composition, incentives, and process.

General Collective Intelligence

It appears that, until recently, no one had asked the question of whether general collective intelligence even exists, that is, whether there is an equivalent of Spearman's *g* for groups: a single statistical factor that predicts how well the group will perform a wide range of tasks. To answer this question, Woolley and colleagues (2010) sampled across the task types described by McGrath (1984) to assess whether groups exhibit a general ability to perform across the full range of different tasks.

In a factor analysis of all the groups' scores, Woolley and colleagues (2010) found that the first factor accounted for 43 percent of the variance in performance on all of the different tasks. This is comparable to the 30–50 percent of variance typically explained by the first factor in a battery of individual cognitive tasks (Chabris, 2007) and suggests the existence of a general collective intelligence factor for groups. Woolley and colleagues (2010) called this factor *c*, by analogy to Spearman's *g*-factor in individual intelligence. This finding has since been replicated in a number of studies and a recent meta-analysis of data from more than 1,000 teams provides further support for the conclusion that a general collective intelligence factor exists in teams (Woolley et al., 2017).

In response to these findings, some have questioned whether the evidence for a general collective intelligence factor is strong enough to conclude that there is a general factor (Credé & Howardson, 2017) as well as whether the factor depends largely on the intelligence of individual team members (Bates & Gupta, 2016). However, the evidence to date supports the conclusion that a general collective intelligence factor exists and suggests opportunities for research to further refine its measurement (Woolley, Kim, & Malone, 2018).

Collective Intelligence and Collective Stupidity

Just as individual humans vary greatly in how intelligent they are, so, too, do groups. Early literature on this topic, for example, often focused on how groups were less intelligent than the individuals in them (e.g., Janis & Mann, 1977; Mackay, 1841) and on the limitations of group decision-making (e.g., Condorcet, 1785/1976). Very recent research has also shown how false stories ("fake news") spread further and faster in social media networks than true ones (Vosoughi, Roy, & Aral, 2018) and many people have speculated that phenomena like this are influencing the outcomes of democratic elections today (e.g., Chakrabarti, 2018).

Yet it has been more common in recent decades to focus on how groups can be smarter than the individuals in them (e.g., Hill, 1982; Surowiecki, 2004). Groups, for instance, can bring more resources to bear on problems (Hill, 1982) and more diversity of perspectives and knowledge (e.g., Page, 2008). Most of the prominent examples of collective stupidity occur when groups block alternative perspectives or bias how those perspectives are considered (e.g., Janis & Mann, 1977); but when groups bring together diverse perspectives and integrate them effectively, the whole can be well more than the sum of the parts (e.g., Bernstein, Shore & Lazer, 2018; Hong & Page, 2004).

Who

The collective intelligence of a group is clearly affected by the characteristics of the group members, that is, by the group composition (e.g., Mann, 1959). Here, we consider three particularly important aspects of group composition: (1) members' *task-relevant abilities*, (2) members' *interpersonal abilities* that help them work together effectively, and (3) the *diversity* of group members on these or other dimensions.

Task-Relevant Abilities

It is clear that, to solve a problem effectively, a group's members should – together – have all the task-relevant knowledge and abilities needed to solve the problem (Hackman, 1987). Research demonstrates, for example, that individual member intelligence maintains a strong relationship with group performance, particularly on specific kinds of tasks, such as decision-making tasks (Devine & Philips, 2001); and Woolley and colleagues (2010) find a moderate correlation between individual intelligence and general collective intelligence. Yet for many teams in a variety of circumstances, task-relevant abilities are necessary but not sufficient for collective intelligence to emerge.

Interpersonal Abilities

Even if a group's members have all the task-relevant knowledge needed, they can still be very ineffective – and thus have very low collective intelligence – if they do not also have the ability to work together well. For instance, there is a general consensus that emotional intelligence (see Chapter 29 in this volume) enhances group performance (Druskat & Wolff, 2001), at least in the short term (Ashkanasy & Daus, 2005). And, even though they did not use these terms, early research by Williams and Sternberg (1988) suggests that when individual group members have other types of intelligence discussed in this volume (such as social intelligence and leadership intelligence) that could increase the performance, and thus the collective intelligence, of a group.

A specific subset of these social intelligence skills, related to the perception of emotions and mental states, has been extensively studied under the term "theory of mind" (ToM; Baron-Cohen et al., 2001; Premack & Woodruff, 1978). A common assumption in much of this research is that people with greater ToM abilities will be more competent at various kinds of social interaction and this has been shown to be true with both children (Watson et al., 1999) and adults (Krych-Appelbaum et al., 2007; Meslec, Aggarwal, & Curşeu, 2016; Woolley et al., 2010).

For instance, Woolley and colleagues (2010) found that groups whose members had higher average ToM scores (as measured by the "Reading the Mind in the Eyes" (RME) Test; Baron-Cohen et al., 2001) also had significantly higher collective intelligence. Indeed, average ToM scores remained the only significant predictor of collective intelligence even when controlling for individual intelligence or other group composition or process variables, such as the proportion of women in the group or the distribution of communication. RME has also predicted performance in teams working together online (Engel et al., 2014) and in groups playing online video games over a period of months (Kim et al., 2017).

Diversity

Diversity of group composition is one of the most commonly studied team variables (e.g., van Knippenberg & Mell, 2016). Despite its potential value, however, a number of studies and meta analyses have failed to show strong effects of diversity on team performance (Joshi & Roh, 2009). Scholars have, therefore, urged researchers to pay close attention to the type of diversity variable studied. It may be critical, for example, to examine the specific type of diversity that is most relevant to the outcomes being investigated (Harrison & Klein, 2007; Horwitz & Horwitz, 2007; Joshi & Roh, 2009; Milliken & Martins, 1996).

With regard to group composition, groups performing tasks that benefit from a range of skills or expertise will underperform unless composed with the requisite cognitive diversity (Woolley et al., 2007; Woolley et al., 2008) even when compared to groups of higher general intelligence or ability (Hong & Page, 2004). Groups that are too homogeneous will also be less creative than more cognitively diverse groups (Aggarwal & Woolley, 2019) and exhibit lower levels of collective intelligence than moderately cognitively diverse groups (Aggarwal et al., 2019). However, groups that are too cognitively diverse run the risk of making costly errors, particularly when the diversity leads them to have difficulties communicating about how to prioritize task elements (Aggarwal & Woolley, 2013). Thus many researchers focus on the moderating effects of group process, such as the development of transactive memory systems and strategic consensus, in examining the relationship between diversity and performance (Aggarwal & Woolley, 2019).

One particularly intriguing kind of diversity in groups comes from using computers, not just to connect group members to each other but as participants in the groups themselves (Kim et al., 2018; Malone, 2018; Weld et al., 2015).

Why

Another key factor in determining the collective intelligence of a group is the motivation of the group members. The literature has generally looked at two sources of motivation – *extrinsic motivation*, often in the form of money or cash incentives, and *intrinsic motivation*, derived from the internal satisfaction associated with the work itself.

Monetary incentives are the core foundation to induce high levels of effort in traditional organizational and market settings (Lazear, 2000; Prendergast, 1999). At times, they have been shown to increase the quantity but not the quality of work produced (Jenkins et al., 1998). The use of group-level monetary incentives can be tricky, as group-based incentives are highly subject to free riding unless accompanied by highly cooperative work behavior (Alchian & Demsetz, 1972; Wageman & Baker, 1997).

When it is difficult to identify and reward the exact contribution made by each worker in a team, the workers will typically lack incentives to provide the optimal level of effort and thus they will work less than if they were working alone. This has also been referred to as the "moral hazard" problem – and suggests that collaboration, particularly by anonymous workers outside of an employment relationship, can produce moral hazard (Holmstrom, 1982) and social loafing (Latane, Williams, & Harkins, 1979). Recommendations for avoiding this include making individual contributions identifiable and encouraging intrinsic motivation by making the work personally meaningful (Benkler, Shaw, and Hill, 2015; Deci & Ryan, 1985; Hackman & Oldham, 1976).

How – Cognitive Processes

The key cognitive processes needed for collective intelligence are those that are needed for any kind of intelligent system, whether it is a human, a machine, or a group. One useful way of classifying these processes is shown in Figure 32.2. Working backward from the action an intelligent system takes, the figure shows that to act effectively in the world, the system needs to *decide* what action to take. To do this, the system needs to somehow *create* possible options for action. This combination of creating and deciding can be called problem-solving, and to do effective problem-solving a system usually needs to also *sense* the external world and *remember* relevant information. Finally, to do all these things better over time, the system needs to *learn* from its own experience. We'll consider the different cognitive processes in this order, emphasizing results from social and organizational psychology but also including selected examples from other disciplines.

Decide

The ability of groups to make decisions effectively – that is, to share relevant details, weight information appropriately, and arrive at the best decision – is directly tied to team performance (Mesmer-Magnus & DeChurch, 2009). For instance, groups frequently fail to surface relevant information and combine it appropriately; instead, they disregard relevant information while basing their decisions on irrelevant

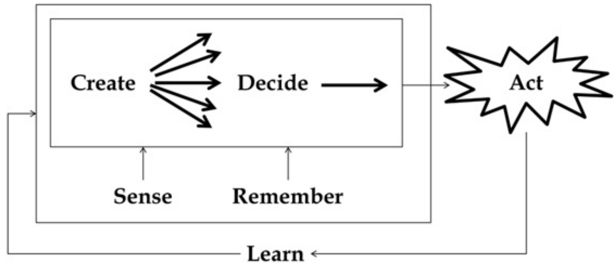

Figure 32.2 *Cognitive processes used by intelligent systems (from Malone, 2018).*

information (Larson, 2009). This can occur as a result of cognitive factors, motivational factors, or as a consequence of bad process.

Cognitive factors. A long line of work on social decision schemes has investigated how predecision preferences of individuals combine to influence a joint decision (Davis, 1973). Groups are also more likely than individual decision-makers to use certain cognitive heuristics and biases (Kerr, MacCoun, & Kramer, 1996). For instance, when there are "hidden profiles," in which members initially prefer different alternatives based on conflicting information they hold, they may need to make a special effort to surface and share all the information they need to reach the correct solution (Stasser & Titus, 1985).

Motivational factors. Motivational approaches to group decision-making focus on group members' motivation to share the information they have, to overlook disconfirming evidence, and to believe in the infallibility of their own group. For instance, work on groupthink, social comparison, and intragroup competition examines various aspects of these motivational issues (Isenberg, 1986; Janis & Mann, 1977; Sanders & Baron, 1977; Toma and Butera, 2009).

Process design factors. The benefits of using collections of *independent* decision-makers have been repeatedly shown in studies where the average of many individuals' estimates is often closer to the true value than almost all of the individual guesses (Galton, 1907; Surowiecki, 2004). However, for this to happen, the individual estimates must be independent of one another and the sample sufficiently large and unbiased to enable errors to be symmetrically distributed (Steyvers & Miller, 2015; Surowiecki, 2004). Even subtle social influence revealing knowledge of others' estimates can create a cascade of effects that reduces the accuracy of crowds (Lorenz et al., 2011).

While independent decision-makers can be useful for some types of decisions, interacting groups are often better when the options are not well defined or when the group needs to buy into a decision for it to be implemented. In these circumstances, outcomes can often be improved by having the group identify key questions to be answered and how to integrate their information to answer those questions (i.e., Woolley et al., 2008).

Another type of intervention involves encouraging a group to grant equal speaking time to all group members on the assumption that this will enable more relevant facts to be brought into the discussion. Equality in speaking time has been associated with higher collective intelligence in groups (Engel et al., 2014; Woolley et al., 2010) and interventions involving real-time feedback on relative contributions to group conversation have also been shown to improve group decision-making performance (DiMicco, Pandolfo, & Bender, 2004).

More generally, there are many ways to organize group decision-making processes beyond the types typically studied by psychologists. For example, Malone (2018) identifies five common types of group decision-making processes as reflected across a number of disciplines: *hierarchies* (in which decisions are made by those with delegated authority), *democracies* (in which decisions are made by voting), *markets* (in which decisions are made by pairwise agreements between buyers and sellers), *communities* (in which decisions are made by informal consensus based on shared norms and reputations), and *ecosystems* (in which decisions are made based on power: the law of the jungle and survival of the fittest).

We believe there are interesting research opportunities to systematically study and compare all these different group decision-making methods both empirically and analytically. Economists, for instance, have shown mathematically that the combination of pairwise agreements between buyers and sellers in markets leads, in specified conditions, to overall group decisions about how resources are allocated that are optimal, in the sense that there is no way they could be improved without making at least some individuals worse off (Arrow & Debreu, 1954; Lo, 2015). And much recent work in behavioral economics has focused on the empirical study of how actual human decision-making departs from the ideal of completely rational decision-making that is assumed by most traditional economic models (e.g., Brennan and Lo, 2012; Camerer, Loewenstein, and Rabin, 2004; Kahneman & Tversky, 1979). What analogous kinds of work could be done for the other group decision-making processes listed above and for comparisons among them?

Create

To make a decision, a group needs to create (implicitly or explicitly) options from which to choose. An effective process for "creating" often involves engaging in divergent thought, where group members bring in as wide a variety of options as possible. However, research on group creativity has repeatedly found that, on average, groups perform worse on brainstorming tasks than nominal groups with the same number of individuals (Girotra, Terwiesch, & Ulrich, 2010; Stroebe, Nijstad, & Rietzschel, 2010). Subsequent research has found, however, that altering the structure of a brainstorming group can enhance creative performance (Girotra, Terwiesch, & Ulrich, 2010; Korde & Paulus, 2017). For instance, early work on the hybrid group format in brainstorming overcomes group process obstacles (i.e., production blocking, social loafing, evaluation apprehension) by employing an individual phase prior to the group phase (Girotra, Terwiesch, & Ulrich, 2010; Larson, 2009; Stroebe, Nijstad, & Rietzschel, 2010).

In addition to small group creativity studies like these, it is important to realize that these sorts of divergent processes underlie creation in larger collectives as well. For instance, product design groups in hierarchical companies delegate different parts of the design process to different individuals and teams (Ulrich & Eppinger, 2015) and markets allow different companies to develop different options from which buyers collectively choose the options that will succeed.

Sense

In order to create and decide intelligently, groups usually need to sense their environment. This happens, for example, when a medical treatment team diagnoses a patient's condition, when voters in a democracy see news stories about what politicians say, and when governments collect and analyze information to assess whether a foreign nation is developing biological weapons. These processes of collective sensing have been less studied by researchers than deciding and creating; but one of the main terms researchers have used to study them is *sensemaking* (Pirolli & Russell, 2011; Weick, 1995; Weick, Sutcliffe, & Obstfeld, 2005) For example, Weick (1993) uses this perspective to analyze a famous disaster in which thirteen members of a firefighting crew died, in part because their expectations that they were fighting a small and easily controlled fire led them to misinterpret what they were seeing. In general, the careful structuring of collective attention lays the groundwork for collectives to notice and respond to critical issues, or to overlook, miss, and underestimate them (Hackman, 2011; Ocasio, 1997).

Even though the allocation of *attention* plays a role in all the cognitive processes, it is central to sensing the environment and determining what stimuli or events require response. Work on attention at the organizational level started with Simon (1947) who examined the channeling, structuring, and allocation of attention as a central concept in studying administrative behavior. Ocasio (1997), in his attention-based theory of the firm, focused on how attention in organizations shapes organizational adaptation. Newer lines of work examine "attentional selection," the development of shared attention in groups (Ocasio, 2011), and ask: What do collectives make the center of their focus? And what do they allow to fall by the wayside?

For instance, teams that focus on *outcomes* tend to produce more innovative or creative outcomes and adapt more effectively to difficulties that arise in their work (Woolley, 2009a, 2009b), while teams that focus on *processes* commit fewer errors (Aggarwal & Woolley, 2013). In other words, where teams focus their attention leads them to sense different issues (while ignoring others) and respond differently. And the degree to which group members agree about the team's strategic priorities can strongly affect how creative they are and the quality of the team's outcomes (Aggarwal & Woolley, 2013, 2019).

Remember

Groups remember things in a variety of ways: in their written records and online systems, in their habitual organizational routines (Nelson & Winter, 1982), in their group norms, and, of course, in the minds of their members. One useful way of thinking about how a group's memory emerges from the individual memories of its members is called transactive memory systems (Wegner, 1987). A transactive memory system (TMS) refers to a shared system that individuals in groups develop to collectively encode, store, and retrieve information or knowledge in different domains (Argote & Ren, 2012; Hollingshead, 2001; Lewis & Herndon, 2011). Groups with a well-developed TMS can efficiently store and make use of a broader range of knowledge than groups without an effective TMS. According to TMS theory, there are three behavioral indicators of an effective TMS: *specialization, credibility*, and *coordination* (Lewis, 2004).

Through performing tasks and answering questions, a member establishes credibility and expertise status. Other members, being aware of the person's expertise, direct new knowledge in the domain to him or her, which reinforces the person's specialization and team members' trust in his or her expertise. Further, members know whom to count on for performing various tasks and whom to consult for information in particular domains, which improves coordination (Argote & Ren, 2012). Dozens of studies have demonstrated the positive effects of TMS on group performance in both laboratory and field settings (Lewis & Herndon, 2011), though work continues to refine measures and conceptualization of the construct and its relationship to performance for different types of tasks (Lewis & Herndon, 2011).

Learn

A great deal of evidence suggests that groups and organizations vary enormously in their ability to learn. The performance of some organizations improves dramatically with experience while the performance of others remains unchanged or even deteriorates (Argote, 1999).

In general, group learning refers to changes in a group – including changes in cognitions, routines, or performance – that occur as a function of experience (Argote & Miron-Spektor, 2011; Fiol & Lyles, 1985). For example, as groups gain experience, they may acquire information about which group members are good at which tasks, how to use a new piece of technology more effectively, or how to coordinate their activities better. This knowledge may in turn improve their performance (Argote, 1999).

It is sometimes useful to distinguish between two kinds of group learning: (1) changes in *knowledge* (which may be gauged from change in performance) and (2) changes in *group processes* or repertoires (Argote & Miron-Spektor, 2011; Edmondson, 1999; Fiol & Lyles, 1985; Wilson, Goodman, & Cronin, 2007). It is also important to realize that groups may learn (e.g., change processes) without any change in performance and they may change performance (e.g., because of changes in the environment) without any corresponding change in the group's knowledge (Argote, 1999). And sometimes knowledge may be *explicit* (easily codifiable and observable; Kogut & Zander, 1992) while at other times it may be only *tacit* (unarticulated and difficult to communicate; Nonaka, 1994).

An organization's overall ability to learn productively – that is, to improve its outcomes through better knowledge and insight (Fiol & Lyles, 1985) – depends on the ability of its teams to learn (Edmondson, 1999; Roloff, Woolley, & Edmondson, 2011; Senge & Sterman, 1992). Much of the work on group learning uses the concept of learning curves originally developed in individual psychology (Ebbinghaus, 1885; Thorndike, 1898) to characterize the rate of improvement and researchers have found considerable variation in this rate for different groups (Argote & Epple, 1990; Dutton & Thomas, 1984; Knott, 2008).

In addition to hierarchical organizations, of course, learning also occurs in other kinds of collectively intelligent groups. Markets, for example, learn as sellers continually try to provide products that will be more desirable to customers than their competitors' products and the sellers that fail to do this well go out of business. At an even more general level, different kinds of collectively intelligent systems are always competing with each other for power and survival in the ecosystem of collectively intelligent systems (Malone, 2018, chap. 10).

How – Coordination Processes

Whenever activities are carried out by a group of individuals instead of a single individual, the individuals' activities need to somehow be coordinated with each other. This coordination can usefully be viewed as managing interdependencies among activities (Faraj & Xiao, 2006; Malone, Laubacher, & Dellarocas, 2010; Okhuysen & Bechky, 2009).

For instance, Thompson (1967) identified three types of interdependencies among activities: *pooled* (where, for instance, activities share a resource such as money or machine time), *sequential* (where resources from one activity flow to another one), and *reciprocal* (where resources flow back and forth between two or more activities). Thompson and later researchers (Malone et al., 1999; Van de Ven, Delbecq, & Koenig, 1976) showed how different kinds of coordination processes are appropriate for different kinds of interdependencies. For instance, pooled (or "shared resource") dependencies can be managed by coordination processes such as "first come-first served," priority order, budgets, managerial decision, or market-like bidding (Malone et al., 1999).

More generally, each of the different types of decision-making structures discussed by Malone (2018) has different implications for how activities can be coordinated in systems that use them. For instance, effective hierarchies usually need to *differentiate*, that is, to divide the overall goal of the organization into different tasks done by different subgroups and also to *integrate* the different subgroups using mechanisms such as informal communication and integrating managers (Galbraith, 2002; Lawrence & Lorsch, 1967).

It is also important to realize that the coordination processes associated with the different decision-making methods have different advantages and disadvantages (Malone, 2018, chap. 11). For instance, organizational theorists and economists have analyzed how the choice of whether it is better to organize activities using hierarchies or markets depends on factors like the transaction costs and incompleteness of contracts under the different arrangements (Hart, 1995; Williamson, 1973, 1981). And new information technologies, by changing the costs of coordination and other activities, can change these tradeoffs and enable new forms of organizational coordination (Malone, 2004; Malone, Yates, & Benjamin, 1987; Malone, Laubacher, & Dellarocas, 2010).

Conclusions

Just as studying the neural basis for intelligence provides a link between cognitive psychology and lower level neural processes, studying collective intelligence provides a link between cognitive psychology and higher-level social, organizational, and economic processes. We have summarized here some of the work so far on these topics and we believe there are substantial opportunities for further research in this area. To the degree that we can better understand the collectively intelligent groups that surround us all the time, we may also be better able to design and use them for solving our most important problems.

Acknowledgments

The work reported in this chapter has been supported by several sponsors over the years, including the US National Science Foundation (grant numbers IIS-0963285, ACI-1322254, IIS-0963451, IIS-1442887, IIS-1047567), the US Army Research Office (grant numbers 56692-MA, 64079-NS, W911NF-15–1-0577, and W911NF-13–1-0422), and other sponsors of the MIT Center for Collective Intelligence.

Portions of this chapter appeared previously in the chapters titled “Introduction,” “Collective intelligence in teams and organizations,” and “Conclusion” in T. W. Malone and M. S. Bernstein (Eds.) (2015), *Handbook of Collective Intelligence.* Cambridge, MA: MIT Press (copyright © 2015 Massachusetts Institute of Technology).

References

ACM (Association for Computing Machinery). (2018). Collective Intelligence Conference Series. https://ci.acm.org/

Aggarwal, I., & Woolley, A. W. (2013). Do you see what I see? The effect of members’ cognitive styles on team processes and performance. *Organizational Behavior and Human Decision Processes*, *122*, 92–99.

Aggarwal, I., & Woolley, A. W. (2019). Team creativity, cognition, and cognitive style diversity. *Management Science*, *64*(4), 1586–1599. https://doi.org/10.1287/mnsc.2017.3001

Aggarwal, I., Woolley, A. W., Chabris, C. F., & Malone, T. W. (2019). The impact of cognitive style diversity on implicit learning in teams. *Frontiers in Psychology*, (*10*), 112. https://doi.org/doi: 0.3389/fpsyg.2019.00112

Alchian, A. A., & Demsetz, H. (1972). Production, information costs, and economic organization. *The American Economic Review*, *62*(5), 777–795.

Argote, L. (1999). *Organizational learning: Creating, retaining, and transferring knowledge*. Norwell, MA: Kluwer Academic.

Argote, L., & Epple, D. (1990). Learning curves in manufacturing. *Science*, *247*, 920–924.

Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization Science*, *22*(5), 1123–1137.

Argote, L., & Ren, Y. (2012). Transactive memory systems: A microfoundation of dynamic capabilities. *Journal of Management Studies*, *49*(8), 1375–1382.

Arrow, K. J., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica*, *22*, 265–290.

Ashkanasy, N. M., & Daus, C. S. (2005). Rumors of the death of emotional intelligence in organizational behavior are vastly exaggerated. *Journal of Organizational Behavior*, *26*(4), 441–452.

Atlee, T., & Por, G. (2000). Collective intelligence as a field of multi-disciplinary study and practice. www.academia.edu/1981875/Collective_Intelligence_as_a_Field_of_Multi-disciplinary_Study_and_Practice

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251.

Bastiaens, T. J., Baumol, U., & Kramer, B. J. (Eds.). (2010). On collective intelligence. In *Proceedings of the Symposium on Collective Intelligence, Hagen, Germany, April 15–16, 2010*. Berlin: Springer-Verlag.

Bates, T. C., & Gupta, S. (2016). Smart groups of smart people: Evidence for IQ as the origin of collective intelligence in the performance of human groups. *Intelligence*, 60, 46–56.

Benkler, Y., Shaw, A., & Hill, B. M. (2015). Peer production: A form of collective intelligence. In T. W. Malone & M. S. Bernstein (Eds.), *Handbook of collective intelligence*. Cambridge, MA: MIT Press.

Bernstein, E., Shore, J., & Lazer, D. (2018). How intermittent breaks in interaction improve collective intelligence. *Proceedings of the National Academy of Sciences*, *115*(35), 8734–8739.

Bigham, J. P., Bernstein, M. S., & Adar, E. (2015). Human-computer interaction and collective intelligence. In T. W. Malone & M. S. Bernstein (Eds.), *Handbook of collective intelligence*. Cambridge, MA: MIT Press.

Brennan, T. J., & Lo, A. W. (2012). An evolutionary model of bounded rationality and intelligence, *PLoS One*, *7*, e50310. doi:10.1371/journal.pone.0050310.

Camerer, C. F., Loewenstein, G., & Rabin, M. (Eds.) (2004). *Advances in Behavioral Economics*. Princeton: Princeton University Press.

Chabris, C. F. (2007). Cognitive and neurobiological mechanisms of the law of general intelligence. In M. J. Roberts (Ed.), *Integrating the mind: Domain general versus domain specific processes in higher cognition* (pp. 449–491). Hove: Psychology Press.

Chakrabarti, S. (2018). Hard questions: What effect does social media have on democracy? *Facebook Newsroom*, January 22. https://newsroom.fb.com/news/2018/01/effect-social-media-democracy/.

Condorcet, M. (1785/1976). Essay on the application of analysis to the probability of majority decisions. In K. M. Baker (Ed.), *Condorcet: Selected writings*, Indianapolis, IN: Bobbs-Merrill.

Credé, M., & Howardson, G. (2017). The structure of group task performance: A second look at "collective intelligence": Comment on Woolley et al. (2010), *Journal of Applied Psychology*, *102*(10), 1483–1492.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review*, *80*(2), 97–125.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. New York: Oxford University Press.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Devine, D. J., & Philips, J. L. (2001). Do smarter teams do better? A meta-analysis of cognitive ability and team performance. *Small Group Research*, *32*(5), 507–532.

DiMicco, J. M., Pandolfo, A., & Bender, W. (2004). Influencing group participation with a shared display. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 614–623). New York: Association for Computing Machinery.

Druskat, V. U., & Wolff, S. B. (2001). Building the emotional intelligence of groups. *Harvard Business Review*, *79*(3), 80–90.

Dutton, J. M., & Thomas, A. (1984). Treating progress functions as a managerial opportunity. *Academy of Management Review*, *9*, 235–247.

Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge: Cambridge University Press.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. New York: Dover Publications.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383.

Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., & Malone, T. W. (2014). Reading the mind in the eyes or reading between the lines? Theory of mind predicts collective intelligence equally well online and face-to-face. *PLoS One*, *9*(12), e115212.

Engelbart, D. C. (1962). *Augmenting human intellect: A conceptual framework*. No. Summary Report AFOSR-3223 under Contract AF 49 (638)-1024, SRI Project 3578 for Air Force Office of Scientific Research, Stanford Research Institute.

Engelbart, D. C. (1995). Toward augmenting the human intellect and boosting our collective IQ. *Communications of the ACM*, *38*(8), 30–32.

Engelbart, D. C., & English, W. K. (1968). A research center for augmenting human intellect. In *Proceedings of the December 9–11,1968, Fall Joint Computer Conference* (Vol. 33, pp. 395–410). San Francisco: Association for Computing Machinery.

Faraj, S., & Xiao, Y. (2006). Coordination in fast-response organizations. *Management Science*, *52*(8), 1155–1169.

Fiol, C. M., & Lyles, M. A. (1985). Organizational learning. *Academy of Management Review*, *10*(4), 803–813.

Franks, N. R. (1989). Army ants: A collective intelligence? *American Scientist*, *7*(2), 138–145.

Galbraith, J. R. (2002). *Designing organizations*. San Francisco: Jossey-Bass.

Galton, F. (1907). Vox Populi. *Nature*, *75*(7), 450–451.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic books.

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea generation and the quality of the best idea. *Management Science*, *56*(4), 591–605.

Gordon, D. M. (2015). Collective behavior in animals: An ecological perspective. In T. W. Malone & M. S. Bernstein (Eds.), *Handbook of collective intelligence*. Cambridge, MA: MIT Press.

Hackman, J. R. (1987). The design of work teams. In J. W. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315–342). Englewood Cliffs, NJ: Prentice Hall.

Hackman, J. R. (2011). *Collaborative intelligence: Using teams to solve hard problems*. San Francisco: Berrett-Koehler Publishers.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Decision Processes*, *16*(2), 250–279.

Hamilton, C. (2004). Come together: Can we discover a depth of wisdom far beyond what is available to individuals alone? What Is Enlightenment, May–July, 57–77.

Harrison, D. D., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199–1228.

Hart, O. (1995). *Firms, contracts, and financial structure*. London: Oxford University Press.

Heylighen, F. (1999). Collective Intelligence and its implementation on the Web: Algorithms to develop a collective mental map. *Computational and Mathematical Organization Theory*, *5*(3), 253–280.

Hill, G.W. (1982). Group versus individual performance: Are N + 1 heads better than one? *Psychological Bulletin*, *91*, 517–539.

Hiltz, S. R., & Turoff, M. (1978). *The network nation: Human communication via computer*. Reading, MA: Addison-Wesley.

Hollingshead, A. B. (2001). Cognitive interdependence and convergent expectations in transactive memory. *Journal of Personality and Social Psychology*, *81*(6), 1080–1089.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, *13*(2), 324–340.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, *101* (46), 16385–16389.

Horwitz, S. K., & Horwitz, I. B. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management*, *33*(6), 987–1015.

Howe, J. (2008). *Crowdsourcing: How the power of the crowd is driving the future of business*. New York: Three Rivers Press.

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, *50*(6), 1141–1151.

Isaacs, W. (1999). *Dialogue and the art of thinking together*. New York: Doubleday Currency.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York, NY: Free Press.

Jenkins, G. D., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance? A meta-analytic review of empirical research. *Journal of Applied Psychology*, *83*(5), 777–787.

Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. *Academy of Management Journal*, *52*(3), 599–627.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision and risk. *Econometrica*, *47*(2), 263–291.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review*, *103*(4), 687–719.

Kim, Y. J., Engel, D., Woolley, A., Lin, J., McArthur, N., & Malone, T. W. (2017). What makes a strong team? Using collective intelligence to predict performance of teams in League of Legends. In *Proceedings of the 20th ACM conference on computer-supported cooperative work and social computing (CSCW 2017)* (pp. 2316–2329). New York: Association for Computing Machinery.

Kim, Y. J., Gupta, P., Glikson, E., Woolley, A.W., & Malone, T. W. (2018). Enhancing collective of human-machine teams. Paper presented at the ACM Conference on Collective Intelligence, Zurich, Switzerland, July 7–8.

Knott, A. M. (2008). R&D/returns causality: Absorptive capacity or organizational IQ. *Management Science*, *54*(12), 2054–2067.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, *3*(3), 383–397.

Korde, R., & Paulus, P. B. (2017). Alternating individual and group idea generation: Finding the elusive synergy. *Journal of Experimental Social Psychology*, *70*, 177–190.

Kowalczyk, R. (Ed.). (2009). Computational collective intelligence. Paper presented at the Proceedings of the First International Conference on Computational Collective Intelligence, Wroclaw, Poland, October 5–7.

Krych-Appelbaum, M., Law, J. B., Barnacz, A., Johnson, A., & Keenan, J. P. (2007). The role of theory of mind in communication. *Journal of Interaction Studies*, *8*(2), 267–280.

Larson, J. R. (2009). *In search of synergy in small group performance*. New York: Psychology Press.

Latane, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, *37*(6), 822–832.

Lawrence, P. R., & Lorsch, J. W. (1967). *Organization and environment: Managing differentiation and integration*. Cambridge, MA: Harvard Business School Press.

Lazear, E. P. (2000). Performance pay and productivity. *The American Economic Review*, *90* (5), 1346–1361.

Levy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace*. New York: Plenum Press.

Lewis, K. (2004). Knowledge and performance in knowledge-worker teams: A longitudinal study of transactive memory systems. *Management Science*, *50*, 1519–1533.

Lewis, K., & Herndon, B. (2011). The relevance of transactive memory systems for complex, dynamic group tasks. *Organization Science*, *22*(5), 1254–1265.

Lo, A. (2015). The wisdom of crowds vs. the madness of mobs. In T. W. Malone & M. S. Bernstein (Eds.), *Handbook of collective intelligence*. Cambridge, MA: MIT Press.

Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, *108*(22), 9020–9025. https://doi.org/10.1073/pnas.1008636108.

Mackay, C. (1841). *Memoirs of extraordinary popular delusions and the madness of crowds* (1st ed.). London: Richard Bentley.

Malone, T. W. (2004). *The future of work: How the new order of business will shape your organization, your management style and your life*. Boston: Harvard Business School Press.

Malone, T. W. (2018). *Superminds: The surprising power of people and computers thinking together*. Boston: Little, Brown and Company.

Malone, T. W., & Bernstein, M. S. (2015). Introduction. In T. W. Malone & M. S. Bernstein (Eds.), *Handbook of collective intelligence*. Cambridge, MA: MIT Press.

Malone, T. W., Crowston, K., Lee, J., Pentland, B., Dellarocas, C., Wyner, G., et al. (1999). Tools for inventing organizations: Toward a handbook of organizational processes. *Management Science*, *45*(3), 425–443.

Malone, T. W., Laubacher, R., & Dellarocas, C. (2010). The collective intelligence genome. *MIT Sloan Management Review*, *51*(3), 21–31.

Malone, T. W., & von Ahn, L. (2012). *Proceedings of the Collective Intelligence 2012 Conference*. Cambridge, MA. http://arxiv.org/abs/1204.2991v3

Malone, T. W., Yates, J., & Benjamin, R. I. (1987). Electronic markets and electronic hierarchies. *Communications of the ACM*, *30*(6), 484–497.

Mann, R. D. (1959). A review of the relationships between personality and performance in small groups. *Psychological Bulletin*, *56*(4), 241–270. https://doi.org/10.1037/h0044587

Mataric, M. J. (1993). Designing emergent behaviors: From local interactions to collective intelligence. In A. Meyer, H. Roitblat, & S. Wilson (Eds.), *Proceedings of the Second International Conference on Simulation of Adaptive Behavior* (pp. 432–441). Cambridge, MA: MIT Press.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Englewood Cliffs, NJ: Prentice-Hall.

Meslec, N., Aggarwal, I., & Curşeu, P. L. (2016). The insensitive ruins it all: Compositional and compilational influences of social sensitivity on collective intelligence in groups. *Frontiers in Psychology*, 7, 676. https://doi.org/10:3389/fpsyg.2016.00676

Mesmer-Magnus, J. R., & DeChurch, L. A. (2009). Information sharing and team performance: A meta-analysis. *Journal of Applied Psychology*, *94*(2), 535.

Milliken, F. J., & Martins, L. L. (1996). Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *Academy of Management Review*, *21*(2), 402–433.

Mulgan, G. (2017). *Big mind: How collective intelligence can change our world*. Princeton: Princeton University Press.

Nelson, R. R., & Winter, S. (1982). *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, *5*(1), 14–37.

Ocasio, W. (1997). Towards an attention-based view of the firm. *Strategic Management Journal*, *18*(S1), 187–206.

Ocasio, W. (2011). Attention to attention. *Organization Science*, 22(5), 1286–1296.

Okhuysen, G. A., & Bechky, B. A. (2009). Coordination in organizations: An integrative perspective. *The Academy of Management Annals*, *3*(1), 463–502.

Page, S. E. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton: Princeton University Press.

Pirolli, P., & Russell, D. M. (2011). Introduction to this special issue on sensemaking. *Human-Computer Interaction*, *26*(1–2), 1–8.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(04), 515–526.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature*, *37*(1), 7–63.

Ridley, M. (2010). *The rational optimist: How prosperity evolves*. New York: Harper Collins.

Roloff, K. S., Woolley, A. W., & Edmondson, A. C. (2011). The contribution of teams to organizational learning. In M. Easterby-Smith & M. Lyles (Eds.) *Handbook of Organizational learning and knowledge management* (2nd ed., pp. 249–271). London: Blackwell.

Sanders, G. S., & Baron, R. S. (1977). Is social comparison irrelevant for producing choice shifts? *Journal of Experimental Social Psychology*, *13*(4), 303–314.

Senge, P. M., & Sterman, J. D. (1992). Systems thinking and organizational learning: Acting locally and thinking globally in the organization of the future. In T. A. Kochan & M. Useem (Eds.) *Transforming organization* (pp. 353–371). New York: Oxford University Press.

Simon, H. A. (1947). *Administrative behavior: A study of decision-making processes in administrative organizations*. Chicago: Macmillan.

Smith, J. B. (1994). *Collective intelligence in computer-based collaboration*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *57*, 67–78.

Steiner, I. D. (1966). Models for inferring relationships between group size and potential group productivity. *Behavioral Science*, *11*(4), 273–283.

Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.

Steyvers, M., & Miller, B. (2015). Cognition and collective intelligence. In T. W. Malone & M. S. Bernstein (Eds.), *Handbook of collective intelligence*. Cambridge, MA: MIT Press.

Stroebe, W., Nijstad, B. A., & Rietzschel, E. F. (2010). Beyond productivity loss in brainstorming groups: The evolution of a question. *Advances in Experimental Social Psychology*, *43*, 157–203.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*. New York: Anchor Books.

Szuba, T. M. (2001). *Computational collective intelligence*. New York: John Wiley & Sons.

Tapscott, D., & Williams, A. D. (2006). *Wikinomics: How mass collaboration changes everything*. New York: Portfolio/Penguin.

Thompson, J. D. (1967). *The structure of complex organizations*. Harmondsworth: Penguin Books.

Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review*, *2*(4), 1–109.

Toma, C., & Butera, F. (2009). Hidden profiles and concealed information: Strategic information sharing and use in group decision making. *Personality and Social Psychology Bulletin*, *35*(6), 793–806.

Ulrich, K. T., & Eppinger, S. D. (2015). *Product design and development*. New York: McGraw-Hill.

Van de Ven, A. H., Delbecq, A. L., & Koenig, R. (1976). Determinants of coordination modes within organizations. *American Sociological Review*, *41*(2), 322–338.

van Knippenberg, D., & Mell, J. N. (2016). Past, present, and potential future of team diversity research: From compositional diversity to emergent diversity. *Organizational Behavior and Human Decision Processes*, *136*, 135–145.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151.

Wageman, R., & Baker, G. (1997). Incentives and cooperation: The joint effects of task and reward interdependence on group performance. *Journal of Organizational Behavior*, *18*(2), 139–158.

Watson, A. C., Nixon, C. L., Wilson, A., & Capage, L. (1999). Social interaction skills and theory of mind in young children. *Developmental Psychology*, *35*(2), 386.

Wechsler, D. (1971). Concept of collective intelligence. *American Psychologist*, *26*(10), 904–907.

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York: Springer-Verlag.
Weick, K. E. (1993). The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly*, *38*, 628–652.
Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage Publications.
Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, *16*(4), 409–421.
Weld, D. S., Mausam, Lin, C. H., & Bragg, J. (2015). Artificial intelligence and collective intelligence. In T. W. Malone & M. S. Bernstein (Eds.), *Handbook of collective intelligence* (pp. 89–118). Cambridge, MA: MIT Press.
Williams, W. M., & Sternberg, R. J. (1988). Group intelligence: Why some groups are better than others. *Intelligence*, *12*(4), 351–377.
Williamson, O. E. (1973). Markets and hierarchies. *American Economic Review*, 63, 316–325.
Williamson, O. E. (1981). The economics of organization: The transaction cost approach. *American Journal of Sociology*, *87*(3), 548–577.
Wilson, J. M., Goodman, P. S., & Cronin, M. A. (2007). Group learning. *Academy of Management Review*, *32*(4), 1041–1059.
Woolley, A. W. (2009a). Means versus ends: Implications of outcome and process focus for team adaptation and performance. *Organization Science*, *20*, 500–515.
Woolley, A. W. (2009b). Putting first things first: Outcome and process focus in knowledge work teams. *Journal of Organizational Behavior*, *30*, 427–452.
Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330* (6004), 686–688.
Woolley, A. W., Gerbasi, M. E., Chabris, C. F., Kosslyn, S. M., & Hackman, J. R. (2008). Bringing in the experts: How team composition and work strategy jointly shape analytic effectiveness. *Small Group Research*, *39*(3), 352–371.
Woolley, A. W., Hackman, J. R., Jerde, T. J., Chabris, C. F., Bennett, S. L., & Kosslyn, S. M. (2007). Using brain-based measures to compose teams: How individual capabilities and team collaboration strategies jointly shape performance. *Social Neuroscience*, *2*, 96–105.
Woolley, A. W., Kim, Y., & Malone, T. W. (2018). Measuring collective intelligence in groups: A reply to Credé and Howardson. Social Science Research Network Paper No. ID 3187373, Rochester, NY. https://papers.ssrn.com/abstract=3187373
Woolley, A. W., Riedl, C., Kim, Y. J., & Malone, T. W. (2017). More evidence for a general collective intelligence factor in human groups: A meta-analysis. Paper presented at the 2017 Collective Intelligence Conference, New York, June 16, 2017.

33 Leadership Intelligence

Richard E. Boyatzis

Leaders matter to organizational performance and adaptability. Effective leaders matter the most in a dramatic and positive manner. Possibly owing to the power and magnetism or fear and suspicion about leaders, scholars have sought to identify the characteristics that distinguishes distinctive leaders. This search has gone on for millennia, from the ancient Greek and Chinese philosophers to legions of management and leadership researchers today. The search for a missing factor that could account for a significant amount of the variance in any outcome variable about leadership effectiveness has left scholars arguing, methodologists scampering for more and more precise measures, and practitioners vulnerable to fads. Although these projects have provided ample employment for many professors, internal and external consultants, and gurus, the research evidence suggests a more prosaic conclusion.

The "trait theory" of leadership, or what was called the great man theory from a misogynist perspective, was the quest for the illusive characteristic that defined and predicted who among those who became leaders were effective. Although never consistently confirmed in rigorous research, cognitive intelligence was often thought to be the determining variable. After about 2,000 years of searching, this approach gave way to approaches that claimed leadership was determined not by a characteristic of the person but by an interaction of a person's capability and style with characteristics of the situation. In research, this meant looking for interaction effects of individual and situational variables. Research in the last twenty-five years has focused on the study of followers (Riggio, Chaleff, & Lipman-Blumen, 2008; Tee, Ashkanasy, & Paulsen, 2013) and the relationship of leaders with those around them (Graen & Uhl-Bien, 1995).

Regarding capabilities of a person, all of the proposed and hypothesized forms of intelligence in this handbook seem relevant to leadership effectiveness in varying proportions. But most likely they need to be considered in some balance and in their context of each other and of the quality of the relationships among the leader and the people around them. If applied to research, this contribution suggests including variables and measures to ensure interaction effects are examined and various forms of intelligence are accounted for (covaried). In application, this contribution suggests a multilevel, more holistic approach to developing leaders.

Leadership Intelligence or Intelligent Leadership

The clearest answer to the question of whether leadership constitutes an ability and therefore could be claimed to be a form of intelligence would be, "no." Although neural processes are involved in the fulfillment of a leadership role, no serious researcher to date has provided evidence or even a theoretical framework within which to claim that there is a neural region or network that is the "leadership intelligence," or LI network.

The American Psychological Association's Task Force on Intelligence (APA Public Affairs Office, 1997) reported that predicting real-life outcomes is an important part of the standard against which we should judge an intelligence. It went on to add that there should be a consensus within a field as to the definition. Since no one is claiming that LI is a specific form of intelligence, we can discard that argument. But the exploration of the variety of neural networks invoked when a person engages in leadership activities and role enactment does suggest that there are multiple forms of intelligence that are relevant and used during the exercise of leadership and possibly effective leadership.

In an effort to establish emotional intelligence (EI) as a form of intelligence, Mayer, Salovey, and Caruso (2000) claimed that three criteria define an intelligence: (1) it should reflect a "mental performance rather than preferred ways of behaving" (pp. 269–270); (2) tests should show positive correlation with other forms of intelligence; and (3) the measures should increase with experience and age. Boyatzis and Sala (2004) claimed that, to be classified as an intelligence, the concept should be:

> (1) Behaviorally observable; (2) Related to biological and in particular neural-endocrine functioning. That is, each cluster should be differentiated as to the type of neural circuitry and endocrine system involved; (3) Related to life and job outcomes; (4) Sufficiently different from other personality constructs that the concept adds value to understanding the human personality and behavior; and (5) The measures of the concept, as a psychological construct, should satisfy the basic criteria for a sound measure, that is, show convergent and discriminant validity. (Campbell & Fiske, 1959)

This chapter is really about the role of intelligence in leadership, not the claim that the capability to be an effective leader is a distinct individual characteristic or a type of intelligence. Intelligent leadership, therefore, is leadership in which a person uses many forms of intelligence – cognitive, emotional, and social. Yet that might not help with a comprehensive theory or practical approach to leadership development. Not to claim we can complete the picture but we can add some important dimensions; we must examine some other components of the person and even personality if broadly defined.

Cognitive Intelligence

The argument about cognitive intelligence and leadership is between academics who wish to promote the idea that being classically smart is both necessary and sufficient for leadership effectiveness and those that believe it is necessary but far from sufficient. Research confirms that various measures of general mental ability are highly related (Frey & Detterman, 2004). Various scholars have shown that the SAT (formerly called the Scholastic Aptitude Test and then the Scholastic Assessment Test) is significantly predictive of general cognitive ability (*g*) and both are highly correlated with grade point average (GPA). As Frey and Detterman (2004) explain, the SAT was originally seen as an IQ test but then diverged in the 1940s as a test of reasoning.

Another of the standardized tests is the Graduate Management Admissions Test (GMAT), which is for management school admissions, like the Medical College Admissions Test (MCAT) for medical school, the Law School Admissions Test (LSAT) for law school, and the Graduate Record Examination (GRE) for a variety of graduate schools. The GMAT was shown to predict grades in the first year of an MBA program only (Crooks, Campbell, & Rick, 1979), with no relation to actual managerial performance. Years and many studies later, the GMAT was found to be even more valid as a predictor of first-year GPA on an MBA program and also the entire GPA for the MBA program. It may provide evidence of perseverance (Oh et al., 2008). O'Reilly and Chatman (1994) showed that GMAT scores and motivation predicted early career success among MBAs within three to four years after graduation but that the separate measures of GMAT and motivation did not support the conclusion that cognitive ability is necessary but not sufficient for managerial or even leadership success or effectiveness.

Cognitive intelligence emerged over the years as having several crucial components, such as fluid and crystalized intelligence. Fluid intelligence is the ability to think logically and solve problems, especially in new situations (Cattell, 1963). Meanwhile, crystalized intelligence is a person's knowledge base about the world and learned operations such as using specific mathematical formulae (Cattell, 1967). Working memory is often seen as the third leg of the cognitive ability stool. It is seen as a system for processing that simultaneously stores and manipulates information, even when distractions or alternate competing ideas occur (Nisbett et al., 2012). As Nisbett and colleagues (2012) explained, working memory also incorporates verbal and spatial problem-solving, arithmetical reasoning, and abstract reasoning. They claimed that working memory is more closely assessed by tests like the SAT than are other elements of cognitive ability. They also claimed that fluid intelligence is closer to what most contend is general cognitive ability, or *g*, than other components.

Successful and Practical Intelligence as a Broader Concept

An alternative approach to cognitive intelligence were part of a "triarchic theory" by purporting that, internally, a person has "meta-components, and knowledge-acquisition components" (Sternberg, 1985, p. 59). This encompasses the analytic processes involved in thinking about life. Sternberg (1985, 2011) went on to propose that the application of these intelligences to everyday life constituted a "practical intelligence" (Sternberg & Hedlund, 2002), which he later refined to be "successful intelligence" (Sternberg, 1999).

Instead of focusing on analytic processes, speed, and working memory, Sternberg's (2011) theory postulated that such analytic processes were only one aspect of a person's ability to "adapt to the environment and learn from experience" (Sternberg & Detterman, 1986). Successful intelligence was the overarching concept that included the quest for goals in life and work. This incorporated crystalized and fluid intelligence components of the traditional cognitive intelligence theories and added elements from what was later to be called emotional intelligence (see the next section). In particular, practical intelligence was the formulation and use of tacit knowledge gained from one's experiences. This was directly related to the performance of leaders in management simulations (Sternberg, 2011) and leadership effectiveness while controlling for *g* (Hedlund et al., 2002; Sternberg & Hedlund, 2002).

One of the many contributions of this approach was bringing intelligence into the behavioral realm. People could now talk about, theorize, and study how individuals applied their internal capability and how it looked to observers. This raised the question as to whether there are forms of intelligence and neurologically based processing that might be more closely related to life and work outcomes and leadership than traditionally defined cognitive intelligence.

The challenge to the role of cognitive intelligence was further questioned in comprehensive studies. Grossman and colleagues (2013) showed that wise reasoning, which they defined as pragmatic analysis in social settings, especially within emotional and conflict events, predicted well-being, career and life satisfaction, and longevity. Their results showed that various measures of cognitive ability, like the WAIS comprehension assessment or processing-speed scale, were negatively related to wise reasoning and well-being. They went further to claim, as shown in prior research, that cognitive abilities such as crystallized intelligence, processing speed, and working memory showed no systematic, positive relationship to well-being (Grossman et al., 2013). These claims suggest wise reasoning is closely related to Sternberg's (1985) concept of practical intelligence.

Emotional and Social Intelligence

In addition to the expansion of intelligence from Sternberg (1985, 2011), among the seven forms of intelligence conceptualized by Howard Gardner (1983) were intrapersonal and interpersonal intelligence. Today, these closely correspond to what are called emotional and social intelligence (ESI), respectively. Although often classified as variations of EI, they have distinctly different neural networks as origins and different behavioral outcomes. The call for a behavioral approach first came with McClelland's (1973) key article in the *American Psychologist*, which sought to understand competencies and not just traditional intelligence. First labeled as an intelligence by Peter Salovey and Jack Mayer (1990), a flurry of research, opinion essays, theory articles, and practitioner pieces emerged. Daniel Goleman's (1995) book *Emotional Intelligence* brought the ideas into the mainstream of practice within organizations and education. Critiques followed (Matthews et al., 2006) and the research became increasingly rigorous.

By 2005, the research, definitions, and measures of EI were said to fall into three streams (Ashkanasy & Daus, 2005). Stream 1 was a direct measurement of how a person handled emotional information, as exemplified by the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) model and measure (Salovey & Mayer, 1990). Stream 2 measures used the MSCEIT model but were based on self-assessment (Schutte et al., 1998; Law, Wong, & Song, 2004). Stream 3 comprised all other measures (Bar-On, 1997; Petrides & Furnham, 2000, 2001). By 2014, Amdurer and colleagues (2014) put forward a four stream, which comprised models with behavioral measures as examined in detail by Boyatzis (2009, 2017) and differed from Stream 3 in that the source of the information was from coded audiotapes of work samples, videotapes of simulations, or others' observations from 360° assessments. Behavioral ESI is likely measuring many of the components of practical intelligence but is clustered to reflect neural networks more closely.

A series of meta-analyses confirmed that all of these measures and approaches were significantly related to job performance. They included leadership effectiveness (Joseph et al., 2014; O'Boyle et al., 2011); leadership job satisfaction and satisfaction of their subordinates (Miao, Humphrey, & Qian, 2016); authentic leadership (Miao, Humphrey, & Qian, 2018a); and subordinates' task performance and organizational citizenship (Miao, Humphrey, & Qian, 2018b). Without the benefit of Stream 4 as a separate category, these meta-analyses all found that EI assessed in any stream was significantly predictive of the outcomes but that Stream 3 measures were more strongly related to the outcomes. Joseph and colleagues (2014) criticized Stream 3 measures' impact on outcomes in their own and others' meta-analyses because of the contamination of self-assessment measures with personality. Further work documenting Stream 4 measures has shown unique variance in the outcomes from the behavioral EI measure separate from the effect of the Big Five personality traits.

Several research studies have shown that behavioral measures of ESI demonstrate unique variance in several effectiveness measures (Boyatzis, Good, & Massa, 2012; Boyatzis, Rochford, & Cavanaugh, 2017). In these studies, measures of cognitive intelligence and personality in terms of the Big Five traits did not show significance in predicting effectiveness. This helps to confirm that ESI is distinct from cognitive or traditional intelligence and should be included in any model of intelligence explaining leadership. Another study had similarly shown unique and significant variance explained by EI (Cote & Miners, 2006) in academic performance, in contrast to cognitive intelligence. Using a Bayesian analysis, Boyatzis, Batista, and colleagues (2015) showed how behavioral ESI competencies showed no relationship to cognitive ability as measured by the GMAT in a large sample of MBAs.

Successful intelligence (Sternberg, 2011) would appear to be closely related to specific EI competencies, such as achievement orientation, emotional self-control, and adaptability. In the subset of successful intelligence framed as practical intelligence, Sternberg (2011) labeled a further subset as social intelligence (SI). This latter category has considerable overlap with specific SI competencies such as influence and the full range of competencies enabling people to build relationships with others.

Certainly the strong indications are that leadership effectiveness can be better understood by incorporating cognitive and ESI into the model of how intelligence relates to leadership.

Opposing Neural Domains and Cognitive vs. Social/Emotional Networks

To understand the differences between neural networks associated with traditional cognitive intelligence, practical intelligence, and ESI, we must turn to an explanation of the opposing domains of specific neural networks (Jack et al., 2012). The opposing domains are two important neural networks that affect our daily lives and functioning as leaders: the Task Positive Network (TPN) and the Default Mode Network (DMN).

The TPN is activated and enables a person to solve problems; analyze people or situations, especially data or financial information; focus on a task; focus on details and be somewhat resistant to change; and engage in abstract thinking. For leaders, the TPN is engaged in any analytic problem-solving, such as analyzing financial performance. It is a form of convergent problem-solving (Friedman et al., 2015). When activated, the TPN is seen as creating a psychological distance between the leader and others. The TPN consists primarily of the dorsal attention system (Fox et al., 2005); the frontoparietal control network (Vincent et al., 2008); the ventral attention network; and the dorsal anterior cingulate cortex (Jack et al., 2012; Martin & Weisberg, 2003; Prabhakaran et al., 1997).

The DMN activates and enables a person to scan the environment and be open to perceiving new ideas, notice and be open to people and their emotions, and be considerate of moral concerns (i.e., not the good or bad distinction that is more analytic or TPN but the sense of something as either fair and just or unfair and unjust). For leaders, the DMN enables brainstorming and more holistic, global thinking (Friedman et al., 2015). It consists primarily of the medial parietal and dorsal medial prefrontal cortex (dMPFC), the ventral medial prefrontal cortex, the right temporoparietal junction, the posterior cingulate cortex, the orbitofrontal cortex, and the nucleus accumbens (Buckner, Andrews-Hanna, & Schacter, 2008; Decety & Batson, 2007; French & Jack, 2014; Jack et al., 2012; Jack, Boyatzis et al., 2013; Van Overwalle, 2010).

Traditional cognitive intelligence is a product of a person activating and applying the TPN (Jack et al., 2012; Jack, Dawson, & Norr, 2013; Prabhakaran et al., 1997). Social, emotional, and practical intelligence is a product of a person predominantly activating and applying the DMN (Buckner, Andrews-Hanna, & Schacter, 2008; Jack et al., 2012).

The distinctive importance of these two networks is that they have almost no overlap and suppress each other (Fox et al., 2005; Jack et al., 2012). This antagonistic relationship helps to explain why leadership styles and behavior have been divided into task and social dimensions for so long (Boyatzis, Rochford, & Jack, 2014). It also explains why so many developmental efforts at management and leadership education and training fail to produce improvements in the desired behavior that last more than a few weeks or months (Boyatzis et al., 2014) because of the curricular focus on abstract and analytic processes. Subjects taught in MBA programs such as finance, accounting, economics, statistics, supply chain, and operations research are usually taught through formulae, problem sets, and predominantly analytic thought processes. It also helps us understand why the focus on goals, measurement, and dashboards in organizations repeatedly activates the TPN, to the exclusion of the DMN, which may result in expedient but unethical decisions and actions (Rochford et al., 2016), not noticing or valuing people in the organization, and not noticing competitors' actions or changes in market demand.

While it has been conjectured that people may have individual dispositions toward using one or the other of these networks (Epstein et al., 1996), Jack and colleagues (2012) note that the toggle rate between the two networks may be as brief as thousandths of a second. Some leaders may be thought to be using both networks at the same time but they are adept at switching below the conscious threshold that others might notice or they might be self-aware. It is also hypothesized that effective leaders are more attuned to activating a network as appropriate to a situation, for example the TPN in addressing a financial variance issue or the DMN in understanding a shift in customers' preferences. Effective leaders need both networks but they also need to be adept in moving back and forth between them.

While this appears at first glance to be similar to Kahneman's dual-process theory (Kahneman, 1992, 2011), it actually refines it considerably. In Kahneman's (1992, 2011) theory, System 1 thinking is automatic, fast, and seemingly effortless. Meanwhile, System 2 thinking is slow, deliberative, reflective, and controlled. According to Friedman and colleagues (2015), opposing domains explain how both the fast and the slow circuits of the dual-process theory have analytic and social dimensions. That is, the opposing domains can map onto the dual processes and the result is four clusters of consequences. But the opposing domains explain the neural functioning more precisely. For example, for leaders, controlled or slow processes in the social and empathic network (i.e., DMN) would manifest as autobiographical recall, emotional regulation, and the telling of social narratives. While the controlled or slow processes in analytic networks (i.e., TPN) would manifest as any goal-driven logic, analytic and especially empirical analysis of budget variances, and often as a more competitive attitude (Friedman et al., 2015). For the dual-process fast and automatic processes, the analytic network, or TPN, would be activated with practiced actions (i.e., habitual) and rapid calculations. In contrast, fast processes with the empathic network, or DMN, would manifest as emotion-driven statements and involve aspects of social stereotyping as well as brainstorming (Friedman et al., 2015).

EI and SI appear to be primarily in the arena of the DMN (Buckner et al., 2008; Boyatzis, Rockford, & Jack, 2014; Fox et al., 2005; Jack et al., 2012, 2013; Martin, & Weisberg, 2003;). Except for those aspects of EI and practical intelligence involving emotional self-control and adaptability, which are more likely within the domain of the TPN, all other aspects would require activation of the DMN. To understand the conflicts among the different forms of intelligence, more research is needed to examine the battle in the brain for mindshare.

Orientation and Dispositions to Leadership

Leadership Motive Profile and the Role of Need for Power

Leadership requires influencing others (McClelland, 1975, 1985; Yukl, 1998; Yukl & Van Fleet, 1990). It is about having an impact on others and making things happen. Of the many attempts to understand or explain a person's motivation for being in a leadership position, perhaps the most illuminating was that of McClelland and his colleagues. They studied the underlying disposition to want to influence and have an impact on others. McClelland (1975, 1985) called this the Need for Power, defined as an unconscious drive or motive in which the person wanted to have an impact on others. It was assessed through a conscious/unconsciousness projective test, the Thematic Apperception Test (TAT). McClelland showed repeatedly that self-assessment measures assessed valuing power but not the motive. Valuing power is how people answer self-report measures or surveys about their own power needs. McClelland (1985) showed with repeated studies that such self-assessment statements reflect a person's values but not their actual behavior. Because of the assessment method, a person's self-assessment of their power needs and desires is a measurement of their values or attitudes about power. Meanwhile, a projective test like the TAT has been shown to get beneath the self-attribution level and reflect deeper, predominantly unconscious drives. As a result, the self-assessment of the motive did not predict expected outcomes and behavior but rather attitudes (McClelland, 1985). Assessment based on the TAT measures provided a more accurate measure of the drive (McClelland, 1985) and the expected behavior and outcomes than self-report surveys of a person's power needs.

The Need for Power was shown to predict a variety of life and job outcomes (McClelland, 1985). Yet, when the need for power was accompanied by a relatively low need for affiliation (being friendly and caring toward others) and a relatively high unconscious desire for self-control, the combination was called the Leadership Motive Profile (Fontana et al., 1987; Jenkins, 1994; Jacobs & McClelland, 1994; McClelland, 1975; McClelland & Boyatzis, 1982). This pattern of combined motives showed the most consistent positive relationship to leadership effectiveness. A related form of a high Need for Power with relatively low self-control was associated with more "personalized power" and what Winter (1973) called the Don Juan Syndrome. People with this pattern of power drive in leadership positions tended to be self-serving, narcissistic, and more concerned about their own reputation and impact than the greater good of the organization and its many stakeholders.

In later work, McClelland applied the concept of the Need for Power to Stewart's developmental stage model of personality (see McClelland, 1985) and classified a high power drive with self-control as "socialized power" and a high power drive with relatively low self-control as "personalized power." These were two forms of Stewart's Stage 3 in personality development. The unconscious drive for self-control of a person's impulses and urges emerged in studies as a form of sacrifice of the person for the good of the organization or group (i.e., family, relationships).

The *leadership motive pattern* was shown to predict increased health problems because of the power stress invoked in leadership roles. The increased demand for the exercise of influence and power was repeatedly shown to activate the sympathetic nervous system (i.e., the human stress response), which compromises a person's immune system and leaves them vulnerable to disease agents and processes. In an interesting anthropological insight, McClelland (1975) showed that countries with this leadership motive pattern in their popular literature and myths had higher rates of cardiovascular disease per capita than those with lower such drives.

Meaning and Purpose

Ancient philosophers claimed that a sense of purpose, or "telos," helped to drive one's behavior (Ross, 1925). In an early stage of American pragmatism, Benjamin Franklin said that people had it within their own power to become more virtuous through intentional actions (Franklin, 2012). In psychology, William James (1890) claimed that a person can exert conscious volition or will in framing and determining their actions in life. The articulation of one's intention or will can be seen as a vision or a dream of a desired future. It has been recognized as a driving force behind sustained, desired change (Boyatzis & Akrivou, 2006; Higgins, 1987). The contrast of the ideal self to the real self or the ought self suggested that a sense of purpose provides meaning for a person (Boyatzis & Akrivou, 2006; Higgins, 1987). When called on in teams or organizations, a shared vision may excite, engage, and inspire others (Boyatzis, Rochford, & Taylor, 2015).

Being able to articulate and remind those involved of their collective sense of purpose, if not their noble purpose, appears to be a well-recognized feature of effective leadership (Bennis & Nanus, 1985; Berg, 2015; Greenleaf, 1970/2015). Research linking vision, purpose, or calling to leadership has taken many forms (i.e., calling, legacy, noble purpose) and only appeared in the last fifteen years in journals. A set of articles showed the potency of shared vision in a special issue of *Frontiers in Psychology* in 2015. Shared vision affected leadership in predicting succession of daughters in family businesses (Overbeke, Bilimoria, & Somers, 2015); financial health over time of family businesses (Neff, 2015); effectiveness of next-generation leaders of family businesses (Miller, 2014); effectiveness of physician leaders (Quinn, 2015); effectiveness of IT managers (Pittenger, 2015); increased corporate social responsibility (Thornton, 2015); and success of mergers and acquisitions when experienced by leaders (Clayton, 2014). In other studies, perceived shared vision in knowledge-worker teams in consulting and manufacturing (Mahon, Taylor, & Boyatzis, 2014), among engineers in project teams (Boyatzis et al., 2017), for community college president's effectiveness (Babu, 2016), and in stimulating innovation in high-tech firms (Kendall, 2016) has moderated and amplified the impact of ESI on engagement.

While Bennis and Nanus (1985) described how an effective leader would focus the attention of others through vision. Kantabutra and Avery (2010) explained more deeply why a sense of vision and purpose helped people organize their collective actions. One study showed that it is possible to help a person build a more comprehensive and compelling personal vision (Mosteo et al., 2016). Another study showed how even thirty minutes of coaching about a person's personal vision activated many of the components of the person's DMN but, specifically, the lateral visual cortex, which allows a person to dream and imagine (Jack et al., 2013).

Whether it is the elicitation of hope from repeatedly being reminded of an organization's shared vision or the focus created by talking about the shared sense of purpose, leaders appear to activate engagement, citizenship, and innovation by others when they remind the people around them of this essence of why they are together. Movements have emerged to foster shared vision among CEOs. Conscious Capitalism is a worldwide association of CEOs committed to develop others, our communities, and a noble purpose through work organizations. They want business leaders to expand their mission to include moral and responsible dimensions within their organizations (see www.consciouscapitalism.org).

Values and Style

The search for the impact of values on effective leadership has continued. The results have been inconsistent. It appears that separate values do not consistently predict the behavior of leaders but that value orientation, which has been called operating philosophy, does (Boyatzis, Murphy, & Wheeler, 2000).

The study led by Bernard Bass (2008) on the characterization of leadership styles as transformational versus transactional produced a major stream of research. The leader using the transformational style emphasizes the big picture, vision and purpose, and the desire to innovate. The leader using the transactional style emphasizes exchanges and quid pro quo approaches to motivation and engagement. Prior to that, in reaction to world events, in the 1940s, leadership styles of democratic, authoritarian, and laissez-faire organizations were studied (Lewin, Lippitt, & White, 1939). The democratic style involved others in decision-making and encouraged a view that the leader was also one of the people in the organization. The authoritarian style involved a concentration and exercise of power coming from the leader onto others who were less potent in any situation. The laissez-faire style appeared as less involved and being more permissive – going along with whatever was occurring. In the 1960s, Theory X and Theory Y were contrasted as a mechanical approach focused on instrumentality and efficiency versus a human approach (MacGregor, 1960).

The transformational leadership style was shown to predict effective leadership in many settings (Avolio, Bass, & Jung, 1999). The transactional style was less effective. Later studies showed that both were useful but in different settings. The transactional style was effective when the work was routine (Bass, 2008). In one study, behavioral ESI (i.e., as seen by others) was strongly related to the use of a transformational style of leadership (Bajaj & Medury, 2013).

In taking a more relational approach, resonant versus dissonant leadership styles were examined in terms of neural activation. The resonant leadership style was the experience of the leader and the people around them as being in sync or in tune with each other. It was often characterized as involving the experience of hope through vision or purpose, caring through compassion and authenticity through mindfulness. The dissonant leadership style was the experience of distance and separation of the leader and the people around them. It was often experienced with leaders who micromanaged others, were negative, controlling, even at times nasty and demeaning to others. Boyatzis and colleagues (2012) examined neural activation in executives in reaction to listening to brief statements about moments each person had with specific resonant and dissonant leaders in their past. Memories of resonant leaders activated many elements of the motor neuron network and the social network within the DMN. Meanwhile, memories of specific moments with dissonant leaders suppressed motor neuron networks and two-thirds of the time suppressed elements of the DMN activated with recollections of moments with resonant leaders. This suggested that both leadership style and the nature of the relationship between the leader and their people around them were important in helping to stimulate more openness.

Eighty years of research depicting leadership styles as different orientations to tasks versus social- or people-oriented was clarified as primarily emerging from two dominant neural networks discussed briefly earlier in this chapter, the TPN and DMN (Boyatzis et al., 2014). These neural networks helped to explain why these two styles were so often seen by scholars and practitioners as alternatives, with the best leadership being the use of both – but that was an elusive aspiration.

The Dark Side of Leadership

Unlike the other forms of intelligence and possible moderators and mediators of effective leadership, there are some characteristics, often called traits or even styles, that have been associated with the dark side of leadership, namely authoritarianism (Adorno et al., 1950), which has also been called Machiavellianism (Christie & Geis, 1970), and narcissism (i.e., ego-centrism or self-centeredness). Along with psychopathy, they have been called the dark triad (Paulhus & Williams, 2002). The choice in framing this chapter was to focus on the characteristics that positively affected leadership. Yet we would be remiss in not declaring that there have been traits and styles that have consistently been associated with less effective leadership or, at best, unevenly associated with leadership effectiveness.

Narcissism does appear related to self-perceived leadership effectiveness but has no relationship to other-perceived effectiveness (Grijalva et al., 2015). It also shows no consistent relationship in either direction to leadership effectiveness with other measures but does suggest a curvilinear relationship in which some of it might be useful but not too much (Grijalva et al., 2015). Similar relationships to leadership effectiveness have been noted with an authoritarian or Machiavellian personality in leaders and even with psychopathology (LeBreton, Schiverdecker, & Grimaldi, 2018).

Quality of Relationships as the Context of Leadership

As statistical methods have become more sophisticated and theories more detailed, the study of leadership has become increasingly complex and subtle. While an observer would likely admit that anyone's individual capabilities, including various forms of intelligence, may appear in any situation in life, the observed behavior is likely to be modified by the situation. Although role clarity, structure, and culture may affect a leader's behavior, it is the quality of one's relationships that might have the most immediate and direct impact on the transformation of their individual abilities. This leads to a need to contemplate and invoke moderators and mediators in leadership research (Fischer, Dietz, & Antonakis, 2017; Miao, Humphrey, & Qian, 2017).

The concept of leadership relationships on which the most academic articles have been published was created by George Graen and called the Leader-Member Exchange (LMX) (Graen & Uhl-Bien, 1995). From this literature, we know that a person's relationship to their immediate supervisor often mediates or moderates the impact of individual characteristics on leadership effectiveness and engagement.

To examine the role of relationships, one, more recent, approach has looked at the degree of shared vision, shared compassion, shared positive mood, and later energy in the relationships. Among knowledge workers in consulting and manufacturing research and development teams, shared vision within the teams moderated the positive impact of average EI as observed by teammates on their engagement (Mahon, Taylor, & Boyatzis, 2014). In family businesses, shared vision between a daughter and her father moderated the relationship of the daughter's self-efficacy on likelihood of succession to the CEO position (Overbeke, Bilimoria, & Somers, 2015). Miller (2014) found that shared vision, among other family climate variables, moderated the impact of behavioral ESI (i.e., as seen by others) on next-generation leader effectiveness. Among physician executives in hospitals, Quinn (2015) showed that shared vision mediated the impact of ESI on leader organizational citizenship, while Pittenger (2015) showed a comparable mediation of ESI on engagement of IT managers and professionals. The effect of philosophy and values on the corporate social responsibility of leaders was mediated by the degree of shared vision (Thornton, 2015). Meanwhile, Neff (2015) found it to be one of the five factors in the business climate among family businesses that predicted multiyear financial success. The perceived degree of shared vision among engineers in project teams contributed 27 percent of the unique variance in their own engagement, suggesting that relationship quality may not just be a moderator or mediator of effectiveness but may instead be another key variable.

Conclusions

While we contend that there is no specific part of the brain and focused ability that constitutes an LI, there are many networks within the brain and personal capabilities that contribute to a person being able to demonstrate intelligent leadership. This appears in the form of effective leadership with the capability of leveraging one's relationships and opportunities toward collective purpose and goals. When used ethically, this composite of a person's talent motivates the human spirit and leads.

References

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper and Row.

Amdurer, E., Boyatzis, R. E., Saatcioglu, A., Smith, M., & Taylor, S. N. (2014). Long-term impact of emotional, social and cognitive intelligence competencies on career and life satisfaction and career success. *Frontiers in Psychology*, *5*(1447), 1–15. https://doi.org/10.3389/fpsyg.2014.01447

APA (American Psychological Association) Public Affairs Office (1997). *Intelligence: Knowns and unknowns*. Washington, DC: APA.

Ashkanasy, N. M., & Daus, C. S. (2005). Rumors of the death of emotional intelligence in organizational behavior are vastly exaggerated. *Journal of Organizational Behavior*, *26*, 441–452.

Avolio, B. J., Bass, B. M., & Jung, D. I. (1999). Re-examining the components of transformational and transactional leadership using the multifactor leadership questionnaire. *Journal of Occupational and Organizational Psychology*, *72*, 441–462.

Babu, M. (2016). Characteristic of effective leadership of community college presidents. Unpublished doctoral dissertation, Case Western Reserve University.

Bajaj, B., & Medury, Y. (2013). Relationship between emotional and social competencies and transformational leadership styles. *International Journal of Research in Commerce and Management*, *3*(2), 56–59.

Bar-On, R. (1997), *Bar-On Emotional Quotient Inventory: Technical manual*. Toronto: Multi-Health Systems.

Bass, B.M. (2008). *Handbook of leadership: Theory, research and managerial applications* (4th ed.). New York: Free Press.

Bennis, W., & Nanus, B. (1985). *Leaders: The strategies for taking charge*. New York: Harper and Row.

Berg, J. L. (2015). The impact of symbiotic visions on performance and a sense of personal accomplishment: Increasing engagement and motivation through vision alignment. *Frontiers in Psychology*, 6, 443. https://doi.org/10.3389/fpsyg.2015.00443

Boyatzis, R. E. (2009). A behavioral approach to emotional Intelligence. *Journal of Management Development*, *28*, 749–770.

Boyatzis, R. E. (2017). Commentary of Ackley (2016). Updates on the ESCI as the behavioral level of emotional intelligence. *Consulting Psychology Journal*, *68*(4), 287–293.

Boyatzis, R. E., & Akrivou, K. (2006). The ideal self as a driver of change. *Journal of Management Development*, *25*(7), 624–642.

Boyatzis, R. E., Batista, J. M., Fernandez-Marin, X., & Truninger, M. (2015). EI competencies as a related but different characteristic than intelligence. *Frontiers in Psychology*, *6*(72), 1–14. https://doi.org/10.3389/fpsyg.2015.00072

Boyatzis, R. E., Good, D., & Massa, R. (2012). Emotional, social and cognitive intelligence as predictors of sales leadership performance. *Journal of Leadership and Organizational Studies*, *19*, 191–201.

Boyatzis, R. E., Murphy, A., & Wheeler, J. (2000). Philosophy as the missing link between values and behavior. *Psychological Reports*, *86*, 47–64.

Boyatzis, R. E., Passarelli, A. M., Koenig, K., Lowe, M., Mathew, B., Stoller, J. K., & Phillips, M. (2012). Examination of the neural substrates activated in memories of experiences with resonant and dissonant leaders. *The Leadership Quarterly*, *23*(2), 259–272.

Boyatzis, R. E., Rochford, K., & Cavanagh, K. (2017). The role of emotional and social intelligence competencies in engineer's effectiveness and engagement. *Career Development International*, *22*(1), 70–86.

Boyatzis, R. E., Rochford, K., & Jack, A. I. (2014). Antagonistic neural networks underlying differentiated leadership roles. *Frontiers in Human Neuroscience*, https://doi.org/10.3389/fnhum.2014.00114

Boyatzis, R. E., Rochford, K., & Taylor, S. N. (2015). The role of the positive emotional attractor as vision and shared vision: Toward effective leadership, relationships and engagement. *Frontiers in Psychology*, 21, 670. http://dx.doi.org/10.3389/fpsyg.2015.00670

Boyatzis, R. E., & Sala, F. (2004). Assessing emotional intelligence competencies. In G. Geher (Ed.), *The measurement of emotional intelligence* (pp. 147–180). Hauppauge, NY: Nova Science Publishers.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *J. Educ. Psychology*, *54*, 1–22. https://doi.org/10.1037/h0046743

Cattell, R. B. (1967). The theory of fluid and crystallized general intelligence checked at the 5–6 year-old level. *British Journal of Educational Psychology*, *37*, 209–224. https://doi.org/10.1111/j.2044-8279.1967.tb01930.x

Christie, R., & Geis, F. (1970). *Studies in Machiavellianism*. New York: Academic Press.

Clayton, B. (2014). Shared vision in corporate acquisitions. *Frontiers in Psychology*, *5*, 1335. https://doi.org/10.3389/fpsyg.2014.01466

Cote, S., & Miners, C. T. H. (2006). Emotional intelligence, cognitive intelligence and job performance. *Administrative Science Quarterly*, *51*, 1–28.

Crooks, L. A., Campbell, J. T., & Rock, D. A. (1979). *Predicting career progress of graduate students in management*. Princeton: Educational Testing Service.

Decety, J., & Batson, C. D. (2007). Social neuroscience approaches to interpersonal sensitivity. *Social Neuroscience*, *2*, 151–157.

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology*, *71*(2), 390–405.

Fischer, T., Dietz, J., & Antonakis, J. (2017). Leadership process models: A review and synthesis. *Journal of Management*, *43*(6), 1726–1753.

Fontana, A. F., Rosenberg, R. L., Marcus, J. L., & Kerns, R. D. (1987). Type A behavior pattern, inhibited power motivation, and activity inhibition. *Journal of Personality and Social Psychology*, 52, 177–183.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anti-correlated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102

French, S. E., & Jack, A. I. (2014). Dehumanizing the enemy: The intersection of neuroethics and military ethics. In D. Whetham (Ed.), *The responsibility to protect: Alternative perspectives*. Leiden: Martinus Nijhoff.

Franklin B. (2012). *The autobiography of Benjamin Franklin, Norton Critical Edition* (ed. J. Chaplin). New York: W. W. Norton.

Friedman, J., Jack, A. I., Rochford, K., & Boyatzis, R. (2015). Antagonistic neural networks underlying organizational behavior. *Organizational Neuroscience (Monographs in Leadership and Management, Volume 7)*, 115–141.

Frey, M. C., & Detterman, D. (2004). Scholastic achievement or *g*? The relationship between the Scholastic Achievement Test (SAT) and General Cognitive Ability (*g*). *Psychological Science*, *15*(6), 373–378.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Goleman, D. (1995), *Emotional intelligence*. New York: Bantam Books.

Graen, G. B., & Uhl-Bien, M. (1995). Relationship-based approach to leadership: Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective. *Leadership Quarterly*, *6*, 219–247.

Greenleaf, R. (1970/2015). *The servant as leader*. Atlanta, GA: The Greenleaf Center for Servant Leadership.

Grijalva, E., Harms, P. D., Newman, D. A., Gaddis, B. H., & Fraley, R. C. (2015). Narcissism and leadership: A meta-analytic review of the linear and non-linear relationships. *Personnel Psychology*, *68*(1), 1–47.

Grossman, I., Na, J., Varnum, M. E. W., Kitayama, S., & Nisbett, R. E. (2013). A route to well-being: Intelligence versus wise reasoning. *Journal of Experimental Psychology General*, *1423*(3), 944–953. https://doi.org/10.1037/a0029560

Hedlund, J., Forsythe, G. B., Horvath, J. A., Williams, W. M., Snook, S., & Sternberg, R. J. (2002). Identifying and assessing tacit knowledge: Understanding the practical intelligence of military leaders. *Leadership Quarterly*, *14*, 117–140.

Higgins, E. T. (1987). Self-discrepancy theory: Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*, 319–340.

Jack, A. I., Boyatzis, R. E., Khawaja, M. S., Passarelli, A. M., & Leckie, R. L. (2013). Visioning in the brain: An fMRI study of inspirational coaching and mentoring. *Social Neuroscience*, *8*(4), 369–384. https://doi.org/10.1080/17470919.2013.808259

Jack, A. I., Dawson, A. J., Begany, K. L., Leckie, R. L., Barry, K. P., Ciccia, A. H., & Snyder, A. Z. (2012). fMRI reveals reciprocal inhibition between social and physical cognitive domains. *NeuroImage*, *66*, 385–401. https://doi.org/10.1016/j.neuroimage.2012.10.061

Jack, A. I., Dawson, A. J., & Norr, M. (2013). Seeing human: Distinct and overlapping neural signatures associated with two forms of dehumanization. *NeuroImage*, *79*(1), 313–328.

Jacobs, R. L., & McClelland, D. C. (1994). Moving up the corporate ladder: A longitudinal study of the leadership motive pattern and managerial success in women and women. *Consulting Psychology Journal Practice and Research*, *46*, 32–41.

James, W. (1890). *The principles of psychology*. New York: Henry Holt.

Jenkins, S. R. (1994). Need for power and women's careers over 14 years: Structural power, job satisfaction, and motive change. *Journal of Personality and Social Psychology*, *66*, 155–165.

Joseph, D., Jin, J., Newman, D., & O'Boyle, E. H. (2014). Why does self-reported emotional intelligence predict job performance? A meta-analytic investigation of mixed EI. *Journal of Applied Psychology*, *100*, 298–342.

Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes*, *51*(2), 296–312.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.
Kantabutra, S., & Avery, G. C. (2010). The power of vision: Statements that resonate. *Journal of Business Strategy*, *31*(1), 37–45.
Kendall, L. (2016). A theory of micro-level dynamic capabilities: How technology leaders innovate with human connection. Unpublished doctoral dissertation, Case Western Reserve University.
Law, K. S., Wong, C., & Song, L.J. (2004). The construct and criterion validity of emotional intelligence and its potential utility in management research. *Journal of Applied Psychology*, *87*(3), 483–496.
LeBreton, J. M., Schiverdecker, L. K., & Grtimaldi, E. M. (2018). The dark triad and work-place behavior. *Annual Review of Organizational Psychology and Organizational Behavior*, *5*, 387–414. https://doi.org/10.1146/annurev-orgpsych-032117–104451
Lewin, K., Lippitt, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created "social climates." *Journal of Social Psychology*, *10*, 271–299.
Mahon, E. G., Taylor, S. N., & Boyatzis, R. E. (2014). Antecedents of organizational and job engagement: Exploring emotional and social intelligence as moderators. *Frontiers in Psychology*, *5*, 129–139. https://doi.org/10.3389/fpsyg.2014.01322
Martin, A., & Weisberg, J. (2003). Neural foundations for understanding social and mechanical concepts. *Cognitive Neuropsychology*, *20*(3), 575–587.
Matthews, G., Emo, A. K., Funke, G., Zeidner, M., Roberts, R. D., Costa, P. T., Jr., & Schulze, R. (2006). Emotional intelligence, personality, and task-induced stress. *Journal of Experimental Psychology: Applied*, *12*(2), 96–107.
Mayer, J. D., Salovey, P., & Caruso, D. R. (2000). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, *27*(4), 267–298.
McClelland, D. C. (1973). Testing for competence rather than intelligence. *American Psychologist*, *28*(1), 1–40.
McClelland, D. C. (1975). *Power: The inner experience*. New York: Irvington Press.
McClelland, D. C. (1985). *Human motivation*. Glenview, IL: Scott, Foresman & Co.
McClelland, D. C., & Boyatzis, R. E. (1982). The leadership motive pattern and long term success in management. *Journal of Applied Psychology*, *67*, 737–743.
McGregor, D. (1960). *The human side of enterprise*. New York: McGraw Hill.
Miao, C., Humphrey, R., & Qian, S. (2016). Leader emotional intelligence and subordinate job satisfaction: A meta-analysis of main, mediator, and moderator effects. *Personality and Individual Differences*, *102*, 13–24.
Miao, C., Humphrey, R. H., & Qian, S. (2017). A meta-analysis of emotional intelligence effects on job satisfaction mediated by job resources, and a test of moderators. *Personality and Individual Differences*. *116*, 281–288.
Miao, C., Humphrey, R. H., Qian, S. (2018a). Emotional intelligence and authentic leadership: A meta-analysis. *Leadership and Organizational Development Journal*, *39* (50), 679–690.
Miao, C., Humphrey, R. H., Qian, S. (2018b). A cross-cultural meta-analysis of ho leader emotional intelligence influences subordinate task performance and organizational citizenship. *Journal of World Business*, *53*, 463–474.
Miller, S. P. (2014). Next-generation leadership development in family businesses: The critical roles of shared vision and family climate. *Frontiers in Psychology*, *5*, 1335. https://doi.org/10.3389/fpsyg.2014.01335
Mosteo, L., Batista-Foguet, J. M., McKeever, J. D., & Serlavos, R. (2016). Understanding cognitive-emotional processing through a coaching process: The influence of

coaching on vision, goal-directed energy, and resilience. *The Journal of Applied Behavioral Science*, *52*(1), 1–33. https://doi.org/10.1177/0021886315600070

Neff, J. E. (2015). Shared vision and family firm performance. *Frontiers in Psychology*, *5*, 1335. https://doi.org/10.3389/fpsyg.2014.01335

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern. D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, *67*(2), 130–159. https://doi.org/10.1037/a0026699

O'Boyle, E. H., Jr., Humphrey, R. H., Pollack, J. M., Hawver, T. H., & Story, P. A. (2011). The relation between emotional intelligence and job performance: A meta-analysis. *Journal of Organizational Behavior*, *32*, 788–818.

Oh, I.-S., Schmidt, F. L., Shaffer, J. A., & Le, H. (2008). The Graduate Management Admissions Test (GMAT) is even more valid than we thought: A new development in meta-analysis and its implications for the validity of the GMAT. *Academy of Management Learning and Education*, *7*(4), 563–570.

O'Reilly, C. A., III, & Chatman, J. A. (1994). Working smarter and harder: A longitudinal study of managerial success. *Administrative Science Quarterly*, *39*, 603–627.

Overbeke, K., Bilimoria, D., & Somers, T. (2015). The critical role of the "ideal self" in the self-selection of next-generation leaders of family businesses. *Frontiers in Psychology*, *5*, 1335. https://doi.org/10.3389/fpsyg.2014.01335

Paulhus, D. L., & Williams, K.M. (2002). The dark triad of personality. *Journal of Research in Personality*, 36, 556–563. https://doi.org/10.1016/s0092-6566(02)00505-6

Petrides, K. V., & Furnham, A. (2000). On the dimensional structure of emotional intelligence. *Personality and Individual Differences*, *29*, 313–320.

Petrides, K. V., & Furnham, A. (2001). Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. *European Journal of Personality*, *17*, 425–448.

Pittenger, L. M. (2015). IT professionals: Recipe for engagement. *Frontiers in Psychology*, *5*, 1335. https://doi.org/10.3389/fpsyg.2015.00623

Prabhakaran, V., Smith, J. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology*, *33* (1), 43–63.

Quinn, J. F. (2015). The effect of vision and compassion upon role factors in physician leadership. *Frontiers in Psychology*, *6*, 442. https://doi.org/10.3389/fpsyg.2015.00442

Riggio, R. E., Chaleff, I., & Lipman-Blumen, J. (2008). *The art of followership: How great followers create great leaders and organizations*. New York: John Wiley & Sons.

Rochford, K. C., Jack, A. I., Boyatzis, R. E., & French, S. E. (2016). Ethical leadership as a balance between opposing neural networks. *Journal of Business Ethics*, *144*(4), 755–770.

Ross, D. (1925/1980). *Aristotle The Nicomachean Ethics: Translated with an Introduction* (ed. J. L. Ackrill & J. O. Urmson). Oxford: Oxford University Press.,

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, *9*, 185–211.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional Intelligence. *Personality and Individual Differences*, *25*, 167–177.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology*, *3*. 292–316. https://doi.org/10.1037/1089-2680.3.4.292

Sternberg, R. J. (2011). The theory of successful intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 504–527). New York: Cambridge University Press.

Sternberg, R. J., & Detterman, D. K. (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex.

Sternbger, R. J., & Hedlund, J. (2002). Practical intelligence, *g* and work psychology. *Human Performance*, *15*, 143–160.

Tee, E., Ashkanasy, N. M., & Paulsen, N. (2013). The influence of follower mood on leader mood and task performance: An affective, follower-centric perspective of leadership. *Leadership Quarterly*, *24*(4), 496–515. https://doi.org/10.1016/j.leaqua.2013.03.005

Thornton, J. C. (2015). Shared vision: Mediating CSR in US food and beverage SMEs. *Frontiers in Psychology*, *5*, 1335. https://doi.org/10.3389/fpsyg.2015.00781

Van Overwalle, F. (2010). A dissociation between social mentalizing and general reasoning. *NeuroImage*, *54*(2), 1589–1599. https://doi.org/10.1016/j.neuroimage.2010.09.043

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, *100*(6), 3328–3342.

Winter, D.G. (1973). *The power motive*. New York: Free Press.

Yukl, G. A. (1998). *Leadership in organizations* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Yukl, G. A., & Van Fleet, D. D. (1990). Theory and research on leadership in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 3). Palo Alto, CA: Consulting Psychologists Press.

34 Cultural Intelligence

Soon Ang, Kok Yee Ng, and Thomas Rockstuhl

Introduction and Historical Background

Cultural intelligence refers to an individual's capability to function effectively in situations characterized by cultural diversity (Ang & Van Dyne, 2008; Earley & Ang, 2003). Earley and Ang (2003) first introduced the concept of cultural intelligence in their book *Cultural Intelligence: Individual Interactions Across Cultures*.

Cultural intelligence (CQ) was conceived at the turn of the twenty-first century, when the world was experiencing unprecedented globalization and interconnectedness, driven by advances in communication and transportation technologies. It was also conceived at a time in which ideological clashes and cultural conflict culminated in the tragic events of September 11, 2001. Nobel Prize laureate Elie Wiesel identified "cultural hatred" – hatred directed toward culturally different individuals – as *the* major source of problems between people, across all times. The *Los Angeles Times* estimates that there are more than fifty hot spots in the world where cultural conflicts occur every day. Amid the promises and perils of globalization, CQ becomes an essential capability for individuals, not only to harness the benefits of cultural diversity but also to manage the conflicts that come with it.

The driving question behind the idea of cultural intelligence is: *Why do some but not other individuals easily and effectively adapt their views and behaviors cross-culturally?* (Van Dyne, Ang, & Livermore, 2010). This question has long interested researchers across diverse disciplines in psychology, sociology, management, health care, the military, education, and other fields. Thus, it is not surprising that a wide array of frameworks and intercultural instruments purport to assess cultural competencies (see Leung, Ang, & Tan, 2014).

A challenge to this body of work, however, is the lack of a coherent theoretical foundation. In a review of the literature, Gelfand, Imai, and Fehr (2008) described the existing cultural competency models as suffering from the "jingle and jangle fallacy – where constructs with the same meaning are labeled differently while constructs with different meanings are labeled similarly" (p. 375). As a result, concerns of construct validity arise and compromise the practical utility of the concept.

It is within this context that the concept of CQ was formulated. Earley and Ang (2003) conceptualized CQ as a set of four capabilities based on Sternberg's (1986) theory of multiple loci of intelligence. Accordingly, CQ is a "cleaner" construct that offers a theoretically grounded, comprehensive, and coherent framework.

Since 2003, the concept of CQ has attracted significant attention worldwide and across diverse disciplines, including applied, cognitive, and social psychology; mental health; management; education; decision sciences; the military; engineering; and religious missions. This rapid growth of research attention on CQ attests to Matsumoto and Hwang's (2013) conclusion on the "promising evidence for assessing CQ" (p. 867).

More importantly, we have witnessed significant and exciting advancements in the theorizing and empirical research on CQ in recent years. Many of these advancements address the future research directions that we offered in the 2011 review in the first edition of this handbook. They include a deeper conceptualization of CQ and its dimensions, complementary measures of CQ, and a richer nomological network. In this chapter, we provide a comprehensive and up-to-date review of research on CQ, with a focus on these recent developments. We conclude with future directions to stimulate new theorizing and empirical research and to foster practical applications in diverse countries and cultures across the globe.

Conceptualization of Cultural Intelligence

Although early research tended to view intelligence narrowly as the ability to grasp concepts and solve problems in academic settings, there is now a consensus that intelligence applies beyond the classroom. The growing interest in "real-world" intelligence has identified new types of nonacademic intelligences (Sternberg, 1997) that focus on specific content domains such as social intelligence (Thorndike & Stein, 1937), emotional intelligence (Mayer & Salovey, 1993), and practical intelligence (Sternberg & Wagner, 2000).

Motivated by the practical reality of globalization, CQ builds on some of these ideas but with a focus on a specific domain – intercultural settings (Earley & Ang, 2003). Just as emotional intelligence (EQ) complements cognitive intelligence (IQ) in predicting work effectiveness in interdependent domestic work contexts (Joseph & Newman, 2010), CQ is another important form of intelligence that can increase our prediction of effectiveness in coping with diversity and functioning in new cultural settings (Rockstuhl et al., 2011).

Cultural Intelligence as a Multidimensional Construct

Earley and Ang (2003) conceptualized CQ as a multidimensional construct based on Sternberg's (1986) "multiple loci" of intelligence argument. Specifically, Sternberg proposed that there are different loci of intelligence within the person – metacognition, cognition, motivation, and behavior – and that a more complete understanding of intelligence requires the consideration of all four loci.

Adopting the multiple loci argument, Earley and Ang (2003) described cultural intelligence as an aggregate multidimensional construct that comprises four dimensions (commonly referred to as the four factors of CQ) – metacognitive, cognitive, motivational, and behavioral CQ. In a major conceptual refinement, Van Dyne and colleagues (2012) advanced more granular subdimensions to allow for a better-articulated conceptual space for each CQ factor. We describe the four CQ factors and their respective subdimensions next (see also Table 34.1 for a summary of the CQ factors and example items).

Metacognitive CQ. This CQ factor refers to an individual's level of conscious cultural awareness during cross-cultural interactions. Metacognitive CQ involves higher-level cognitive strategies – strategies that allow individuals to develop new heuristics and rules for social interactions in novel cultural environments. More specifically, Van Dyne and colleagues (2012) proposed three subdimensions of metacognitive CQ. They are planning (i.e., strategizing before intercultural encounters), awareness (i.e., having real-time consciousness of cultural influences on self, others, and the situation), and checking (i.e., reviewing assumptions and adjusting mental models when actual experiences differ from expectations).

People with high metacognitive CQ are more likely to be deliberate and intentional when they encounter cross-cultural interactions. They tend to plan for an interaction by taking the perspective of culturally diverse others and anticipating the actions and reactions of various parties in that cultural context. During the interaction, they are more likely to pay attention to meaningful cues, suspend judgments until sufficient information is available for accurate sensemaking, and adjust their original assumptions on new information. For instance, a Western executive with high metacognitive CQ may be more aware and mindful about when to speak up during meetings with Asians. They may also consciously look for cues during meetings to interpret what is said and not said, to develop a more accurate understanding of their Asian counterparts.

Metacognitive CQ is a critical component of CQ. It promotes active thinking about people and situations in different cultural settings, challenges rigid reliance on culturally bounded thinking and assumptions, and drives individuals to adapt and revise their strategies dynamically to achieve desired outcomes in cross-cultural encounters.

Cognitive CQ. While metacognitive CQ focuses on higher-order cognitive processes, cognitive CQ reflects knowledge of norms, practices, and conventions in different cultures acquired from education and personal experiences. It is an individual's level of knowledge of the cultural environment and knowledge of the self as embedded in the cultural context of the environment. Cognitive CQ includes subdimensions of culture-general knowledge (i.e., declarative knowledge of the universal elements that constitute a cultural environment) and culture-specific knowledge (i.e., declarative and procedural knowledge about cultural universals in a specific domain, for instance leading people across different cultures) (Van Dyne et al., 2012).

Table 34.1 *Summary of CQ factors with sample items from the 11-Dimension Expanded CQ Scale (E-CQS).* [a]

CQ dimensions	*Definition* / Sample Items
METACOGNITIVE CQ	
Subdimensions	*An individual's level of conscious cultural awareness and executive processing during intercultural interactions.*
Planning	I develop action plans before interacting with people from a different culture
Awareness	I am aware of how my culture influences my interactions with people from different cultures
Checking	I adjust my understanding of a culture while I interact with people from that culture
COGNITIVE CQ	
Subdimensions	*An individual's knowledge structures about cultural institutions, norms, practices, and conventions in different cultural settings.*
Culture-General Knowledge	I can describe the different cultural value frameworks that explain behaviors around the world
Context-Specific Knowledge	I can describe the ways that leadership styles differ across cultural settings
MOTIVATIONAL CQ	
Subdimensions	*An individual's capability to direct attention and energy toward learning about and functioning in situations characterized by cultural differences.*
Intrinsic Interest	I truly enjoy interacting with people from different cultures
Extrinsic Interest	I value the status I would gain from living or working in a different culture
Self-Efficacy to Adjust	I am confident that I can persist in coping with living conditions in different cultures
BEHAVIORAL CQ	
Subdimensions	*An individual's capability to enact a wide repertoire of verbal and nonverbal actions when interacting with people from different cultures.*
Verbal behavior	I change my use of pause and silence to suit different cultural situations
Nonverbal behavior	I modify how close or far apart I stand when interacting with people from different cultures
Speech Acts	I modify the way I disagree with others to fit the cultural setting

[a] © Cultural Intelligence Center 2011. Used by permission of Cultural Intelligence Center. Note: Use of these items and scale is granted to academic researchers for research purposes only. For information on using the items and scale for purposes other than academic research (e.g., consultants and nonacademic organizations), please send an email to cquery@culturalq.com

Individuals with high cognitive CQ are likely to have more elaborate knowledge structures about cultural institutions, norms, practices, and conventions in different cultural settings. Understanding the elements that constitute the cultural environment helps individuals appreciate how these elements shape and cause patterns of behaviors and interactions within a culture and why behaviors and interactions differ across different cultural environments (Ang & Van Dyne, 2008). For instance, understanding how a family system works becomes critically relevant when a leader develops ways to reward and motivate their employees in cultures that expect respect and care for senior members of their extended family.

Cognitive CQ is a critical component of CQ because it aids in making isomorphic attributions of behaviors observed in different cultural contexts. This in turn is critical for sound judgment and decision-making in culturally diverse settings (Ang et al., 2007). In addition, cognitive CQ can help to reduce uncertainty and anxiety during intercultural interactions.

Motivational CQ. Broadly, motivational CQ can be understood as approach versus avoidance motivation (Elliot & Covington, 2001). Those with higher motivational CQ are more likely to approach, rather than avoid, intercultural situations. More specifically, motivational CQ reflects the capability to direct attention and energy toward learning about, and functioning in, culturally diverse situations. Kanfer and Heggestad (1997) argued that such motivational capacities "provide agentic control of affect, cognition and behavior that facilitates goal accomplishment" (p. 39). The subdimensions of motivational CQ include intrinsic interest (i.e., valuing intercultural experiences in and of themselves), extrinsic interest (i.e., valuing the tangible, personal benefits that can be derived from intercultural experiences), and self-efficacy to adjust (i.e., having task-specific confidence in intercultural situations).

People with high motivational CQ are more likely to be attracted to intercultural situations because they value the tangible and intangible benefits of these interactions. They also tend to be more confident in coping with the inherent challenges of cultural differences. Motivational CQ is a critical component of CQ because it determines whether a person will approach or avoid intercultural situations and whether such interactions will be sustained. Possessing such a drive is important for crossing cultures since intercultural interactions are often fraught with cultural and language challenges. For example, a Chinese executive who likes and values interacting with people from other cultures may be less hesitant to approach a colleague from Japan, even if they do not speak Japanese or English well. In contrast, another Chinese executive who places little value on cross-cultural encounters will more likely avoid such cross-cultural interactions even if language is not a barrier.

Behavioral CQ. Behavioral CQ reflects an individual's capability to enact a wide range of verbal and nonverbal actions when interacting with people from different cultures. Behavioral CQ enables people to manage and regulate social behaviors in intercultural encounters so as to minimize misperceptions and misattributions (Gudykunst, 1993). The subdimensions of behavioral CQ include verbal behavior (i.e., flexibility in vocalization, including accent, tone), nonverbal behavior (flexibility in communication via gestures, facial expressions, and body language, etc.), and speech acts (flexibility in using words to communicate specific types of messages such as requests, invitations, apologies, gratitude, disagreement) (Van Dyne et al., 2012).

People with high behavioral CQ are more likely to overcome the natural human tendency to exhibit habitual behaviors. Instead, they are more likely to display a wide repertoire of verbal behaviors, nonverbal behaviors, and speech acts to suit the cultural context. For example, a traveler with high behavioral CQ and a low-context communication style (i.e., who prefers to convey meaning explicitly and without reference to contextual understanding; Adair et al., 2016; Hall, 1959) may show behavioral flexibility and display high-context communication behaviors (e.g., say "no" indirectly) when it is more culturally appropriate to do so.

Behavioral CQ is a critical component of CQ because behaviors are salient and visible to others during social interactions. Mental capabilities for cultural understanding and motivation are rendered useless if they are not complemented with the ability to exhibit appropriate verbal and nonverbal actions during intercultural interactions (Hall, 1959). As such, the three behavioral CQ subdimensions may be the most critical aspects of CQ from the perspective of observers.

Conceptual Distinctiveness

To further clarify the nature of CQ, we describe what CQ is not. Specifically, we distinguish CQ from personality and other forms of intelligence.

Cultural intelligence and personality. CQ is a set of abilities, which refer to personal characteristics that relate to the capability to perform the behavior of interest. As such, CQ is clearly different from personality traits, which are nonability individual differences. CQ focuses on culturally relevant capabilities. Thus, it is more specific than personality or general cognitive ability. In addition, CQ is malleable and can be enhanced through experience, education, and training. Hence, CQ is a state-like individual difference that can evolve over time, while personality is a relatively stable, trait-like individual difference.

Cultural intelligence in relation to other intelligence constructs. CQ is similar to general cognitive ability (e.g., Schmidt & Hunter, 1998) and emotional intelligence (Mayer & Salovey, 1993) because it deals with a set of abilities. CQ differs, however, from the two intelligences in the nature of the ability examined. General cognitive ability – the ability to learn – predicts performance across many jobs and settings but it is not specific to certain contexts – such as culturally diverse situations. In addition, it does not include behavioral or motivational aspects of intelligence. Emotional intelligence (EQ) is the ability to deal with personal emotions. Thus, it is similar to CQ because it goes beyond academic and mental intelligence but it differs from CQ because it focuses on the general ability to perceive and manage emotions without consideration of the cultural context. Given that emotional cues are symbolically constructed within a culture, a person who is emotionally intelligent in one culture is not necessarily emotionally intelligent in another culture (Earley & Ang, 2003). In other words, EQ is culture-bound. In contrast, CQ is not culture-specific and refers to a general set of capabilities with relevance to situations characterized by cultural diversity.

Empirical research has supported the conceptual distinctiveness of CQ from cognitive ability (e.g., Klafehn, Li, & Chiu, 2013; Rockstuhl et al., 2011; Varela & Gatlin-Watts, 2014) and EQ (e.g., Groves, Feyerherm, & Gu, 2015; Lin, Chen, & Song, 2012; Şahin et al., 2013). In addition, empirical evidence suggests that CQ has incremental predictive validity over cognitive ability and EQ in predicting cross-border leadership effectiveness (Rockstuhl et al., 2011) and negotiation effectiveness (Groves et al., 2015).

Measurement of Cultural Intelligence

Individual differences in cultural intelligence are measured using diverse methods, including report-based and performance-based measures. We emphasize that alternative measures of CQ are complementary and capture unique information about a person's CQ. Rather than embark on a search for the methodological "Holy Grail," we suggest that the choice of the measure should be guided by the nature of the outcome of interest (Campbell & Fiske, 1959). For instance, research shows that matching the measure and outcome in terms of observability could improve prediction (Lance et al., 2008). This suggests that a self-reported measure of CQ may be more suitable in predicting affective states and, hence, outcomes such as intercultural adaptation. By contrast, a peer-reported measure of CQ may be more suitable for observable behaviors in intercultural interaction and, hence, for outcomes such as interpersonal effectiveness.

Report-Based Measures of Cultural Intelligence

Ang and associates (2007) and Van Dyne, Ang, and Koh (2008) initiated a series of studies to develop, validate, and cross-validate ($N > 1{,}500$) the first twenty-item Cultural Intelligence Scale (CQS). From an initial item pool of fifty-three items (thirteen to fourteen items per CQ dimension), ten best items for each dimension (i.e., a total of forty items) were retained based on clarity, readability, and definitional fidelity. To validate these items, Ang and colleagues conducted five studies. In Study 1, business school undergraduates in Singapore ($N = 576$) completed the forty items. Items with high residuals, low factor loadings, small standard deviations or extreme means, and low item-to-total correlations were dropped, resulting in twenty items with the strongest psychometric properties. Confirmatory factor analysis (CFA) (LISREL 8: maximum likelihood estimation and correlated factors) demonstrated a good fit for the hypothesized four-factor model with the data. Scholars interested in the use of the CQS for research purposes may contact the first author for information.

In a series of four studies (Studies 2–5), Ang and colleagues (2007) cross-validated the CQS across samples, time, and countries to assess its factor validity and cross-cultural measurement equivalence. CFA results confirmed the four-factor structure in different samples and demonstrated the temporal stability of the scale across a four-month period, as well as measurement equivalence across Singapore and US samples. Finally, they cross-validated the CQS across methods using an observer-report version of the scale. Multitrait-multimethod (MTMM) analysis provided evidence of convergent, discriminant, and criterion validity of the CQS across self- and peer-ratings of CQ and interaction adjustment in a sample of executive managers.

Subsequent studies also support the psychometrics, factor structure validity, and generalizability of the CQS. The four-factor structure has been replicated across multinational samples (e.g., Shannon & Begley, 2008; Shokef & Erez, 2008; Ward et al., 2009) and multiple countries, including India (Jyoti & Kour, 2015), Korea (T. Moon 2010a), the Philippines (Presbitero, 2016), Turkey (Şahin et al., 2013), and Saudi Arabia (Al-Dossary, 2016). Across studies, the CQS also shows good internal consistency reliability.

More recently, Van Dyne and colleagues (2012) introduced the Expanded CQS (E-CQS; see Table 34.1 for example items), a thirty-seven–item scale that measures subdimensions of the four CQ factors. They also provided evidence for the convergent and discriminant validity of the E-CQS in a sample of 286 individuals from more than thirty countries.

Performance-Based Measures of Cultural Intelligence

Moving beyond the report-based measures of CQS, Ang, Rockstuhl, and Ng (2014) developed a performance-based measure of CQ in the form of an intercultural situational judgment test (iSJT). The iSJT presents respondents with short video scenarios of intercultural conflict in the workplace. For example, the video may show the conflict between a team member who prefers to schedule a detailed work plan based on cultural values of high uncertainty avoidance and a team member who prefers to proceed with trial and error based on low uncertainty avoidance. At the end of each video scenario, respondents are asked what they would do next in this situation and their open-ended responses are scored for how effectively they would resolve the underlying cultural conflict.

Rockstuhl and colleagues (2015) validated the iSJT across a series of studies with students working in culturally diverse teams, as well as professionals working in multicultural consulting teams. Across all three studies, performance on the iSJT predicted peer-rated task performance and citizenship behaviors, over and above Big Five personality, general cognitive ability, international experience, and demographic characteristics. In sum, emerging empirical evidence highlights the potential of the iSJT to complement report-based measures of CQ as a predictor of intercultural effectiveness.

Evidence of CQ Nomological Network

Empirical research on CQ has flourished since the publication of Ang and colleagues' (2007) validated CQS. Today, we have a richer understanding of the nomological network of CQ and accumulated evidence of its predictive validity in a myriad of contexts and disciplines. In this section, we review empirical articles organized around the antecedents and outcomes of CQ. We further organize research on the outcomes of CQ by three levels of analysis – individual, dyad/team, and firm.

Antecedents of CQ

Research has examined three primary types of antecedents of CQ. They are personality, identity, and international experience.

Personality. In the section on conceptual distinctiveness, we distinguished personality, a distal, trait-like individual difference, from CQ, which is a state-like and malleable capability. Consistent with this conceptual distinction, Ang, Van Dyne, and Koh (2006) showed that the Big Five personality traits are distinct from, but related to, CQ. Further, CQ is shown to mediate the effects of personality traits such as openness to experience on adaptive performance (Oolders et al., 2008) and job performance (Sri Ramalu et al., 2012).

Identity. A relatively new antecedent of CQ is multicultural identity. Multiculturals, defined as people who identify with two or more cultures (Brannen & Thomas, 2010), are increasingly common because of immigration, intercultural marriages, or extensive multicultural experiences. This poses an interesting question for CQ research: Are multiculturals more culturally intelligent?

Addressing this question using professionals in an international agency in the Netherlands, Korzilius, Bücker, and Beerlage (2017) found that individuals who reported greater multiculturalism tend to have higher overall CQ. In another study of bicultural students in the United States, Dheer and Lenartowicz (2018) found that biculturals who perceive their cultural identities as integrated and compatible (versus dissociated or difficult to integrate) are more likely to have higher CQ. Interestingly, a superordinate identification with a global culture appears to overcome challenges associated with nonintegrated identities related to specific cultures. In a study of MBA students working in culturally diverse teams, Lee and colleagues (2018) found that, when global identity was low, students with integrated or balanced home and host country identities were perceived as more culturally intelligent by their peers, compared with students who identified with either their home or their host country. However, when global identity was high, CQ was high regardless of whether or not home and host country identities were balanced.

International experience. The relationship between international experience and CQ is a fairly well-established one. A number of studies show that international work experience (e.g., Crowne, 2008; Shannon & Begley, 2008; Tay, Westman, & Chia,

2008) and nonwork experience (e.g., Moon, Choi, & Jung, 2012; Tarique & Takeuchi, 2008) are positively related to CQ. Studies have also examined more complex models, including the role of CQ in mediating the effects of international experience on intercultural outcomes and moderators of the international experience – CQ relationship. For instance, in a study of expatriates in Korea, H. Moon and colleagues (2012) found that CQ mediated the effects of prior international work and nonwork experiences on intercultural adaptation.

Interestingly, CQ does not always mediate the effect of international experience. In a study by Kim and Van Dyne (2012), CQ mediates the effects of international experience (i.e., number of countries lived in) on international leadership potential only for majority members but not for minority members. Another moderator of the international experience – CQ relationship – is learning styles. In a study of international MBA students, Li, Mobley, and Kelly (2013) found that people with a divergent learning style (which emphasizes engaging in concrete experiences and reflecting on one's observations) are more likely to translate their international experience into higher CQ.

Outcomes of CQ – Individual Level

Much of the research on outcomes of CQ is conducted at the individual level and can be grouped into four major outcomes – adaptation, job performance, leadership, and the change in CQ as a result of an intervention. Within each outcome, we further organize our review by the study context, which primarily involves global professionals, foreign workers and migrants, the military, and students.

Adaptation. Cultural adaptation comprises two dimensions: sociocultural and psychological adjustment. Sociocultural adaptation includes general adjustment to foreign living conditions; work adjustment to foreign work culture; and interactional adjustment – the extent of getting along with those from another culture. Psychological adjustment refers to a person's general mental well-being when immersed in another culture (Church, 1982).

The majority of adaptation studies focus on global professionals and expatriates. In an early study, Ang and colleagues (2007) demonstrated that IT consulting professionals with higher motivational and behavioral CQ have better general, work, and interactional adjustment, as well as enhanced mental well-being in multicultural settings. Since then, a number of studies have demonstrated that CQ, especially motivational CQ, is an important predictor of expatriate adjustment (e.g., Chen et al., 2010; Firth et al., 2014; Guðmundsdóttir, 2015; Huff, Song, & Gresch, 2014; Zhang & Oczkowski, 2016). For instance, Zhang and Oczkowski (2016) found that motivational CQ is the only CQ factor that predicted both sociocultural and psychological adjustment in a sample of Australian expatriates. Similarly, Huff and colleagues (2014) found that motivational CQ predicted general, interactional, and work adjustment of expatriates in Japan.

In a multilevel study of expatriates across thirty-one subsidiaries, Chen and colleagues (2010) showed that the effects of expatriates' motivational CQ on work adjustment depend on the context. Specifically, results showed that motivational CQ affects expatriates' work adjustment more when cultural distance and subsidiary support are low. This is an interesting finding as it suggests that motivational CQ is less likely to be activated when the context renders it unnecessary for expatriates to put in effort (e.g., high subsidiary support). At the same time, motivational CQ alone is not sufficient when the cultural environment is difficult, such as when the cultural distance between the home and the host cultures is high.

In the first longitudinal study that tracks the adjustment of expatriates, Firth and colleagues (2014) found that motivational CQ positively relates to initial levels of work adjustment. Surprisingly, motivational CQ is negatively related to subsequent change in work adjustment. Firth and colleagues explained that, consistent with control theory, expatriates with higher initial adjustment are more likely to experience smaller discrepancy between actual and desired levels of work adjustment and, hence, tend to devote less effort subsequently.

Research has also begun to examine CQ in foreign workers (e.g., Chen, 2015; Le, Jiang, & Nielsen, 2018) and students (e.g., Crowne & Engle, 2016; Peng, Van Dyne, & Oh, 2015; Racicot & Ferry, 2016; Shu, McAbee, & Ayman, 2017). For instance, in a study of foreign laborers in Taiwan, Chen (2015) found that CQ has an indirect effect on workers' job involvement via work adjustment. This effect is accentuated for workers who received intercultural training. Peng, Van Dyne, and Oh (2015) found that motivational CQ positively predicted the cultural well-being of students in a short-term business study-abroad program.

Job performance. An important outcome of CQ is performance, which includes work performance, creative performance, and negotiation performance.

Work performance is a multidimensional construct (Campbell, 1990) and empirical evidence shows that CQ predicts different aspects of performance. In the context of global professionals, Ang and colleagues (2007) showed that international managers with higher metacognitive CQ and cognitive CQ performed better at cultural decision-making and those with higher metacognitive CQ and behavioral CQ demonstrated higher task performance. In a study of call center employees in the Philippines, Presbitero (2017) found that motivational CQ mediates the effects of language ability on agents' service performance. In a study of expatriates, Chen and colleagues (2010) showed that motivational CQ influences job performance indirectly by enhancing cultural adjustment. Chen, Liu, and Portnoy (2012) showed that motivational CQ predicted cross-cultural sales but not overall sales of real estate agents in the United States and that this effect was stronger in firms with higher organizational-level motivational CQ.

In the context of the military, Şahin and Gürbüz (2014) examined CQ and adaptive performance – a form of work performance that emphasizes the management of changing work and novel requirements (Hesketh & Neal, 1999). Results based on a sample of Turkish troops deployed in the European Union Force in Bosnia and Herzegovina showed that soldiers with higher motivational CQ and behavioral CQ are more likely to display adaptive performance as rated by peers.

Negotiation performance is another important outcome of CQ. Studies on CQ and negotiation have mostly used negotiation tasks involving undergraduate and graduate students. At the individual level, Groves and colleagues (2015) found that cognitive and behavioral CQ affect negotiation performance indirectly through interest-based negotiation behaviors. Studying dyads of American and East Asian students, Imai and Gelfand (2010) found that the minimum motivational CQ score of the dyad predicted integrative negotiation behaviors, which in turn predicted joint profits. In an experimental study of intercultural negotiation and dispute mediation involving American and Turkish students, Salmon and colleagues (2013) found that the effects of the negotiation dyad's CQ are moderated by the style of mediation. Specifically, the dyad's motivational CQ has a positive effect on negotiation outcomes when there is no mediation or when the mediation is formulative in nature (i.e., moving parties forward by offering constructive and specific suggestions). Surprisingly, the dyad's motivational CQ is negatively related to the negotiation outcome when the mediation is manipulative in nature (i.e., moving parties off a previously held position through threats and rewards).

A relatively new outcome variable examined in CQ research is creative performance. In the context of global professionals, Xu and Chen (2017) found that expatriates' metacognitive CQ and motivational CQ affect cross-cultural job creativity through cultural learning. Similarly, Lorenz and colleagues (2018) found that metacognitive CQ and cognitive CQ influence expatriates' innovations in products, services, and processes because of their greater ability to recognize new opportunities. Chua, Morris, and Mor's (2012) study of executives from diverse backgrounds showed that metacognitive CQ increases affect-based trust among culturally different members of multicultural professional networks, which in turn fosters the exchange and cross-pollination of ideas.

In the context of students, Chua and Ng (2017) postulated an interesting interaction effect of cognitive and metacognitive CQ on individuals' creativity. Their findings showed that, while cognitive CQ provides domain knowledge that aids creativity, too much cognitive CQ (i.e., cultural knowledge) could be detrimental because of cognitive overload and entrenchment. This "too-much-of-a-good-thing" effect, however, is seen only for individuals with low metacognitive CQ. This finding suggests that individuals with high metacognitive CQ have better self-regulated mental processes to avoid the potential dark side of cognitive CQ when generating ideas for a cross-cultural creative task.

Global leadership. Global leadership has been a long-standing area of interest for CQ scholars. In the context of organizational leaders from culturally diverse backgrounds, Groves and Feyerherm (2011) showed that leaders' CQ predicts followers' ratings of leader performance and team performance only when the team is culturally diverse. CQ has no effects when team diversity is low.

In the military context, Rockstuhl and colleagues (2011) conducted a classic study that simultaneously examined domestic and global leadership effectiveness. Based on a sample of professional military officers who had leadership roles in both domestic and cross-border contexts (e.g., UN peacekeeping missions), Rockstuhl and colleagues (2011) showed that EQ is a stronger predictor of leadership effectiveness in domestic contexts while CQ is a stronger predictor of leadership effectiveness in cross-border contexts. General intelligence (IQ) predicts leadership effectiveness in both domestic and cross-border contexts.

In the context of students, Lisak and Erez (2015) examined a group of MBA students working in self-managing virtual teams to determine the characteristics of emergent leaders. Using logistic regressions, the authors found that MBA students with high CQ, coupled with high global identity and openness to cultural diversity, are more likely to emerge as leaders than other team members.

Change in CQ. In light of the predictive validity of CQ, an important question is whether and how CQ can be developed in individuals. There is now growing empirical evidence to show that direct and authentic intercultural experiences, when coupled with appropriate interventions, can facilitate the development of CQ. Most of these studies are conducted with students in education contexts such as working in virtual multicultural teams (Erez et al., 2013; Taras et al., 2013), intercultural contact experiences (MacNab, 2012), short-term overseas trips (Wood & Peters, 2014), study-abroad programs (Chao, Takeuchi, & Farh, 2017; Ramirez, 2016), and international service–learning programs (Engle & Crowne, 2014; Pless, Maak, & Stahl, 2011).

More recent studies have progressed beyond a pre-post assessment of CQ to examine trajectories of CQ development over multiple time points. In a study of Chinese exchange students in the United States, Wang and colleagues (2015) assessed students' CQ at four time points from predeparture to the third month in the United States. Results revealed four intriguing CQ trajectories – consistently high CQ scores, decreasing CQ scores, increasing CQ scores, or a sharp decrease in CQ scores over the first two months followed by a rebound. This important finding highlights the presence of boundary conditions in the development of CQ.

Existing research has suggested several boundary conditions that could affect the development of CQ. For instance, MacNab, Brislin, and Worthley (2012) found that higher-quality intercultural contact based on four conditions derived from contact theory (Allport, 1954): equal status, common goals, personalized contact, and authority support led to greater increases in CQ. Rosenblatt and colleagues (2013) replicated and extended this finding by showing that the increase is mediated by disconfirmation of expectations. In a similar vein, Alexandra (2018) found that students with lower social dominance orientation were more likely to change their stereotypes during intercultural contact and in turn were more likely to improve in their CQ.

In the context of the military, Şahin, Gürbüz and Köksal (2014) compared the CQ of military personnel before and after international deployments and found interesting personality moderators. Results showed that extraverted individuals are more likely to increase in their metacognitive and behavioral CQ, whereas those with greater openness to experience tend to increase in their motivational CQ.

Outcomes of CQ – Dyad/Team Level

With the prevalence of multicultural teams, a number of studies have examined the role of CQ in affecting team processes (e.g., Adair, Hideg, & Spence, 2013; Chen & Lin, 2013; Moynihan, Peterson, & Earley; 2006) and outcomes (e.g., Crotty & Brett, 2012; Magnusson, Schuster, & Taras, 2014). Most of these studies are conducted with self-managing student teams. In this section, we review studies that explicitly address team processes and outcomes at the dyad and team level.

In a study that examined dyads within teams, Rockstuhl and Ng (2008) found that higher metacognitive and cognitive CQ enhance affect-based trust in culturally diverse dyad partners in the team. Further, higher behavioral CQ displayed by a dyad partner positively influences affect-based trust in the dyad partner. CQ, however, has no effect on trust in culturally homogeneous dyads. In another study, Li and colleagues (2017) examined dyads comprising German and Chinese students working virtually on a project. Findings show that dyads with higher average CQ scores reported greater satisfaction with the virtual collaboration.

Using teams comprising undergraduate students, Adair and colleagues (2013) examined the relationship between team-level CQ and teams' shared values. Results showed that team-level metacognitive CQ and behavioral CQ are positively related to shared values in culturally heterogeneous teams but not in homogeneous teams. A surprising finding was that team-level motivational CQ and metacognitive CQ are negatively related to shared values in culturally homogeneous teams, suggesting that CQ could create unintended effects in teams that do not require it. In another study of 145 global virtual teams, Magnusson and colleagues (2014) found that teams' psychic distance (i.e., perception of differences among members) is positively related to team performance, mediated by team-level efforts. This mediated relationship is strengthened in teams with higher motivational CQ.

Outcomes of CQ – Firm Level

Ang and Inkpen (2008) developed a conceptual model of firm-level CQ that comprises three components: managerial CQ, competitive CQ, and structural CQ. Adopting a resource-based view of the firm, Ang and Inkpen argued that firm-level CQ is an important competitive resource for international business ventures. Complementing Ang and Inkpen's model with a dynamic capability framework, Moon (2010b) proposed three CQ organizational capabilities. They are (1) process capability (comprising cross-cultural coordination, organizational learning, and cross-cultural reconfiguration); (2) position capability (comprising managerial, competitive, and structural CQ); and (3) path capability (comprising cross-cultural initiation, experience, and resource fungibility).

Recent studies have empirically examined the role of CQ at the firm level in three different contexts – international alliances (Pesch & Bouncken, 2017; Yitmen, 2013), exporting firms (Magnusson et al., 2013; Golgeci, Swiatowiec-Szczepanska, & Raczkowski, 2017), and small businesses (Charoensukmongkol, 2015, 2016). For instance, in a study involving German firms in the photonics and biotechnology industries, Pesch and Bouncken (2017) found that firm-level managerial CQ is positively related to task discourse – the extent to which alliance partners communicate and challenge each other's views and problem-solving methods. Further, managerial CQ determines whether socializing practices (e.g., social events, interorganizational teams) are effective in building trust between alliance partners. Interestingly, Pesch and Bouncken found that socializing practices are effective in building trust only when firm managerial CQ is high but can backfire when managerial CQ is low.

In the context of exporting firms, Golgeci and colleagues (2017) found that a leader's metacognitive CQ accentuates the positive relationship between firm absorptive capacity and firm innovativeness. This is because leaders with high metacognitive CQ are more likely to be adept in recognizing opportunities and exploring new knowledge in the external environment.

Another interesting context for the study of firm-level CQ is small and medium businesses. For instance, Charoensukmongkol (2015) examined the relationship between owners' CQ and firm performance using 129 small and medium manufacturing firms in Thailand. Results showed that business owners' CQ is positively related to firm export performance, mediated by owners' strong relationships with foreign stakeholders such as customers, competitors, and suppliers.

Future Directions

Our review summarizes the evolution of CQ research in the past decade, from establishing the incremental predictive validity of CQ (over and above related constructs such as personality and other forms of intelligences) to a more complex understanding of the mediators and moderators of CQ; from a focus on the individual-level to more sophisticated multilevel theorizing; from examining why CQ is important to how we can develop CQ. Many of these research developments are aligned with the research directions outlined in our review in the first edition of this handbook. Given the accumulating empirical studies on CQ, the time may be ripe for a meta-analysis to synthesize research findings. We also encourage future research to continue to deepen our understanding of the conceptualization, measurement, and nomological network of CQ. In the following section, we propose new directions for future research along these lines of inquiry.

Expand Conceptualizations of CQ

Biological loci of CQ. Beyond the four-factor model of CQ, a new and exciting development is the exploration of a biological loci of CQ. Building on recent advances in sociocognitive neuroscience research, Rockstuhl and colleagues (2010) elucidated the neurological basis of CQ by mapping the four CQ factors to distinct cortical regions in the brain. Specifically, they identified the anterior rostral medial frontal cortex (including the paracingulate cortex) as the neurological mediator for metacognitive CQ; the orbitofrontal cortex for motivational CQ; and the posterior rostral medial frontal cortex and dorsal anterior cingulate cortex for behavioral CQ. Beyond identifying the neurological mediators of the CQ factors, Rockstuhl and colleagues proposed that high overall CQ may be associated with a greater capability to tune one's patterns of neural activity to different cultural contexts. The conceptualization of the biological loci of CQ offers fertile ground for future research and empirical testing.

CQ and broader diversity markers. The existing conceptualization of CQ has focused on national cultural differences because they represent highly salient markers of cultural diversity (Ang, Van Dyne, & Rockstuhl, 2015). Future research could broaden the existing conceptualization to focus on other diversity markers, such as functional, generational, gender, and socioeconomic status diversity. One approach is to repeat the construct development and validation process, starting with developing grounded theories based on qualitative research on individuals who manage such diversity challenges effectively.

Diversify Measurement of CQ

Future research should continue to develop complementary measures of CQ. For example, big data and data science present many exciting opportunities to diversify the measurement of CQ. A key advantage of big data is that it offers data granularity, which allows us to examine the "most theoretically proximal measurement of a phenomenon or unit of analysis" (George et al., 2016, p. 1494). A specific example is in the use of wearable devices to monitor interaction patterns of individuals in culturally diverse settings and to measure "sociometrics" (e.g., tone of voice, gestures, turn-taking, interruption) (Pentland, 2012). This technology not only enables an unobtrusive and real-time assessment of behavioral CQ but also allows us to ask new questions. For instance, we can begin to ask questions on the proximal processes through which culturally intelligent leaders influence their followers; or how members in multicultural teams interact with (or avoid) one another and how they develop shared values over time; or how ideas are exchanged and fused between culturally diverse individuals to spark new and creative ideas.

Broaden the Nomological Network of CQ

Future research should also continue to broaden our understanding of the nomological network of CQ. Here, we suggest several areas that seem promising.

Antecedents and outcomes of CQ. Our review highlighted several emerging constructs within the nomological network of CQ that future research could shed more light on. On the antecedents of CQ, we urge future research to explore the interplay of language and CQ. According to the Sapir-Whorf hypothesis, language and culture are inextricably linked (Kramsch, 2014). It is therefore surprising that language has not been studied as much in CQ research (for exceptions, see Peyrols-Wu & Ng, 2018; Presbitero, 2017). For instance, does language capability compensate for low CQ (and vice versa) during intercultural interactions? In a study that explores the interactive effects of English-language self-efficacy and CQ on avoidance behaviors, Peyrols-Wu and Ng (2018) found that both language self-efficacy and CQ are important to minimize avoidance behaviors, suggesting that language capabilities and CQ are not compensatory. More research is needed to replicate and extend this initial finding to a broader range of outcomes.

On the outcomes of CQ, we encourage future CQ research to examine understudied dimensions of performance such as organizational citizenship behaviors, voice behaviors, and adaptive performance. For example, scholars may examine the effects of CQ on different types of voice behaviors such as supportive, constructive, defensive, and destructive voice (Maynes & Podsakoff, 2014) in culturally diverse settings. Future research could also deepen our existing understanding of CQ effects on creativity and negotiation performance by examining proximal mediators of these relationships.

CQ actions and CQ reactions. To date, CQ research has typically focused on one "actor" – a focal person and how their CQ will affect the effectiveness of the interactions. Consider this faux pas: An American enters a Japanese friend's home without removing his shoes. In this scenario, CQ research would have much to prescribe on how the American could avoid committing such a blunder by increasing his CQ capabilities. What is neglected, though, is the role of the CQ of the "partner." In this example, the Japanese host is equally important in ensuring the effectiveness of the interaction. Research shows that violations of social norms can evoke aversive emotional reactions (Berthoz et al., 2002) and negative evaluations of the offender (Molinsky, 2005). This brings up the interesting question of what constitutes a culturally intelligent reaction to a cultural faux pas. Should the Japanese host simply ignore the incident? Should he educate the American about the appropriate etiquette? And what might be the outcomes of either course of action? Are some responses more effective than others and, if so, under what conditions? In addressing these questions, we advance CQ research toward the truism that "it takes two to tango."

Factor-specific CQ training interventions. Another important future research direction relates to interventions for developing CQ in individuals. Whereas research to date has focused primarily on the development of overall CQ, future research could explore how different training interventions might differentially impact the development of specific CQ factors. For example, mindfulness interventions hold great potential in the development of metacognitive CQ (Allen et al., 2012; Thomas, 2006). Understanding interventions that develop specific CQ factors would increase the precision and effectiveness of CQ developmental methodologies, as well as deepen our understanding of the CQ development process. In addition to interventions targeted at developing individuals' CQ, future research could also explore interventions targeted at developing team-level CQ.

On another note, most research on CQ interventions such as study-abroad programs involved participants who have self-selected into the programs. As a result, findings from existing studies are based on samples where motivational CQ is already fairly high. Yet Marin (1990) pointed out that apathy, which refers to a lack of motivation, is perhaps a more common reaction in the face of cultural differences. Thus, we urge future research to examine CQ development in people who are in a state of cultural apathy.

Culturally intelligent team processes. Although scholars have begun to examine outcomes of CQ at the team level, culturally intelligent team processes and norms have received far less attention. One such team process that requires future research relates to the management of intercultural conflict. Intercultural conflict is particularly complex as the disagreement is amplified by cultural differences in the perception of conflict and what constitutes effective conflict management. We suggest that future CQ research adopts a grounded theory approach to uncover specific conflict management moves. Moves are subunits of analysis that compose a broader concept (Goffman, 1981) and have been used in prior research to study complex behavioral units (Clark et al., 2018). Discovering effective conflict management moves in intercultural conflicts will contribute greatly to our understanding of effective multicultural team functioning.

Culturally intelligent organizational routines. To date, empirical research on firm-level CQ has focused primarily on managerial CQ. Conceptual research on organizational CQ offers many promising ideas that have yet to be empirically validated, particularly in areas related to culturally intelligent organizational routines and practices (Ang & Inkpen, 2008; Moon, 2010; Ng, Tan, & Ang, 2011). Future research will offer great theoretical and practical contributions by assessing culturally intelligent organizational routines and their effects on firm performance.

CQ in the education system. Given the incessant pace of globalization, it is imperative that we educate and equip our future generations in CQ to manage growing cultural challenges surrounding issues of race, class, gender, politics, and religion. Goh (2012) argued that, to prepare citizens for the challenges of a global society, educational systems will have to incorporate CQ into their curricula. As he aptly pointed out, "the conversation in teacher education has shifted from *whether* we should train teachers to be culturally intelligent to *how*" (Goh, 2012, p. 412). We concur with this important need and urge future research to examine a systematic infusion of CQ into school curricula that would reach every student, beyond short-term interventions that target only a selective segment of the student population (e.g., study-abroad programs). Such broader attempts to infuse CQ into broader citizen education programs appear all the more important in light of the increasing populist and xenophobic movements witnessed worldwide in recent years.

Artificial cultural intelligence. Rapid advances in the areas of affective computing and artificial intelligence (AI) have given rise to conversational agents, such as chatbots. As Krakovsky (2018) rightly pointed out, the true success of these digital assistants lies in their ability to detect and respond to human emotions with "some emotional *savoir faire*" (p. 18), which is the foray of artificial emotional intelligence. Building on this point, we suggest that the true success of such digital assistants lies in their ability to detect and respond to human emotions *across cultures*. We believe that integrating affective computing and CQ will contribute greatly to the development of smart robots or chatbots that can cross cultures. For example, most dialogue systems still follow handcrafted rules for responding to human interaction (Schuller, 2018). A recent study by Morris, Savani, and Fincher (2019) on the role of metacognition highlights the potential of using alternative processes in cultural learning such as error monitoring and reactive error-based updating. We believe that integrating CQ work into the emerging field of affective computing could advance both fields and shape the future of how humans and machines interact.

Conclusions

CQ was conceived at the turn of the twenty-first century, when the world was experiencing unprecedented global trade, while, at the same time, witnessing the growth of ideological clashes and cultural conflicts. Almost two decades later, the relevance of CQ has remained, if not grown, as evidenced by the widespread interest in the construct in a vast array of disciplines. As summarized in this review, we have gained much insight and understanding into the nature, function, and boundary conditions of CQ at different levels of analysis. More importantly, we hope to spur future research to explore new conceptualizations, measurements, and applications of CQ. Ultimately, we hope that the continuing research on CQ will bring about real-world impact in this age of globalization.

References

Adair, W. L., Buchan, N. R., Chen, X. P., & Liu, D. (2016). A model of communication context and measure of context dependence. *Academy of Management Discoveries*, *2*, 198–217.

Adair, W. L., Hideg, I., & Spence, J. R. (2013). The culturally intelligent team: The impact of team cultural intelligence and cultural heterogeneity on team shared values. *Journal of Cross-Cultural Psychology*, *44*, 941–962.

Al-Dossary, S. A. (2016). Psychometric properties of the cultural intelligence scale in a Saudi Arabian context. *International Journal of Selection and Assessment*, *24*, 305–311.

Alexandra, V. (2018). Predicting CQ development in the context of experiential cross-cultural training: The role of social dominance orientation and the propensity to change stereotypes. *Academy of Management Learning and Education*, *17*, 62–78.

Allen, M., Dietz, M., Blair, K. S., van Beek, M., Rees, G., Vestergaard-Poulsen, P., . . . & Roepstorff, A. (2012). Cognitive-affective neural plasticity following active-controlled mindfulness intervention. *Journal of Neuroscience*, *32*, 15601–15610.

Allport, G. (1954). *The nature of prejudice*. Reading, MA: Addison Wesley.

Ang, S., & Inkpen, A. C. (2008). Cultural intelligence and offshore outsourcing success: A framework of firm-level intercultural capability. *Decision Sciences*, *39*, 337–358.

Ang, S., Rockstuhl, T., & Ng, K. Y. (2014). *Performance-based cultural intelligence (CQ): Development and validation of an intercultural situational judgment test (iSJT).* Singapore: Nanyang Technological University, Center for Leadership and Cultural Intelligence.

Ang, S., & Van Dyne, L. (2008). Conceptualization of cultural intelligence: Definition, distinctiveness, and nomological network. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence* (pp. 3–15). New York: M.E. Sharpe.

Ang, S., Van Dyne, L., & Koh, C. (2006). Personality correlates of the four-factor model of cultural intelligence. *Group and Organization Management*, *31*, 100–123.

Ang, S., Van Dyne, L., Koh, C., Ng, K.Y., Templer, K. J., Tay, C., & Chandrasekar, N.A. (2007). Cultural intelligence: Its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance. *Management and Organization Review*, *3*, 335–371.

Ang, S., Van Dyne, L., & Rockstuhl, T. (2015). Cultural intelligence: Origins, conceptualization, evolution and methodological diversity. In M. Gelfand, C. Y. Chiu, & Y. Y. Hong (Eds.), *Handbook of Advances in Culture and Psychology* (Vol. 5, pp. 273–323). New York: Oxford University Press.

Berthoz, S., Armony, J. L., Blair, R. J. R., & Dolan, R. J. (2002). An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain*, *125*(8), 1696–1708.

Brannen, M. Y., & Thomas, D. C. (2010). Bicultural individuals in organizations: Implications and opportunity. *International Journal of Cross-Cultural Management*, *10*, 5–16.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 687–732). Palo Alto, CA: Consulting Psychologists Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Chao, M. M., Takeuchi, R., & Farh, J. L. (2017). Enhancing cultural intelligence: The roles of implicit culture beliefs and adjustment. *Personnel Psychology*, *70*, 257–292.

Charoensukmongkol, P. (2015). Cultural intelligence of entrepreneurs and international network ties: The case of small and medium manufacturing firms in Thailand. *Management Research Review*, *38*, 421–436.

Charoensukmongkol, P. (2016). Cultural intelligence and export performance of small and medium enterprises in Thailand: Mediating roles of organizational capabilities. *International Small Business Journal*, *34*, 105–122.

Chen, A. S. (2015). CQ at work and the impact of intercultural training: An empirical test among foreign laborers. *International Journal of Intercultural Relations*, *47*, 101–112.

Chen, G., Kirkman, B. L., Kim, K., Farh, C. I., & Tangirala, S. (2010). When does intercultural motivation enhance expatriate effectiveness? A multilevel investigation of the moderating roles of subsidiary support and cultural distance. *Academy of Management Journal*, *53*, 1110–1130.

Chen, M. L., & Lin, C. P. (2013). Assessing the effects of cultural intelligence on team knowledge sharing from a socio-cognitive perspective. *Human Resource Management*, *52*, 675–695.

Chen, X. P., Liu, D., & Portnoy, R. (2012). A multilevel investigation of motivational cultural intelligence, organizational diversity climate, and cultural sales: Evidence from U.S. real estate firms. *Journal of Applied Psychology*, *97*, 93–106.

Chua, R. Y. J., Morris, M. W., & Mor, S. (2012). Collaborating across cultures: Cultural metacognition and affect-based trust in creative collaboration. *Organizational Behavior and Human Decision Processes*, *118*, 116–131.

Chua, R. Y., & Ng, K. Y. (2017). Not just how much you know: Interactional effect of cultural knowledge and metacognition on creativity in a global context. *Management and Organization Review*, *13*, 281–300.

Church, A. (1982). Sojourner adjustment. *Psychological Bulletin*, *91*, 540–572.

Clark, C., Tan, M. L., Murfett, U. M., Rogers, P., & Ang, S. (2018). The call center agent's performance paradox. A mixed-methods study of discourse strategies and paradox resolution. *Academy of Management Discoveries* (Advance online publication.) https://doi.org/10.5465/amd.2016.0024

Crotty, S. K., & Brett, J. M. (2012). Fusing creativity: Cultural metacognition and teamwork in multicultural teams. *Negotiation and Conflict Management Research*, *5*, 210–234.

Crowne, K. A. (2008). What leads to cultural intelligence? *Business Horizons*, *51*, 391–399.

Crowne, K. A., & Engle, R. L. (2016). Antecedents of cross-cultural adaptation stress in short-term international assignments. *Organization Management Journal*, *13*, 32–47.

Dheer, R. J., & Lenartowicz, T. (2018). Multiculturalism and entrepreneurial intentions: Understanding the mediating role of cognitions. *Entrepreneurship Theory and Practice*, *42*, 426–466.

Earley, P. C., & Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures*. Palo Alto, CA: Stanford University Press.

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review*, *13*, 73–92.

Engle, R. L., & Crowne, K. A. (2014). The impact of international experience on cultural intelligence: An application of contact theory in a structured short-term programme. *Human Resource Development International*, *17*(1), 30–46.

Erez, M., Lisak, A., Harush, R., Glikson, E., Nouri, R., & Shokef, E. (2013). Going global: Developing management students' cultural intelligence and global identity in culturally diverse virtual teams. *Academy of Management Learning and Education*, *12*, 330–355.

Firth, B. M., Chen, G., Kirkman, B. L., & Kim, K. (2014). Newcomers abroad: Expatriate adaptation during early phases of international assignments. *Academy of Management Journal*, *57*, 280–300.

Gelfand, M. J., Imai, L., & Fehr, R. (2008). Thinking intelligently about cultural intelligence: The road ahead. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence* (pp. 375–387). New York: M.E. Sharpe.

George, G., Osinga, E. C., Lavie, D., & Scott, B. A. (2016). Big data and data science methods for management research. *Academy of Management Journal*, *59*, 1493–1507.

Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.

Goh, M. (2012). Teaching with cultural intelligence: Developing multiculturally educated and globally engaged citizens. *Asia Pacific Journal of Education*, *32*, 395–415.

Gölgeci, I., Swiatowiec-Szczepanska, J., & Raczkowski, K. (2017). How does cultural intelligence influence the relationships between potential and realised absorptive capacity and innovativeness? Evidence from Poland. *Technology Analysis and Strategic Management*, *29*, 857–871.

Groves, K. S., & Feyerherm, A. E. (2011). Leader cultural intelligence in context: Testing the moderating effects of team cultural diversity on leader and team performance. *Group and Organization Management*, *36*, 535–566.

Groves, K. S., Feyerherm, A., & Gu, M. (2015). Examining cultural intelligence and cross-cultural negotiation effectiveness. *Journal of Management Education*, *39*, 209–243.

Gudykunst, W. B. (1993). Toward a theory of effective interpersonal and intergroup communication. In R. J. Wiseman & J. Koester (Eds.), *Intercultural communication competence, international and intercultural communication annual* (Vol. 16, pp. 3–17). Newbury Park, CA: Sage.

Guðmundsdóttir, S. (2015). Nordic expatriates in the US: The relationship between cultural intelligence and adjustment. *International Journal of Intercultural Relations*, *47*, 175–186.

Hall, E. T. (1959). *The silent language*. New York: Anchor Books.

Hesketh, B., & Neal, A. (1999). Technology and performance. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of performance: Implications for staffing, motivation, and development* (pp. 21–55). San Francisco: Jossey-Bass.

Huff, K. C., Song, P., & Gresch, E. B. (2014). Cultural intelligence, personality and cross-cultural adjustment: A study of expatriates. *International Journal of Intercultural Relations*, *38*, 151–157.

Imai, L., & Gelfand, M. J. (2010). The culturally intelligent negotiator: The impact of cultural intelligence (CQ) on negotiation sequences and outcomes. *Organizational Behavior and Human Decision Processes*, *112*, 83–98.

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology*, *95*, 54–78.

Jyoti, J., & Kour, S. (2015). Assessing the cultural intelligence and task performance equation: Mediating role of cultural adjustment. *Cross Cultural Management*, *22*, 236–258.

Kanfer, R., & Heggestad, E. D. (1997). Motivational traits and skills: A person-centered approach to work motivation. *Research in Organizational Behavior*, *19*, 1–56.

Kim, Y. J., & Van Dyne, L. (2012). Cultural intelligence and international leadership potential: The importance of contact for members of the majority. *Applied Psychology: An International Review*, *61*, 272–294.

Klafehn, J., Li, C., & Chiu, C. Y. (2013). To know or not to know, is that the question? Exploring the role and assessment of metacognition in cross-cultural contexts. *Journal of Cross-Cultural Psychology*, *44*, 963–991.

Korzilius, H., Bücker, J. J., & Beerlage, S. (2017). Multiculturalism and innovative work behavior: The mediating role of cultural intelligence. *International Journal of Intercultural Relations*, *56*, 13–24.

Krakovsky, M. (2018). Artificial (emotional) intelligence. *Communications of the ACM*, *61*, 18–19.

Kramsch, C. (2014). Language and Culture. *AILA Review*, *27*, 30–55.

Lance, C. E., Hoffman, B. J., Gentry, W. A., & Baranik, L. E. (2008). Rater source factors represent important subcomponents of the criterion construct space, not rater bias. *Human Resource Management Review*, *18*, 223–232.

Le, H., Jiang, Z., & Nielsen, I. (2018). Cognitive cultural intelligence and life satisfaction of migrant workers: The roles of career engagement and social injustice. *Social Indicators Research*, *139*, 237–257.

Lee, Y.-T., Masuda, A., Fu, X., & Reiche, S. (2018). Navigating between home, host, and global: Consequences of multicultural team members' identity configurations. *Academy of Management Discoveries*, *4*, 180–201.

Leung, K., Ang, S., & Tan, M. L. (2014). Intercultural competence. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 489–519.

Li, M., Mobley, W. H., & Kelly, A. (2013). When do global leaders learn best to develop cultural intelligence? An investigation of the moderating role of experiential learning style. *Academy of Management Learning and Education*, *12*, 32–50.

Li, Y., Rau, P. P., Li, H., & Maedche, A. (2017). Effects of a dyad's cultural intelligence on global virtual collaboration. *IEEE Transactions on Professional Communication*, *60*, 56–75.

Lin, Y.-C., Chen, A., S.-Y., & Song, Y.-C. (2012). Does your intelligence help to survive in a foreign jungle? The effects of cultural intelligence and emotional intelligence on cross-cultural adjustment. *International Journal of Intercultural Relations*, *36*, 541–552.

Lisak, A., & Erez, M. (2015). Leadership emergence in multicultural teams: The power of global characteristics. *Journal of World Business*, *50*, 3–14.

Lorenz, M. P., Ramsey, J. R., & Richey, R. G. (2018). Expatriates' international opportunity recognition and innovativeness: The role of metacognitive and cognitive cultural intelligence. *Journal of World Business*, *53*, 222–236.

MacNab, B. R. (2012). An experiential approach to cultural intelligence education. *Journal of Management Education*, *36*, 66–94.

MacNab, B., Brislin, R., & Worthley, R. (2012). Experiential cultural intelligence development: Context and individual attributes. *The International Journal of Human Resource Management*, *23*, 1320–1341.

Magnusson, P., Schuster, A., & Taras, V. (2014). A process-based explanation of the psychic distance paradox: Evidence from global virtual teams. *Management International Review*, *54*, 283–306.

Magnusson, P., Westjohn, S. A., Semenov, A. V., Randrianasolo, A. A., & Zdravkovic S. (2013). The role of cultural intelligence in marketing adaptation and export performance. *Journal of International Marketing*, *21*(4), 44–61.

Marin, R. S. (1990). Differential diagnosis and classification of apathy. *The American Journal of Psychiatry*, *147*, 22–30.

Matsumoto, D., & Hwang, H. C. (2013). Assessing cross-cultural competence: A review of available tests. *Journal of Cross-Cultural Psychology*, *44*, 849–873.

Maynes, T. D., & Podsakoff, P. M. (2014). Speaking more broadly: An examination of the nature, antecedents, and consequences of an expanded set of employee voice behaviors. *Journal of Applied Psychology*, *99*, 87–112.

Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, *17*, 433–442.

Molinsky, A. (2005). Language fluency and the evaluation of cultural faux pas: Russians interviewing for jobs in the United States. *Social Psychology Quarterly*, *68*, 103–120.

Moon, H. K., Choi, B. K., & Jung, J. S. (2012). Previous international experience, intercultural training, and expatriates' intercultural adjustment: Effects of cultural intelligence and goal orientation. *Human Resource Development Quarterly*, *23*, 285–330.

Moon, T. (2010a). Emotional intelligence correlates of the four-factor model of cultural intelligence. *Journal of Managerial Psychology*, *25*, 876–898.

Moon, T. (2010b). Organizational cultural intelligence: Dynamic capability perspective. *Group and Organization Management*, *35*, 456–493.

Morris, M. W., Savani, K., & Fincher, K. (2019). Metacognition fosters cultural learning: Evidence from individual differences and situational prompts. *Journal of Personality and Social Psychology*, *116*, 46–68.

Moynihan, L. M., Peterson, R. S., & Earley, P. C. (2006). Cultural intelligence and the multinational team experience: Does the experience of working in a multinational team improve cultural intelligence? *Research on Managing Groups and Teams*, *9*, 299–323.

Ng, K. Y., Tan, M. L., & Ang, S. (2011). Global culture capital and cosmopolitan human capital: The effects of global mindset and organizational routines on cultural intelligence and international experience. In A. Burton & J. C. Spender (Eds.), *The Oxford handbook of human capital* (pp. 96–119), New York: Oxford University Press.

Oolders, T., Chernyshenko, O. S., & Stark S. (2008). Cultural intelligence as a mediator of relationships between openness to experience and adaptive performance. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence* (pp. 145–158). New York: M.E. Sharpe.

Peng, A. C., Van Dyne, L., & Oh, K. (2015). The influence of motivational cultural intelligence on cultural effectiveness based on study abroad: The moderating role of participant's cultural identity. *Journal of Management Education*, *39*, 572–596.

Pentland, A. (2012). The new science of building great teams. *Harvard Business Review*, *90* (4), 60–69.

Pesch, R., & Bouncken, R. B. (2017). How to achieve benefits from diversity in international alliances: Mechanisms and cultural intelligence. *Global Strategy Journal*, *8*, 275–300.

Peyrols-Wu, C., & Ng, K. Y. (2018). Avoidance in intercultural work contexts: Predictors and mediators. Paper presented at the 78th Annual Meeting of the Academy of Management, Chicago, USA, August 10–14.

Pless, N. M., Maak, T., & Stahl, G. K. (2011). Developing responsible global leaders through international service-learning programs: The Ulysses experience. *Academy of Management Learning and Education*, *10*, 237–260.

Presbitero, A. (2016). Cultural intelligence (CQ) in virtual, cross-cultural interactions: Generalizability of measure and links to personality dimensions and task performance. *International Journal of Intercultural Relations*, *50*, 29–38.

Presbitero, A. (2017). It's not all about language ability: Motivational cultural intelligence matters in call center performance. *The International Journal of Human Resource Management*, *28*, 1547–1562.

Racicot, B. M., & Ferry, D. L. (2016). The impact of motivational and metacognitive cultural intelligence on the study abroad experience. *Journal of Educational Issues*, *2*, 115–129.

Ramirez R, E. (2016). Impact on intercultural competence when studying abroad and the moderating role of personality. *Journal of Teaching in International Business*, *27* (2–3), 88–105.

Rockstuhl T., & Ng, K. Y. (2008). The effects of cultural intelligence on interpersonal trust in multicultural teams. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence* (pp. 206–220). New York: M.E. Sharpe.

Rockstuhl, T., Ang, S., Ng, K. Y., Lievens, F., & Van Dyne, L. (2015). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *Journal of Applied Psychology*, *100*, 464–480.

Rockstuhl, T., Hong, Y.-Y., Ng, K. Y., Ang, S., & Chiu, C.-Y. (2010). The culturally intelligent brain: From detecting to bridging cultural differences. *Journal of Neuroleadership*, *3*, 1–15.

Rockstuhl, T., Seiler, S., Ang, S., Van Dyne, L., & Annen, H. (2011). Beyond general intelligence (IQ) and emotional intelligence (EQ): The role of cultural intelligence (CQ) on cross-border leadership effectiveness in a globalized world. *Journal of Social Issues*, *67*, 825–840.

Rosenblatt, V., Worthley, R., & MacNab, B. (2013). From contact to development in experiential cultural intelligence education: The mediating influence of expectancy disconfirmation. *Academy of Management Learning and Education*, *12*, 356–379.

Şahin, F., & Gürbüz, S. (2014). Cultural intelligence as a predictor of individuals' adaptive performance: A study in a multicultural environment. *International Area Studies Review*, *17*, 394–413.

Şahin, F., Gürbüz, S., & Köksal, O. (2014). Cultural intelligence (CQ) in action: The effects of personality and international assignment on the development of CQ. *International Journal of Intercultural Relations*, *39*, 152–163.

Şahin, F., Gürbüz, S., Köksal, O., & Ercan, Ü. (2013). Measuring cultural intelligence in the Turkish context. *International Journal of Selection and Assessment*, *21*, 135–144.

Salmon, E. D., Gelfand, M. J., Çelik, A. B., Kraus, S., Wilkenfeld, J., & Inman, M. (2013). Cultural contingencies of mediation: Effectiveness of mediator styles in intercultural disputes. *Journal of Organizational Behavior*, *34*, 887–909.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262–274.

Schuller, B. W. (2018). Speech emotion recognition: Two decades in a nutshell, benchmarks, and ongoing trends. *Communications of the ACM*, *61*(5), 90–99.

Shannon L. M., & Begley, T. M. (2008). Antecedents of four-factor model of cultural intelligence. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence* (pp. 41–55). New York: M.E. Sharpe.

Shokef, E., & Erez, M. (2008). Cultural intelligence and global identity in multicultural teams. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence* (pp. 177–191). New York: M.E. Sharpe.

Shu, F., McAbee, S. T., & Ayman, R. (2017). The HEXACO personality traits, cultural intelligence, and international student adjustment. *Personality and Individual Differences*, *106*, 21–25.

Sri Ramalu, S., Shamsudin, F. M., & Subramaniam, C. (2012). The mediating effect of cultural intelligence on the relationship between openness personality and job performance among expatriates on international assignments. *International Business Management*, *6*, 601–610.

Sternberg, R. J. (1986). A framework for understanding conceptions of intelligence. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition* (pp. 3–15). Norwood, NJ: Ablex.

Sternberg, R. J. (1997). *Successful intelligence: How practical and creative intelligence determine success in life*. New York: Plume.

Sternberg, R. J., & Wagner, R. J. (2000). Practical intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 380–395). New York: Cambridge University Press.

Taras, V., Caprar, D. V., Rottig, D., Sarala, R. M., Zakaria, N., Zhao, F., . . . & Bryła, P. (2013). A global classroom? Evaluating the effectiveness of global virtual collaboration as a teaching tool in management education. *Academy of Management Learning and Education*, *12*, 414–435.

Tarique, I., & Takeuchi, R. (2008). Developing cultural intelligence: The roles of international nonwork experiences. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence: Theory, measurement, and applications* (pp. 56–70). New York: M. E. Sharpe.

Tay, C., Westman, M., & Chia, A. (2008). Antecedents and consequences of cultural intelligence among short-term business travelers. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural Intelligence: Theory, measurement, and applications* (pp. 126–144). New York: M. E. Sharpe.

Thomas, D. C. (2006). Domain and development of cultural intelligence: The importance of mindfulness. *Group and Organization Management*, *31*, 78–99.

Thorndike, R. L., & Stein, S. (1937). An evaluation of the attempts to measure social intelligence. *Psychological Bulletin*, *34*, 275–285.

Van Dyne, L., Ang, S., & Koh, C. (2008). Development and validation of the CQS. In S. Ang & L. Van Dyne (Eds.), *Handbook of cultural intelligence* (pp. 16–38). New York: M. E. Sharpe.

Van Dyne, L., Ang, S., & Livermore, D. (2010). Cultural intelligence: A pathway for leading in a rapidly globalizing world. In K. M. Hannum, B. McFeeters, & L. Booysen (Eds.), *Leading across differences* (pp. 131–138). San Francisco: Pfeiffer.

Van Dyne, L., Ang, S., Ng, K. Y., Rockstuhl, T., Tan, M. L., & Koh, C. (2012). Subdimensions of the four factor model of cultural intelligence: Expanding the conceptualization and measurement of cultural intelligence. *Social and Personality Psychology Compass*, *6*, 295–313.

Varela, O. E., & Gatlin-Watts, R. (2014). The development of the global manager: An empirical study on the role of academic international sojourns. *Academy of Management Learning and Education*, *13*, 187–207.

Wang, K. T., Heppner, P. P., Wang, L., & Zhu, F. (2015). Cultural intelligence trajectories in new international students: Implications for the development of cross-cultural competence. *International Perspectives in Psychology: Research, Practice, Consultation*, *4*, 51–65.

Ward, C., Fischer, R., Lam, F. S. Z., & Hall, L. (2009). The convergent, discriminant, and incremental validity of scores on a self-report measure of cultural intelligence. *Educational and Psychological Measurement*, *69*, 85–105.

Wood, E. D., & St. Peters, H. Y. (2014). Short-term cross-cultural study tours: Impact on cultural intelligence. *The International Journal of Human Resource Management*, *25*, 558–570.

Xu, X.-J., & Chen, X.-P. (2017). Unlocking expatriates' job creativity: The role of cultural learning, and metacognitive and motivational cultural intelligence. *Management and Organization Review*, *13*, 767–794.

Yitmen, I. (2013). Organizational cultural intelligence: A competitive capability for strategic alliances in the international construction industry. *Project Management Journal*, *44* (4), 5–25.

Zhang, Y., & Oczkowski, E. (2016). Exploring the potential effects of expatriate adjustment direction. *Cross Cultural and Strategic Management*, *23*, 158–183.

35 Mating Intelligence

Glenn Geher, Scott Barry Kaufman, Julie A. Planke, and Jacqueline M. Di Santo

Mating Intelligence Defined

In the broadest terms, we see *mating intelligence* as the cognitive abilities that bear on mating-relevant outcomes – in short: the mind's reproductive system (Geher, Miller, & Murphy, 2008). *Mating intelligence* differs from the broader field of *mating psychology* per se as mating intelligence focuses on relatively high-level cognitive processes – intelligence that underlies the domain of human mating – while mating psychology writ large has focused on relatively basic, unconscious, low-level psychological processes – such as the effects of ovulation on attraction (Miller, Tybur, & Jordan, 2007) or the nature of the human voice as a courtship device (e.g., Locke, 2017; Pipitone & Gallup, 2008). A mountain of research on human mating makes it abundantly clear that many basic psychological processes comprise evolved mating adaptations in our species.

Mating intelligence is different in that it focuses on the richer, more abstract, and more intellectual nature of human psychology in the domain of mating. Clearly, there are low-level, physiological, and emotional aspects of human mating that seem like important products of our evolutionary heritage. Mating intelligence suggests that there are also high-level, cognitive aspects of human psychology that also primarily reflect mating-relevant adaptations resulting from our evolutionary heritage.

Summary of Geher, Camargo, and O'Rourke's (2008) Model

In summarizing the first fifteen chapters of the book *Mating Intelligence*, Geher, Camargo, and O'Rourke (2008) provide a framework for conceptualizing this new construct. First, these authors draw important distinctions between the *fitness indicator* component of mating intelligence and the *cognitive mating mechanisms* component.

Rooted in Miller's (2000a) conception of high-order human intelligence as having evolved for courtship purposes, the fitness indicator component of mating intelligence corresponds to areas of intelligence that are uniquely human (including, for instance, artistic and linguistic elements), that vary dramatically from person to person, that are partly heritable, and that are attractive in the mating domain. Such forms of intelligence may include, for instance, art (Nettle & Clegg, 2006), creative writing (Nettle, 2009), humor (see Kaufman et al., 2010; Kaufman et al., 2008), and vocabulary (see Rosenberg & Tunney, 2008). Importantly, while these hypothesized mental fitness indicators have been shown to act as courtship signals, they do not necessarily directly bear on mating issues. Thus, the fitness indicator component of mating intelligence is thought to comprise higher-order intellectual processes (e.g., the ability to write and recite a high-quality poem) but the links between these processes and mating outcomes are conceptualized as indirect. So, while poetic ability, for instance, may have evolved partly because success in this area was related to success in attracting high-quality mates, the thoughts that underlie poetry need not be directly mating-relevant or, indeed, publicly advertised as part of courtship (Nettle, 2009; although they may be – see Gottschall & Wilson, 2005).

On the other hand, *cognitive mating mechanisms* are proposed to be relatively high-level cognitive abilities that bear directly on mating-relevant issues. In successful mating, one must effectively engage in a host of such processes – such as accurate cross-sex mind reading (to know whether a potential mate is interested, to know what a current mate wants, etc.), strategic flexibility in mating strategies (knowing when it is optimal to pursue long-term versus short-term strategies), being able to read cues that reliably indicate that a mate has cheated in a relationship, being able to outcompete intrasexual rivals while keeping an eye toward presenting oneself as kind and other-oriented, and so on. In short, there are many cognitive processes that are directly relevant to the domain of mating. We conceptualize these processes as the cognitive mating mechanisms of mating intelligence.

Two important superordinate variables underlie the nature of mating intelligence in the model proposed by Geher, Camargo, and O'Rourke (2008). The first is biological sex. In many regards, human mating processes have been shown to be sex-differentiated. While dramatic intrasex variability tends to exist for mating-relevant variables, consistent sex differences on such variables are reliably found – often across disparate cultures – suggesting that males are more likely than females to pursue short-term mating strategies across the gamut of mating-relevant behavioral traits (see Buss, 2003). As such, male mating intelligence is predicted to be more honed toward optimizing short-term mating opportunities while female mating intelligence is predicted to be more honed toward optimizing long-term opportunities. This prediction follows from asymmetries in parental investment across the sexes that benefit males, the lower investing sex, in short-term strategies and that benefit females, the higher investing sex, in long-term strategies (see Buss, 2003).

Consistent with Buss's (2003) predictions, using the newly developed Mating Success Scale (MSS), Camargo and colleagues (2013) found that their MSS short-term subscale positively correlated with the number of short-term partners for males and their MSS long-term subscale positively correlated with the number of long-term partners for females. These findings provide empirical support for sex differences in optimizing short- and long-term mating opportunities when it comes to mating success.

Life-history strategy is a similarly important superordinate variable (see Figueredo et al., 2008). This idea, adapted from evolutionary ecology, suggests that organisms unconsciously strategize to find an optimal balance between somatic effort (facilitating their own survival) and reproductive effort (facilitating the replication of their genes into future generations via reproduction). This concept was initially designed by biologists to characterize different kinds of species – those that are *k-selected* – defined as "expecting" a long life within a stable environment (e.g., elephants) versus those that are *r-selected* – defined as "expecting" an unpredictable life, within an unstable environment (e.g., rabbits; see MacArthur & Wilson, 1967). While humans are k-selected as a species, there are clearly differences among human environments in terms of predictability of resources and long-term stability. With this idea in mind, Figueredo and colleagues (2008) and others (see Giosan, 2006) propose that people differ in terms of the degree to which they follow a prototypical k-selected strategy. As such, these scholars conceptualize a k-differential continuum as typifying humans, with some people being relatively *high-k* (these would be individuals who are raised in relatively resource-rich and stable backgrounds) and others being relatively *low-k* (individuals raised in harsh and relatively resource-poor and unstable backgrounds and/or high in mortality).

A great deal of recent research has shown that the differential-k continuum is strongly predictive of general behavioral strategies – with *high-k* individuals being more likely to delay gratification and take long-term approaches to solving problems (mating and otherwise) and *low-k* individuals being more likely to seek instant gratification and to take short-term approaches to solving problems (see Kruger, Reischl, & Zimmerman, 2008). Such a strategies approach allows for plasticity and malleability of human adaptations and is in concordance with evolutionary principles in behavioral ecology, in which adaptations vary by specific environmental demands, as these constraints influence the expression of adaptations (Wilson, 2007). Consequently, this plasticity of adaptations also allows for considerable individual differences, the focus of the mating intelligence construct.

Geher, Camargo, and O'Rourke (2008) propose that the differential-k continuum is a major variable that underlies mating intelligence. To the extent that the elements of mating intelligence are adaptations, designed to facilitate long-term reproductive success, it makes sense that the nature of mating intelligence would change as a function of an individual's placement on the differential-k continuum. High-k individuals are expected to be most likely to pursue long-term mating strategies and to ultimately engage in high levels of parental effort while low-k individuals are expected to be most likely to pursue short-term mating opportunities. As such, high-k individuals are predicted to have cognitive sets that facilitate long-term mating, often at a cost to success in the area of short-term mating, while low-k individuals are predicted to be characterized by cognitive sets that, on the other hand, facilitate success in short-term mating. Thus, the nature of mating intelligence likely takes on different forms in light of the k-differential continuum. Someone high in general intelligence who comes from an unstable childhood background and develops a low-k life-history strategy may well make mating decisions in adulthood that seem highly unintelligent (consider Bill Clinton's scandal with Monica Lewinsky, as an example).

Finally, Geher, Camargo, and O'Rourke (2008) propose that the different elements of mating intelligence – including the fitness indicators and cognitive mating mechanisms – ultimately should predict Darwin's bottom line of reproductive success. In fact, from an evolutionary perspective, all adaptations are adaptations because they gave our ancestors reproductive advantages. Biologists who study nonhumans are able to see whether certain traits are more likely to lead to higher numbers of viable offspring compared with other traits. However, the study of humans from an evolutionary perspective runs into an idiosyncratic quagmire regarding this issue: birth control. The presence of birth control in most Westernized societies makes it nearly impossible to study contemporary human behavior optimally from an evolutionary perspective, as hypothesized human evolutionary adaptations cannot typically be examined vis-à-vis reproductive success. A researcher who, for instance, hypothesizes that relatively deep voices in males evolved because women are attracted to such voices and ultimately are more willing to become pregnant and bear children of men with deep voices runs into a problem – such women may well be taking oral contraceptives – so this researcher will have a difficult time counting viable offspring as a way of testing their adaptationist hypothesis.

This problem, which ends up as a major concern for all evolutionary approaches to humans, needs to be addressed. Geher, Camargo, and O'Rourke (2008) and others (e.g., Pérusse, 1993) propose that we need to measure indicators of mating success as a proxy for reproductive success to be better able to test evolutionary hypotheses. If mating intelligence does comprise an important set of adaptations, then measures of mating intelligence should predict reproductive success. In this sense, we refer to *reproductive* success purely in the Darwinian sense, corresponding to the number of viable offspring produced in one's lifetime (and other markers of inclusive fitness).

Since we cannot typically measure reproductive success effectively in large samples of modern humans, predictions regarding mating intelligence should seek to predict mating success that may be addressed in terms of behavioral outcomes as well as potential reproductive fitness outcomes such as sperm quality (see Arden et al., 2009). Interestingly, Arden and colleagues (2009) and others have found that sperm quality actually tracks markers of physical attractiveness and health.

Mating success is defined largely as including outcomes that would have likely led to reproductive success under precontraceptive conditions. For males, such outcomes would include, for a straightforward example, having had sexual intercourse with multiple women (of high mate value, including having strong physical health, etc.) and, in particular, attracting women who are physically attractive. Of course, males also are often motivated to pursue long-term strategies (see Simpson & Gangestad, 2000) and, as such, a measure of mating success for males should also include such outcomes as being courted by kind, intelligent, and socially connected females for long-term relationships. For females, outcomes associated with mating success would include, for instance, having a history of dating relatively successful men (i.e., men who are high in objective measures of mate value) and having had multiple men allocate resources toward them (for a thorough treatment of operationalizing mating success in modern humans, see Camargo et al., 2013). Importantly, mating success, in this context, refers to outcomes that would have led to increased fitness relative to same-sex competitors under ancestral conditions – we are not referring to more intuitive conceptions, such as relationship happiness in long-term mateships. The working assumption here, which is consistent with the general approach of evolutionary psychology (see Geher, 2014), is that our minds are the products of ancestral conditions and, as such, we are focusing on factors that would have increased reproductive success of our ancestors as opposed to modern proximate outcomes associated with relationships.

In sum, this model of mating intelligence suggests that it (1) is broken into fitness indicators and cognitive mating mechanisms, (2) is moderated importantly by the superordinate variables of biological sex and the differential-k continuum, and (3) ultimately predicts mating success.

What Is New Here?

What is new here? Any time someone proposes a novel psychological construct, educated psychological researchers automatically raise a skeptical eye – rightfully so. The modern behavioral sciences are rooted in methods for objectively collecting and analyzing observable data. Psychology is an empirical science – and psychologists demand empirical evidence for any and all claims. While this skeptical approach may make psychological research difficult to conduct and to publish, it is, generally, a good thing. The scientifically rigorous approach that underlies modern research psychology makes it so that the material taught to students in psychology classes in modern universities is based largely on data.

When Geher and Miller (2008) launched the construct of mating intelligence in their book by the same name, they knew full well that this construct would be under a good bit of scrutiny. In fact, several of the chapters in that edited volume on mating intelligence included comments that were critical of the concept writ large. Never one to mince words, Satoshi Kanazawa (2008) wrote: “Intelligence, in its original definition, referred to purely cognitive abilities . . . I personally would have preferred to keep it that way” (p. 283). Similar concerns are expressed in chapters by Figueredo and colleagues (2008) as well as in David Buss’s (2008) foreword to the book.

While the basic idea of mating intelligence has generally been well received in both academic (see Springer, 2009) and popular circles (see Perina, 2007), we think it is important to address criticisms of this construct up front. As is true of any newly introduced psychological construct, the main criticism launched at mating intelligence has been essentially this: What is new here?

The Heuristic Value of Mating Intelligence (What Is New Here)

Sometimes, progressive scientific ideas form from stepping back and looking at things from a new angle (see Dawkins, 2005). We believe that the unification of the fields of mating and intelligence, implied in the mating intelligence construct, provides such a new angle on many areas of the behavioral sciences. In a thorough consideration of the areas potentially illuminated by this construct, Miller (2008) argues that mating intelligence has the potential to improve our understanding of such disparate facets of human functioning as medicine, psychiatry, economics, marketing, political science, sociology, education, and law. Here, we discuss specific areas of psychological research that may benefit – or that have already benefited – from the mating intelligence construct.

The study of individual differences from an evolutionary perspective has been, to this point, largely incomplete. With a major focus on human universals, evolutionary psychology has often either dismissed or ignored individual differences in important behavioral traits. While there are some important exceptions to this generalization, such as Nettle and Clegg's (2008) work on understanding superordinate personality trait dimensions (such as introversion/extraversion) in terms of balancing selection forces and Simpson and Gangestad's (1991) groundbreaking work on individual differences in sociosexuality (i.e., promiscuity; see also Penke, Denissen, & Miller, 2007), by and large, mating research conducted from an evolutionary perspective focuses on human universals such as sex-specific tactics to derogate competitors for mates (e.g., Buss & Schmitt, 1996), universals in the nature of human jealousy (Buss et al., 1992), universals in features of attractive faces and bodies (Hughes & Gallup, 2003), and universals in qualities desired in long-term versus short-term mates (Gangestad & Simpson, 2000).

While the universalist approach that characterizes most evolutionary psychology research clearly has shed light on many important aspects of the human condition, it fails to do justice to the myriad traits in our species that demonstrate reliable individual differences. Our conception of mating intelligence as including both mental fitness indicators and cognitive mating mechanisms opens the door for two important areas of individual-differences research. The study of mental fitness indicators addresses many cognitively laden traits (see Klasios, 2013) that seem to act as courtship mechanisms. Such traits include verbal fluency (Rosenberg & Tunney, 2008), humor (Greengross & Miller, 2008, 2011; Kaufman et al., 2008; Tornquist & Chiappe, 2015), conspicuous altruism (in a general sense – not just toward a specific mate; see Arnocky et al., 2017; Miller, 2007), and creative writing (Kaufman & Kaufman, 2009; Lange & Euler, 2014).

The study of cognitive mating mechanisms has the potential to provide insights into many areas of mating psychology that have been primarily studied from a universalist perspective. For instance, while mating psychologists have previously documented sex-specific features of deception in the mating domain (e.g., Haselton et al., 2005), a mating intelligence approach to this issue may address individual differences in mate-deception efficacy (e.g., O'Brien et al., 2010). Similarly, while prior researchers have addressed universals in responses to infidelity, it may be that there are individual differences in such processes as (1) the ability to accurately detect infidelity, (2) the ability to engage in infidelity with a high-quality mate, (3) the ability to deceive a partner about one's history of infidelity, and so forth. The study of individual differences in mating-relevant trait dimensions should be a major product of the mating intelligence construct.

In formulating our model of mating intelligence (Geher, Camargo, & O'Rourke, 2008), the importance of mating success became clear. Intelligence research of all kinds focuses on predicting success in some area. Research on cognitive, or general intelligence, has focused on predicting success in various academic arenas (see Sternberg, 1996); research on social intelligence has sought to pinpoint the predictors of success in such areas as marriage and career (Cantor & Kihlstrom, 1987); research on emotional intelligence has examined the predictors of success in such areas as intimate relationships (Casey et al., 2008), health (Matthews, Zeidner, & Roberts, 2002), education (Brackett et al., 2007), and, most recently, creativity (Geher, Betancourt, & Jewell, 2017). Given the evolutionary roots that underlie mating intelligence, it quickly becomes clear that the main kind of success that should result from mating intelligence would be *reproductive success*, which is essentially Darwin's bottom line – ultimately bearing on the number of viable descendants that reach future generations (taking physical quality of descendants into account [conceptually] as well, to the extent that quality facilitates gene proliferation overall across generations). Whether a trait is adaptive in the Darwinian sense corresponds, ultimately, to whether certain levels of that trait led to increases in reproductive success in our ancestors. As such, the main outcome that should be predicted by any adaptation is reproductive success – often framed in terms of the number of viable offspring produced. To further elucidate the primacy of Darwin's bottom line, individuals higher in mating intelligence show particular preferences for sex acts that would facilitate reproduction, that is, vaginal intercourse (Peterson, Geher, & Kaufman, 2011).

An important hurdle to the study of mating intelligence, then, becomes apparent. Given the widespread use of birth control in so many modern societies, reproductive success, operationalized in terms of number of offspring, has little construct validity. A tall, muscular, symmetrical, dominant, and intelligent male in a modern society may well attract many high-quality (attractive, healthy, and free from debilitating mental illness) sexual partners but his consistent use of birth control may reduce his reproductive success to zero. As such, the widespread use of birth control renders reproductive success nearly impossible to operationalize in modern human populations. For this reason, we propose that reproductive success needs to be approximated with measures of mating success – defined (based on our prior work in this area) as including outcomes that would have corresponded to reproductive success under precontraceptive conditions (see Geher, Camargo, & O'Rourke, 2008). While previous scholars have considered the importance of operationalizing mating success (e.g., Pérusse, 1993), the mating intelligence framework makes the need for valid measures of mating success extremely clear.

Camargo and colleagues (2013) sought to create a valid and reliable MSS to serve as a proxy for reproductive success for research within the mating domain. Importantly, this scale focused on the quality of sexual relationships rather than simply the number of previous sexual partners. Accurately measuring the mating success of males compared to females may therefore be a relatively simpler feat, with the metric of the number of sex partners alone sufficiently serving to capture the construct. One of the important outcomes of the mating intelligence construct should pertain to thorough psychometric work on mating success.

Intelligence and Mental Fitness Indicators

Human courtship has a distinct flavor compared to the courtship behavior of other species. We sing tunes designed to coordinate with lyrics, write poems, and paint wonderfully complex and aesthetic pictures to attract mates. We go on dates, exchange witty banter, and engage in long conversations about preferences and values. Why do we bother?

When we seek a mate, we surely look for someone whom we can connect with on a personal level, who shares our hopes, desires, goals, and fears. As such, mate selection in humans consistently focuses on qualities that are optimal for short-term as well as long-term partners. But, at another level, our genes pull us toward individuals high in fitness (heritable genetic quality). Most animals in the animal kingdom advertise fitness by displaying elaborate structures that do not appear to serve a survival function. The peacock's tail, the elk's antlers, and the nightingale's voice are all examples of adaptations that signal fitness.

Humans are unique, however, in the amount of fitness information that is contained in the brain. And, because the brain is the source of human intelligence, intelligence is fair game for sexual selection. According to the principle of sexual selection, reproduction is just as much a struggle as survival. Thus, while adaptations for survival surely come to typify organisms via evolutionary processes, adaptations that are primarily about successful reproduction share the front seat. Sexually selected traits (as opposed to traits operating under the forces of natural selection) display high variance because there is competition for individuals to mate with those who exhibit traits that are metabolically expensive, hard to maintain, not easily counterfeited, and highly sensitive to genetic mutations. Such traits that display these properties are the most reliable indicators of genetic fitness. According to Zahavi's (1997) handicap principle, even though fitness indicators may impair the odds of survival (creating a handicap), they can offer reproductive benefits that outweigh the survival costs. The peacock's tail may make it difficult for the peacock to walk, and may make the peacock more visible to predators, but it attracts mates. Likewise with the human brain – while there may be metabolic costs associated with having such a heavy brain, the costs may be outweighed by reproductive benefits (as well as, perhaps, benefits in other domains). Those animals who can display such structures that go beyond survival are advertising that they not only have the resources to survive but they also have resources left over to invest in excess. An analogy can be found in Veblen's (1899) idea of conspicuous consumption. According to Veblen, a wasteful display of wealth is a reliable indicator of wealth since the poor cannot afford such waste. From Zahavi's perspective, such characteristics represent *costly signals*, which evolve as hard-to-fake, honest advertisements of heritable qualities. Importantly, while modern technologies have changed the playing field a bit in terms of how people present themselves to one another, many basic behavioral attributes, such as verbal fluidity, artistic talent, and so on, still fit within this costly signaling framework.

In recent years, Geoffrey Miller has applied Zahavi's handicap principle to the evolution of human intelligence, arguing that sexual selection played a much greater role than natural selection in shaping the most distinctively human aspects of our minds, including storytelling, art, music, sports, dance, humor, kindness, and leadership (Miller, 1998, 2000a, 2000b, 2000c, 2001; Kaufman et al., 2008). Miller argues that these behaviors are the result of complex psychological adaptations whose primary functions were to attract mates, yielding reproductive rather than survival benefits. Germs of this idea can be traced back to Darwin: "It appears probable that the progenitors of man, either the males or females or both sexes, before acquiring the power of expressing mutual love in articulate language, endeavored to charm each other with musical notes and rhythm" (Darwin, 1871, p. 880).

In his model of the mind as a courtship device, Miller (2000) makes several assumptions regarding the nature of human intelligence. In this model, he assumes that the general factor of human intelligence (i.e., *g*) is synonymous with human intelligence writ large. Miller also argues that behaviors that show a strong influence of general intelligence (i.e., are highly *g*-loaded) should be sexually attractive since they are indicators of a superordinate *fitness factor* (*f* factor). Indeed, evidence has been accumulating that suggests the existence of an *f* factor. Various threads of research show a correlation between *g* and many biological traits such as height, health, longevity, bodily symmetry, and even sperm quality (Arden et al., 2009; Banks, Batchelor, & McDaniel, 2010; Bates, 2007; Calvin et al., 2010; Furlow et al., 1997; Jensen, 1998; Prokosch, Yeo, & Miller, 2005; Silventoinen et al., 2006; Sundet et al., 2005); *g* may therefore be an indicator of deleterious mutation load, which would affect many interacting genes and thereby have an effect on the entire biological system.

There is also accumulating evidence that intelligence and creativity (which Miller argues are an important facet of intelligence) are sexually attractive traits. Buss (1989) investigated mate preferences across thirty-seven cultures and found that intelligence was the second-most-desired trait in a sexual partner, right below kindness. Experimental research shows that intelligent and creative individuals are considered more attractive and have a higher number of sexual partners (Buss, 1989; Griskevicius, Cialdini, & Kenrick, 2006; Haselton & Miller, 2006; Nettle & Clegg, 2006; Prokosch et al., 2009).

Various scholars have elaborated and clarified Miller's theory. Feist (2001) notes that Miller focuses on sexual selection and artistic creativity to the exclusion of the evolution of scientific creativity and technology, which Feist argues is more likely to have been shaped by natural selection pressures. Further, Feist (2001) argues that natural selection has driven mainly the more applied or technological aspects of creativity that have clear survival benefits, such as advances in science and engineering, whereas sexual selection may have driven more ornamental or aesthetic aspects of creativity, including art, music, dance, and humor – forms of creativity that have come along more recently on the evolutionary scene.

Therefore, not all creative displays may be considered equally as sexually attractive. "Nerdier" displays of creativity, such as in math, engineering, and the sciences, may be considered less attractive, on average, than more "artistic" displays of creativity such as in poetry, music, and art. Recent research does suggest that, collapsing over individual differences, more artistic forms of creativity are considered more sexually attractive than are more scientific forms of creativity (Kaufman et al., 2009). This trend, of course, may well depend on the target audience (as female scientists might find scientific creativity to be relatively attractive).

In any case, individual differences were found in that those who reported higher levels of creative achievement in scientific forms of creativity did tend to find scientific forms of creativity sexually attractive (as well as some artistic forms of creativity), whereas those who reported higher levels of creative achievement in artistic forms of creativity did tend to find artistic forms of creativity sexually attractive but did not report finding scientific forms of creativity sexually attractive. Future research should clarify these issues, testing Feist's hypothesis at both the group and the individual level of analysis.

In a related line of thought, Feist argues that Miller's account of sexual selection does not fully connect with the creativity literature. In this body of literature, creativity is defined as both novel and adaptive behavior (Sternberg, 1998), not as novel creative displays that attract the attention of potential mates. Feist also notes that there is evidence that creative people tend to be less likely to marry and, when they do, they have relatively *few* children (Harrison, Moore, & Rucker, 1985), a factor that surely also impacts on reproductive success. Also, it should be noted that time spent on creative projects may be time taken away from mating and child-rearing (Gabora & Kaufman, 2010). And it is also possible that creative individuals may have trouble in relationships, on average, as well, due to their unique constellation of personality traits, including being less conventional and conscientious and more driven, ambitious, dominant, hostile, and impulsive than less creative individuals (see Feist, 1998). However, additional research has suggested that, since negative dispositional qualities (e.g., anxiety, neuroticism, schizotypy) associated with mental illness sometimes go hand in hand with creativity, these attributes may actually be beneficial in regard to short-term mating in males (Beaussart, Kaufman, & Kaufman, 2012).

In a related line of research, Mithen (2006) presents evidence that the musicality of our ancestors and relatives may in fact have had considerable *survival value* as a means of communicating emotions, intentions, and information as well as facilitating cooperation. Thus, sexual selection may not be the primary selective pressure for musicality. He also notes that, while it may appear at first blush that creative men have more short-term sexual partners because of their creativity, their attractiveness may be more the combination of good looks, style, and an antiestablishment persona. Mithen also points out that the finding (Miller, 1999) that males produce at least ten times more music than females and are the most productive around the age of thirty (in which men are in their peak mating effort and activity) could more parsimoniously be explained by the particular structure and attitudes of twentieth-century Western society. Perhaps the most reasonable conclusion is that sexual selection helped ramp up the evolution of intelligence and creativity, exaggerating certain forms, or making them not only functional but also ornamental. In this way, they went beyond the realm of practicality to the realm of aesthetic functionality. However, more recent research provides support for Miller's general hypothesis that artistic ability is at least partially sexually selected. Crocchiola (2014) found both male and female artists, compared to nonartist controls, have higher prenatal testosterone levels, as indicated by having a relatively low ratio between the second to fourth digit finger length.

From a different angle, Kanazawa (2008) argues that individuals with greater general intelligence do not have greater mating intelligence, except in areas where the mechanisms underlying mating intelligence operate on evolutionarily novel stimuli. Kanazawa (2004, 2010) proposed that general intelligence evolved as a domain-specific psychological mechanism to solve evolutionarily novel problems (for a different perspective on the evolution of general intelligence, see Borsboom & Dolan, 2006; Chiappe & MacDonald, 2005; Geary, 2004, 2009; Girotto & Tentori, 2008; Kaufman et al., 2011). With this theory as a foundation, Kanazawa (2008) argues that general intelligence is independent of other adaptations, including mating intelligence. Kanazawa presents evidence that those higher in verbal intelligence are relatively ineffective at evolutionarily familiar tasks such as finding mates, having children, and getting and staying married (for further evidence on the negative association between IQ and marriage, see Taylor et al., 2005). Kanazawa presents evidence that those with higher verbal intelligence are better, however, at voluntarily controlling fertility, a finding Kanazawa interprets as reflecting the better ability of those with higher verbal intelligence in dealing with evolutionarily novel means of contraception in the current environment. Accordingly, Kanazawa and others see this tendency for individuals high in general intelligence to take steps to inhibit reproduction as consistent with the dysgenic hypothesis, that low intelligence drives out high intelligence.

Perhaps it is important to distinguish between the sexual attractiveness of intelligence and the use of human intelligence to navigate the mating domain. An interesting irony may be that, while intelligence might be a sexually attractive trait, those with high intelligence may have no advantage in actually navigating the mating domain (unless the domain consists of evolutionary novelty). It is to the cognitive mechanisms underlying mating intelligence that we now turn.

Mating-Relevant Cognitive Mechanisms

As stated in prior work, we believe that the cognitive mating mechanisms of mating intelligence include both species-typical and individual-differences features (Miller, 2008). Species-typical (i.e., universalist) mating mechanisms include the many mating qualities that have been studied by prior researchers that may be thought of as characterizing a human universal mating intelligence. Such qualities include, as examples, the tendencies to (1) advertise qualities that are attractive to potential mates (Buss & Schmitt, 1996), (2) engage in adaptive mating-relevant self-deception (O'Sullivan, 2008; Lynn, Pipitone, & Keenan, 2014), (3) demonstrate meta-strategic flexibility, by changing one's mating strategy as a function of current ecological conditions (such as the prevailing sex ratios; see Schmitt 2005), and (4) hold biased mating-relevant beliefs that may be evolutionarily adaptive (Haselton &

Buss, 2000; Cyrus, Schwarz, & Hassebrauck, 2011). To a large extent, the edifice of mating psychology comprises the species-typical portion of mating intelligence's cognitive mating mechanisms.

As a recent example of a mating-relevant psychological process framed as a cognitive mating mechanism, consider Geher's (2009) work on cross-sex mind reading. Rooted in methods borrowed from the field of emotional intelligence research (Geher, 2004), this work explored, in a large sample of heterosexual adults, the ability to accurately guess the mating desires of the opposite sex. Being able to read the thoughts of the opposite sex (literally, not in an extrasensory manner!) comprises an important set of cognitive skills that are crucial for mating success. Thus, this ability is a crucial cognitive mating mechanism that underlies mating intelligence. In this research, participants were presented with real personal ads written by members of their own sex – and they were asked to judge which ad (in clusters of three) was rated as most attractive for either a long-term or short-term mating partner by members of the opposite sex. In a separate part of the study, members of the opposite sex rated these same ads, so the actual answers could be determined. Ads were all content-coded for the presence of sexual content in a blind process by two independent judges.

Across both short- and long-term items, women showed a strong tendency to overestimate the degree to which males were attracted to ads of women who included sexual content. These findings are consistent with an *adaptive bias* account of cross-sex mind reading, suggesting that women may be particularly prone to think that men are only interested in sex; such a judgment may encourage women to be especially skeptical of men's intentions. Such commitment skepticism may be part of a broad long-term female mating strategy designed to reduce the likelihood of a female's being impregnated by a nonfaithful male and, thus, bearing the evolutionary tax of raising an offspring alone.

In terms of accuracy in cross-sex mind reading, the findings were revealing. Each sex turned out to be relatively expert at guessing the mating-relevant thoughts of the opposite sex when the judgments corresponded to the dominant strategy of the opposite sex. Thus, females outperformed males in guessing short-term desires, while males outperformed females in guessing long-term desires. Accordingly, it seems that cross-sex mind reading seems particularly honed when it comes to knowing what the opposite sex wants in the areas that are prioritized by the opposite sex. The Analogical Peacock Hypothesis, proposed by McKeown (2013), further suggests that mind-reading abilities (in addition to creative displays) are the product of sexual selection pressures that are advantageous in social communication and, in turn, can increase reproductive success. This hypothesis basically says that conspicuous social-analytic abilities may be the products of sexual selection.

While Geher (2009) explicates the utility of the mating intelligence construct to generate new research and new findings, this study was limited when it came to understanding cross-sex mind reading in terms of individual differences. An attempt to measure cross-sex mind reading in terms of individual differences did not yield internally reliable scales. While this fact was somewhat disappointing, it is worth noting that this same issue typified the earliest attempts to create ability-based measures of emotional intelligence (Mayer & Geher, 1996). Attempts to operationalize emotional intelligence in terms of individual differences have increased markedly in their success across time (Geher, 2004). We expect that attempts to measure the mating mechanisms of mating intelligence as individual-differences variables will also succeed in time.

In fact, another thread of recent work has demonstrated that mating intelligence may prove to be a valid individual-differences construct. Geher and Kaufman (2007) created a self-report measure of mating intelligence published in *Psychology Today* (Perina, 2007). While this scale was not initially designed with scholarly goals in mind, several recent studies that have included this measure have demonstrated its internal reliability as well as its predictive utility (O'Brien et al., 2010). Male and female versions of this scale, created primarily for use with heterosexual populations, tap several major dimensions that underlie mating intelligence, including (1) accuracy in cross-sex mind reading, (2) effective deception in the mating domain (a characteristic that likely pertains to effective self-presentation in both short-term and long-term mating strategies), (3) adaptive self-deception in the mating domain, (4) adaptive mating-relevant bias (with the male subscale corresponding to overestimating the degree to which women find males sexually attractive and the female subscale corresponding to being hyper-skeptical of males' intentions), and (5) effective behavioral courtship display. Thus, this scale is designed to tap both mental fitness indicators and mating mechanisms in terms of individual differences.

It is important to note that this measure uses self-report methods and that, without question, work on this scale represents the nascent stage of psychometric efforts on this construct that are needed. Previous research on aspects of human intelligence using self-report methods has generally cast a critical eye on such approaches (Geher & Renstrom, 2004). Ultimately, ability-based measures would likely have more face validity as well as, perhaps, more predictive validity. Still, both the male and the female versions of this measure (based on total scale scores) demonstrated high internal-consistency reliability. Further, in two studies on young heterosexual adults, this scale demonstrated a strong ability to predict important variables related to reproductive success. In the first study, males' scores were positively predictive of having had more sexual partners in the past year as well as more lifetime partners, whereas females' scores showed a more nuanced pattern, with high mating intelligence for females corresponding to having had sexual relations relatively early in life but not having a relatively high number of sexual partners in the last year (a pattern that may be adaptive for females). Thus, for males, high mating intelligence seems to correspond to more sexual partners overall whereas, for females, high mating intelligence corresponds to having more sexual experience but *not* a more promiscuous current strategy (O'Brien et al., 2010).

A second study explored mating intelligence in the context of hookups, generally defined as short-term sexual relationships with no explicit long-term relationship attached (Garcia & Reiber, 2008; for review, see Garcia et al., 2012). In addition to measuring mating intelligence, this study asked participants if they had ever engaged in Type-I hookups (with strangers), Type-II hookups (with acquaintances), and Type-III hookups (with individuals they defined as *friends*). Again, the mating intelligence scale demonstrated sensitivity to important sex-differentiated features of relationships. For males, higher mating intelligence corresponded to having engaged in each kind of hookup, whereas, for females, high mating intelligence corresponded to having engaged in hookups with acquaintances (Type-II) but not either of the other kinds. These findings make sense from an evolutionary perspective, as it may be particularly costly for a female to engage in sex with a stranger, about whom she has little information. Such relationships, started with minimal baseline information, could put a female at high risk for such adverse outcomes as violence, desertion, or disease. On the other hand, prior research has demonstrated that it is not adaptive for females to have sexual relations with close opposite-sex friends; and, in fact, females typically do not report having opposite-sex friends for sexual reasons (Bleske-Recheck & Buss, 2001). Relations with individuals defined as acquaintances may well strike a balance.

The findings from the aforementioned studies (Geher, 2009; O'Brien et al., 2010) are presented to give a face to the field of mating intelligence. Some of these findings bear primarily on species-typical mating mechanisms whereas others focus on individual differences in the different elements of mating intelligence. While this work provides an important first step in carving out the nature of mating intelligence and its contribution to the field of psychology, more research is surely needed to help the mating intelligence construct realize its potential.

The Future of Mating Intelligence

By proposing the mating intelligence construct, we hope to stimulate research on the connection between human sexuality and human intelligence. A large part of the relatively nascent field of evolutionary psychology includes the study of human mating (see Buss, 2005). However, evolutionary psychology has traditionally focused on human universals instead of individual differences. Recent work has even suggested there may be individual variation in mating intelligence over the course of the ovulatory cycle in naturally cycling females, with mating intelligence scores increasing with probability of conception (Peterson, Carmen, & Geher, 2013). The field has also traditionally focused on lower-level cognitive processes instead of higher-level cognitive functions. We hope the mating intelligence construct will provide a missing piece of the human cognitive puzzle for the fields of both human intelligence and evolutionary psychology and will stimulate cross-talk between the two fields of inquiry.

The integrative model of mating intelligence outlined here and first proposed by Geher, Camargo, and O'Rourke (2008) includes two main components. The first class of cognitive processes relates to mating-relevant cognitive domains that are thought to primarily serve courtship display functions. While evolutionary psychology has tended to focus mainly on behavioral displays of physical qualities such as strength, virility, and athleticism, the mating intelligence construct focuses on psychological qualities (*mental fitness indicators*) such as confidence, kindness, creativity, intelligence, resourcefulness, status, humor, and mental health.

According to Miller's (2000) fitness indicator model, humans are particularly attuned to behavioral qualities of potential mates that reveal *good genes* in the evolutionary sense in that they reveal a relatively low mutation load (in other words, a relatively low number of genetic mutations) as well as genes that are generally associated with health, survival, and successful reproductive abilities (see Keller & Miller, 2006). Therefore, much of human mate choice can be explained as an adaptive (unconscious) fear of heritable mutations – as *mutation phobia*. According to this idea of mutation phobia, people are repulsed by features of potential mates that have a strong latent correlation with high mutation load. Importantly, this model is largely based on the idea that various physical and behavioral attributes have significant heritability coefficients (see Miller, 2000).

In the biological literature, body asymmetry or dullness of plumage are often given as examples (see Hasson, 2006). Although physical traits such as bodily asymmetry are thought to be indicators of developmental instability (DI), which negatively relates to fitness, a recent study failed to identify significant relationships between various indicators of DI and mating success (Kordsmeye & Penke, 2017). However, these null findings may be the result of an inadequate assessment of mating success. Additional work would perhaps benefit from the implementation of the newly developed MSS (Camargo et al., 2013) to capture both quantitative and qualitative aspects of mating success and is necessary to fully claim that low-DI individuals (with low mutation loads) are actually more attractive, as indicated by their mating success. Clearly, future research on this topic is needed.

It is not clear, however, whether such mate choice operates in a continuous or categorical manner. It is entirely possible that our mate preferences have been shaped more to avoid mating with high mutation load individuals who have obvious physical or psychological problems than to make very fine discriminations among individuals who seem more or less average in terms of mutation load. The idea here is that, although not all physical and psychological problems are rooted in high mutation load, across evolutionary time, there was enough of a correlation between these variables that we were selected to prioritize these attributes in the mate selection process.

Zebrowitz and Rhodes (2004) offer evidence that, at least in some cases, mate choice operates in a categorical manner. They found that people could accurately predict overall health and intelligence for targets with relatively unattractive faces but not for targets with relatively attractive faces. Facial attractiveness was predictive of health and intelligence for targets and intelligence only at the low-fitness extremes.

Such a curvilinear relationship between indicator quality and sexual attractiveness (concave-downward, with rapidly diminishing returns above the mean of indicator quality) may be seen in the domain of mating intelligence. For example, someone with an IQ of 90 may be much more attractive than someone with an IQ of 70 but a potential mate with an IQ of 150 may only be a little more attractive than one whose IQ is 130. Importantly, IQ scores vary considerably across such groupings as education levels. As such, these points regarding IQ likely need to be considered in such a broader context.

Research should attempt to investigate the (probably nonlinear) functions that relate mutation load to mental fitness indicators and that relate indicator quality to attractiveness in mating. Such research should sample populations from all strata of society. Indeed, if it turns out that fitness indicators correlate differently at low-quality and high-quality extremes, and assortative mating on IQ is a predominant occurrence, then bright, healthy, college sophomores may not be the best and/or only population we should be studying for mating intelligence research on the display, judgment, and sexual attractiveness of fitness indicators!

Another issue in the understanding of mental fitness indicators has to do with the relation of each fitness indicator to general intelligence. In conceiving of *g*-loaded mental traits as having arisen from sexual selection processes, Miller (2000a) posits that *g* is essentially an index of neurodevelopmental stability and brain efficiency that taps any overall fitness factor (roughly, the first principal component of genetic quality across all fitness-related traits). Further, Miller proposes that the existence of this superordinate fitness factor should be manifest as a positive manifold (all-positive correlations) among fitness indicators in general. Future research should attempt to tests Miller's (2000c) predictions and shed light on the nature of the courtship display components of mating intelligence. One such method would be to simply assess the *g*-loadings of a variety of mental fitness indicators and compare the relationship of the *g*-loadings to ratings of sexual attractiveness of each fitness indicator. According to Miller, there should be a positive relationship.

Future research should also try to elucidate the particular characteristics of various mental displays that are sexually attractive. Various forms of creativity (e.g., artistic) may be considered more attractive than other forms of creativity (e.g., scientific), not only due to indications of *g* (indeed, scientific forms of creativity are probably more *g*-loaded than artistic forms of creativity) but also due to fitness indications of kindness, emotional expressivity, and so on. Importantly, these different attributes are not all strongly intercorrelated with one another and many exceptions to the general trends that are postulated in Miller's (2000) model exist.

Future research should also assess the importance of individual differences in preferences for various mental fitness indicators. Preliminary research in this regard is underway (Kaufman et al., 2009) and suggests that, at the group level, artistic forms of creativity are considered more sexually attractive than scientific forms of creativity, with substantial individual differences in preferences for forms of creative display that can at least partly be predicted by an individual's personality, intelligence, and creativity.

The second class of cognitive processes acts as mating mechanisms. Such potentially fruitful domains of mating intelligence that can be classified under the mating mechanisms component of mating intelligence include *mate choice mechanisms* for evaluation and choosing among potential sexual partners (e.g., Penke et al., 2008); *self-evaluation mechanisms* for assessing one's own mate value (O'Brien et al., 2010); *mechanisms for making context-sensitive decisions about mating strategies* (Schmitt, 2005) such as whether to pursue short-term or long-term relationships; *cross-sex mind reading mechanisms* (Geher, 2009) for understanding and influencing the behavior of potential mates, and of their friends, families, and children; and *same-sex mind reading mechanisms* for understanding and influencing the behavior of potential sexual rivals, and of their friends, families, and allies (Fisher, 2004). Future research should also attempt to investigate relations between mental fitness indicators and mating mechanisms. For instance, are those with a higher IQ better able to detect interest in a potential mate? Are those who are higher in fitness displays such as humor production better at assessing their own mate value? Such an investigation of how various fitness indicators relate to one another and with other mating mechanisms will help clarify the structure of mating intelligence.

One step toward this clarification would be to develop a performance measure of mating intelligence. The mating mechanisms in our model may be interrelated much like the abilities that underlie emotional intelligence (see Chapter 29 in this volume). The ability-based model of emotional intelligence presented by the authors of that chapter suggests that there are four basic facets of emotional intelligence, which are somewhat interrelated and mildly *g*-loaded. These facets include the ability to identify emotions, assimilate emotion into thought, understand emotions, and manage emotions (in one's self and others). This framework might be useful for producing a test of mating intelligence as well as understanding the structure of mating intelligence. Just as emotional intelligence may have basic interrelated components that underlie it, mating intelligence may also have basic elements (such as the ability to accurately assess one's own mate value) that may be interrelated and found to comprise a distinct set of mating-relevant cognitive abilities. Kirsner, Figueredo, and Jacobs (2003) created the Mate Value Inventory (MVI), a comprehensive multivariate scale, which assessed qualities desired in social and sexual mates (e.g., intelligence, jealousy, physical fitness, desire to have children, and ambition). Recent research has assessed the relationship between mating intelligence and mate value specifically finding that mating intelligence appears to be a valid predictor of one's self-perceived mate value (Dillon et al., 2016). Moreover, Wagstaff, Sulikowski, and Burke (2015) attempt to further build the bridge between mating intelligence and mate value by showing how making mating-relevant decisions is correlated with one's own perceived mate value.

An important distinction between emotional intelligence and mating intelligence here pertains to content – with emotional intelligence dealing with emotion-relevant stimuli and processes and mating intelligence focusing on content tied to the mating domain. Recent research has shown that individuals engage in assortative mating on emotional intelligence (Śmieja & Stolarski, 2018). Essentially, people tend to mate with others who have a level of emotional intelligence comparable to their own (Dillon et al., 2016). Similar findings may be found regarding assortative mating patterns and mating intelligence specifically, once a more robust mating intelligence measure is established. An important area for future work would be to examine the empirical relationships between measures of emotional intelligence and measures of mating intelligence in detail.

In addition to such basic psychometric qualities as internal reliability of measuring instruments, this work will need to assess (1) whether different elements of mating intelligence are mildly interrelated, (2) whether they are somewhat related to *g*, (3) whether they are not redundant with well-established personality traits such as the Big Five, and (4) whether the abilities that comprise mating intelligence are, indeed, predictive of mating success (such as the abilities to attract, choose, court, and retain high-quality sexual partners and to deter sexual rivals and infidelities). Such psychometric work will be crucial in determining whether mating intelligence is a useful individual-differences construct within psychology writ large. Further, given that emotional intelligence is predictive of success in intimate relationships, research on the interface between emotional intelligence and mating intelligence could be both theoretically and practically valuable. Finally, future research needs to focus on measuring mating intelligence in an ability-based manner. Work on the parallel construct of emotional intelligence has clearly demonstrated that indices of this construct as an ability are not fully correlated with indices of this construct measured via self-report measures (see Geher, 2004). Ability-based measures of mating intelligence might use work in emotional intelligence as a guide, examining such abilities as, for instance, the ability to know what is attractive to a large group of potential mates, the ability to effectively deceive others regarding mating-relevant stimuli, and so on. Future research along these lines should be very fruitful in carving out the nature of this construct.

In terms of the practical value of mating intelligence, there are important potential applications of the mating intelligence framework to society. Awareness of mating intelligence in the larger society should increase our appreciation of psychological and mental qualities in a potential mate in addition to purely physical qualities. Further, sex education in schools can be improved by being informed by the mating intelligence framework. In particular, by embracing the fact that much of the human mind is really about mating, sex education classes could teach students the importance of mental indicators and the various skills necessary to successfully navigate the mating domain. Informed by the complexities of human mating research, such education could address the fact that there are multiple routes to success in mating – with males and females both armed with a variety of long- and short-term strategies that are highly context-sensitive. The mating intelligence idea underscores this complexity but also places these ideas within a coherent framework informed by evolutionary theory. In addition, Geher and Kaufman's (2013) book *Mating Intelligence Unleashed: The Role of the Mind in Sex, Dating, and Love* provides us with practical guidance regarding mating intelligence research for our daily lives by providing insight into questions in the field of mating by using research in areas such as personality, social psychology, developmental psychology, neuroscience, and creativity.

It is our hope that the mating intelligence construct, by providing an evolutionarily informed understanding of human intelligence that takes into account the important domain of human mating, can allow us to move toward a more complete understanding of human intelligence.

Acknowledgments

We greatly appreciate the diligent work of senior editor Robert Sternberg during the editorial process and we thank Justin Garcia for offering his thoughtful guidance and insights regarding biological aspects of this work. We also thank Megan Geher for her excellent editorial help.

References

Arden, R., Gottfredson, L. S., Miller, G., & Pierce, A. (2009). Intelligence and semen quality are positively correlated. *Intelligence*, *37*, 277–282.

Arnocky, S., Piché, T., Albert, G., Ouellette, D., & Barclay, P. (2017). Altruism predicts mating success in humans. *British Journal of Psychology*, *108*(2), 416–435. https://doi.org/10.1111/bjop.12208

Banks, G. C., Batchelor, J. H., & McDaniel, M. A. (2010). Smarter people are (a bit) more symmetrical: A meta-analysis of the relationship between intelligence and fluctuating asymmetry. *Intelligence*, *4*, 393–401.

Bates, T. C. (2007). Fluctuating asymmetry and intelligence. *Intelligence*, *35*, 41–46.

Beaussart, M. L., Kaufman, S. B., & Kaufman, J. C. (2012). Creative activity, personality, mental illness, and short-term mating success. *The Journal of Creative Behavior*, *46*, 151–167.

Bleske-Recheck, A. L., & Buss, D. M. (2001). Opposite-sex friendship: Sex differences and similarities in initiation, selection, and dissolution. *Personality and Social Psychology Bulletin*, *27*, 1310–1327.

Borsboom, D., & Dolan, C. V. (2006). Why g is not an adaptation: A comment on Kanazawa (2004). *Psychological Review*, *113*, 433–437.

Brackett, M. A., Alster, B., Wolfe, C. J., Katulak, N. A., & Fale, E. (2007). Creating an emotionally intelligent school district: A skill based approach In R. Bar-On, J. G. Maree, & M. J. Elias (Eds.), *Educating people to be emotionally intelligent* (pp. 123–137). Westport, CT: Praeger.

Buss, D. M. (1989). Sex differences in human mate selection: Evolutionary hypothesis tested in 37 cultures. *Behavioral and Brain Sciences*, *12*, 1–49.

Buss, D. M. (2003). *The evolution of desire: Strategies of human mating*. New York: Basic Books.

Buss, D. M. (2005). *The handbook of evolutionary psychology*. New York: Wiley.

Buss, D. M. (2008). The future of mating intelligence. *Mating intelligence: Sex, relationships, and the mind's reproductive system*. Mahwah, NJ: Erlbaum.

Buss, D. M., Larsen, R. J., Weston, D., & Semmelroth, J. (1992). Sex differences in jealousy: Evolution, physiology, and psychology. *Psychological Science*, *3*, 251–255.

Buss, D. M., & Schmitt, D. P. (1996). Strategic self-promotion and competition derogation: Sex and conflict effects on perceived effectiveness of mate attraction tactics. *Journal of Personality and Social Psychology*, *70*, 1185–1204.

Calvin, C. M., Deary, I. J., Fenton, C., Roberts, B. A., Der, G., Leckenby, N., & Batty, G. D. (2010). Intelligence in youth and all-cause-mortality: Systematic review with meta-analysis. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyq190.

Camargo, M. A., Geher, G., Fisher, M. L., & Arrabaca, A. (2013). The relationship between hypothesized psychological genetic fitness indicators and indices of mating success. *Journal of Social, Evolutionary, and Cultural Psychology*, *7*(2), 138–162. https://doi.org/10.1037/h0099205

Cantor, N., & Kihlstrom, J. F. (1987). *Personality and social intelligence*. Englewood Cliffs, NJ: Prentice-Hall.

Casey, J. J., Garrett, J., Brackett, M. A., & Rivers, S. (2008). Emotional intelligence, relationship quality, and partner selection. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system* (pp. 263–282). Mahwah, NJ: Erlbaum.

Chiappe, D., & MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *Journal of General Psychology*, *132*, 5–40.

Crocchiola D. (2014). Art as an indicator of male fitness: Does prenatal testosterone influence artistic ability? *Evolutionary Psychology*, *12*(3), 521–533.

Cyrus, K., Schwarz, S., & Hassebrauck, M. (2011). Systematic cognitive biases in courtship context: Women's commitment–skepticism as a life-history strategy? *Evolution and Human Behavior*, *32*(1), 13–20. https://doi.org/10.1016/j.evolhumbehav.2010.07.006

Darwin, C. (1871). *The descent of man, and selection in relation to sex* (2 vols.). London: John Murray.

Dawkins, R. (2005). Afterword. In D. M. Buss (Ed.), *The handbook of evolutionary psychology*. New York: Wiley.

Dillon, H. M., Adair, L. E., Geher, G., Wang, Z., & Strouts, P. H. (2016). Playing smart: The mating game and mating intelligence. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, *35*(3), 414–420. https://doi.org/10.1007/s12144-015-9309-y

Feist, G. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, *2*, 290–309.

Feist, G. (2001). Natural and sexual selection in the evolutionary of creativity. *Bulletin of Psychology and the Arts*, *2*, 11–16.

Figueredo, A. J., Brumbach, B. H., Jones, D. N., Sefcek, J. A., Vasquez, G., & Jacobs, W. J. (2008). Ecological constraints on mating tactics. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system* (pp. 335–361). Mahwah, NJ: Erlbaum.

Fisher, M. L. (2004). Female intrasexual competition decreases female facial attractiveness. *Proceedings of the Royal Society*, *271*, 283–285.

Furlow, B., Armijo-Prewitt, T., Gangestad, S. W., & Thornhill, R. (1997). Fluctuating asymmetry and psychometric intelligence. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *264*, 823–829.

Gabora, L., & Kaufman, S. B. (2010). Evolutionary approaches to creativity. R. J. Sternberg & J. C. Kaufman (Eds.), *The Cambridge Handbook of Creativity* (pp. 279–301). New York: Cambridge University Press.

Gangestad, S.W., & Simpson, J. A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences*, *23*, 573–644.

Garcia, J. R., & Reiber, C. (2008). Hook-up behavior: A biopsychosocial perspective. *Journal of Social, Evolutionary, and Cultural Psychology*, *2*, 192–208.

Garcia, J. R., Reiber, C., Massey, S. G., & Merriwether, A. M. (2012). Sexual hookup culture: A review. *Review of General Psychology*, *16*(2),161.

Geary, D. C. (2004). *Origin of mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: American Psychological Association.

Geary, D. C. (2009). Evolution of general fluid intelligence. In S. M. Platek & T. K. Shackelford (Eds.), *Foundations in evolutionary cognitive neuroscience* (pp. 25–26). Cambridge, MA: MIT Press.

Geher, G. (Ed.). (2004). *Measuring emotional intelligence: Common ground and controversy*. New York: Nova Science.

Geher, G. (2009). Accuracy and oversexualization in cross-sex mind-reading: An adaptationist approach. *Evolutionary Psychology*, *7*, 331–347.

Geher, G. (2014). *Evolutionary Psychology 101*. New York: Springer.

Geher, G., Betancourt, K., & Jewell, O. (2017). The link between emotional intelligence and creativity. *Imagination, Cognition and Personality*, *37*(1), 5–22. https://doi.org/10.1177/0276236617710029

Geher, G., Camargo, M. A., & O'Rourke, S. (2008). Future directions in research on mating intelligence. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system* (pp. 395–425). Mahwah, NJ: Erlbaum.

Geher, G., & Kaufman, S. B. (2007). The mating intelligence scale. *Psychology Today*, *40*, 78–79.

Geher, G., & Kaufman, S. B. (2013). *Mating intelligence unleashed: The role of the mind in sex, dating, and love*. New York: Oxford University Press.

Geher, G., & Miller, G. F. (Eds.). (2008). *Mating intelligence: Sex, relationships, and the mind's reproductive system*. Mahwah, NJ: Erlbaum.

Geher, G., Miller, G., & Murphy, J. (2008). Mating intelligence: Toward an evolutionarily informed construct. In G. Geher & G. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system* (pp. 3–34). Mahwah, NJ: Erlbaum.

Geher, G., & Renstrom, K. L. (2004). Measurement issues in emotional intelligence research. In G. Geher (Ed.), *Measuring emotional intelligence: Common ground and controversy*. New York: Nova Science.

Giosan, C. (2006). High-k strategy scale: A measure of the high-k independent criterion of fitness. *Evolutionary Psychology*, *4*, 394–405.

Girotto, V., & Tentori, K. (2008). Is domain-general thinking a domain-specific adaptation? *Mind & Society*, *7*, 167–175.

Greengross, G., & Miller, G. F. (2008). Dissing oneself versus one's rivals. *Evolutionary Psychology*, *6*, 393–408.

Greengross, G., & Miller, G. (2011). Humor ability reveals intelligence, predicts mating success, and is higher in males. *Intelligence*, *39*(4), 188–192. https://doi.org/10.1016/j.intell.2011.03.006

Griskevicius, V., Cialdini, R. B., & Kenrick, D. T. (2006). Peacocks, Picasso, and parental investment: The effects of romantic motives on creativity. *Journal of Personality and Social Psychology*, *91*, 63–76.

Gottschall, J., & Wilson, D. S. (2005). *The literary animal: Evolution and the nature of narrative*. Evanston, IL: Northwestern University Press.

Harrison, A., Moore, M., & Rucker, M. (1985). Further evidence on career and family compatibility among eminent women and men. *Archivio di Psicologia, Neurologia e Psichiatria*, *46*, 140–155.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, *78*, 81–91.

Haselton, M. G., Buss, D. M., Oubaid, V., & Angleitner, A. (2005). Sex, lies, and strategic interference: The psychology of deception between the sexes. *Personality and Social Psychology Bulletin*, *31*, 3–23.

Haselton, M. G., & Miller, G. F. (2006). Women's fertility across the cycle increases the short-term attractiveness of creative intelligence compared to wealth. *Human Nature*, *17*, 50–73.

Hasson, O. (2006). The role of amplifiers in sexual selection: An integration of the amplifying and the Fisherian mechanisms. *Evolutionary Ecology*, *4*, 277–289.

Hughes, S., & Gallup, G. G., Jr. (2003). Sex differences in morphological predictors of sexual behavior: Shoulder to hip and waist to hip ratios. *Evolution and Human Behavior*, *24*, 173–178.

Jensen A. (1998). *The g factor: The science of mental ability*. London: Praeger.

Kanazawa, S. (2004). General intelligence as a domain-specific adaptation. *Psychological Review*, *111*, 512–523.

Kanazawa, S. (2008). The independence of mating intelligence and general intelligence. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system*. Mahwah, NJ: Erlbaum.

Kanazawa, S. (2010). Evolutionary psychology and intelligence research. *American Psychologist*, *65*, 279–289.

Kaufman, S. B., DeYoung, C. G., Reis, D. L., & Gray, J. R. (2011). General intelligence predicts reasoning ability even for evolutionarily familiar content. *Intelligence*, *39*, 311–322.

Kaufman, S. B., Erickson, J. E., Huang, J. Y., Ramesh, S., Thompson, S., Kozbelt, A., Paul, E., & Kaufman, J. C. (2009). Art as an aphrodisiac. Paper presented at the Northeastern Evolutionary Psychology Society Conference, Oswego, NY.

Kaufman, S. B., Erickson, J. E., Ramesh, S., Kozbelt, A., Magee, M., & Kaufman, J. C. (2010). What are funny people like? Paper presented at the Human Behavior and Evolution Society Conference, Eugene, OR.

Kaufman, S. B., & Kaufman, J. C. (Eds.). (2009). *The psychology of creative writing*. Cambridge: Cambridge University Press

Kaufman, S. B., Kozbelt, A., Bromley, M. L., & Miller, G. F. (2008). The role of creativity and humor in human mate selection. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system* (pp. 227–263). Mahwah, NJ: Erlbaum.

Keller, M., & Miller, G. F. (2006). An evolutionary framework for mental disorders: Integrating adaptationist and evolutionary genetics models. *Behavioral and Brain Sciences*, 29, 429–452.

Kirsner, B. R., Figueredo, A. J., & Jacobs, W. J. (2003). Self, friends, and lovers: Structural relations among Beck Depression Inventory scores and perceived mate values. *Journal of Affective Disorders*, *75*(2), 131–148. https://doi.org/10.1016/S0165-0327(02)00048-4

Klasios, J. (2013). Cognitive traits as sexually selected fitness indicators. *Review of General Psychology*, *17*(4), 428–442. https://doi.org/10.1037/a0034391

Kordsmeyer, T. L., & Penke, L. (2017). The association of three indicators of developmental instability with mating success in humans. *Evolution and Human Behavior*, *38*(6), 704–713. https://doi.org/10.1016/j.evolhumbehav.2017.08.002

Kruger, D. J., Reischl, T. M., & Zimmerman, M. A. (2008). Time perspective as a mechanism for functional developmental adaptation. *Journal of Social, Evolutionary, and Cultural Psychology*, *2*, 1–22.

Lange, B. P., & Euler, H. A. (2014). Writers have groupies, too: High quality literature production and mating success. *Evolutionary Behavioral Sciences*, *8*(1), 20–30. https://doi.org/10.1037/h0097246

Locke, J. L. (2017). Emancipation of the voice: Vocal complexity as a fitness indicator. *Psychonomic Bulletin & Review*, *24*(1), 232–237. https://doi.org/10.3758/s13423-016-1105-7

Lynn, C. D., Pipitone, R. N., & Keenan, J. P. (2014). To thine own self be false: Self-deceptive enhancement and sexual awareness influences on mating success. *Evolutionary Behavioral Sciences*, *8*(2), 109–122. https://doi.org/10.1037/h0097255

MacArthur, R. H., & Wilson, E. O. (1967). *Theory of island biogeography*. Princeton: Princeton University Press.

Matthews, G., Zeidner, M., & Roberts, R. D. (2002). *Emotional intelligence: Science and myth*. Cambridge, MA: MIT Press.

Mayer, J. D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence*, *22*, 89–113.

McKeown, G. J. (2013). The Analogical Peacock Hypothesis: The sexual selection of mind-reading and relational cognition in human communication. *Review of General Psychology*, *17*(3), 267–287. https://doi.org/10.1037/a0032631

Miller, G. F. (1998). How mate choice shaped human nature: A review of sexual selection and human evolution. In C. B. Crawford & D. L. Krebs (Eds.), *Handbook of evolutionary psychology: Ideas, issues, and applications* (pp. 87–129). Mahwah, NJ: Erlbaum.

Miller G. F. (1999). Sexual selection for cultural displays. In R. Dunbar, C. Knight, & C. Power (Eds.), *The evolution of culture* (pp. 71–91). Edinburgh: Edinburgh University Press.

Miller, G. F. (2000a). *The mating mind: How sexual choice shaped the evolution of human nature*. New York: Doubleday.

Miller, G. F. (2000b). Mental traits as fitness indicators: Expanding evolutionary psychology's adaptationism. In D. LeCroy & P. Moller (Eds.), *Evolutionary perspectives on human reproductive behavior* (pp. 62–74). New York: New York Academy of Sciences.

Miller, G. F. (2000c). Sexual selection for indicators of intelligence. *Novartis Foundation Symposium*, *233*, 260–270.

Miller, G. F. (2001). Aesthetic fitness: How sexual selection shaped artistic virtuosity as a fitness indicator and aesthetic preferences as mate choice criteria. *Bulletin of Psychology and the Arts*, *2*, 20–25.

Miller, G. F. (2007). Sexual selection for moral virtues. *Quarterly Review of Biology*, *82*(2), 97–125.

Miller, G. F. (2008). Mating intelligence: Frequently asked questions. In G. Geher & Miller, G. F. (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system* (pp. 367–393). Mahwah, NJ: Erlbaum.

Miller, G. F., Tybur, J., & Jordan, B. (2007). Ovulatory cycle effects on tip earnings by lap-dancers: Economic evidence for human estrus? *Evolution and Human Behavior*, *28*, 375–381.

Mithen, S. (2006). *The singing Neanderthals: The origins of music, language, mind, and body*. London: Weidenfeld and Nicolson.

Nettle, D. (2009). The evolution of creative writing. In S. B. Kaufman & J. C. Kaufman (Eds.), *The psychology of creative writing* (pp. 101–117). Cambridge: Cambridge University Press.

Nettle, D., & Clegg, H. (2006). Schizotypy, creativity and mating success in humans. *Proceedings of the Royal Society: B*, *273*, 611–615.

Nettle, D., & Clegg, H. (2008). Personality, mating strategies, and mating intelligence. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system*. Mahwah, NJ: Erlbaum.

O'Brien, D., Geher, G., Gallup, A., Garcia, J., & Kaufman, S. B. (2010). Self-perceived mating intelligence predicts sexual behavior in college students: Empirical validation of a theoretical construct. *Imagination, Cognition and Personality*, *29*, 341–362.

O'Sullivan, M. (2008). Deception and self-deception as strategies in short- and long-term mating. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system*. Mahwah, NJ: Erlbaum.

Penke, L., Denissen, J. J., & Miller, G. F. (2007). The evolutionary genetics of personality [target article]. *European Journal of Personality*, *21*, 549–587.

Penke, L., Todd, P., Lenton, A. P., & Fasolo, B. (2008). How self-related cognitions can guide human mating decisions. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: Sex, relationships, and the mind's reproductive system*. Mahwah, NJ: Erlbaum.

Perina, K. (2007). Love's loopy logic. *Psychology Today*, *40*, 68–77.

Pérusse, D. (1993). Cultural and reproductive success in industrial societies: Testing the relationship at the proximate and ultimate levels. *Behavioral and Brain Sciences*, *16*, 267–322.

Peterson, A., Carmen, R., & Geher, G. (2013). Ovulatory shifts in mating intelligence. *Journal of Social, Evolutionary, and Cultural Psychology*, *7*(1), 66–75. https://doi.org/10.1037/h0099173

Peterson, A., Geher, G., & Kaufman, S. B. (2011). Predicting preferences for sex acts: Which traits matter most, and why? *Evolutionary Psychology*, *9*(3), 371–389. https://doi.org/10.1177/147470491100900308

Pipitone, R. N., & Gallup, G. G. (2008). Women's voice attractiveness varies across the menstrual cycle. *Evolution and Human Behavior*, *29*, 268–274.

Prokosch, M. D., Coss, R. G., Scheib, J. E., & Blozis, S. A. (2009). Intelligence and mate choice: Intelligent men are always appealing. *Evolution and Human Behavior*, *30*, 11–20.

Prokosch, M. D., Yeo, R. A., & Miller, G. F. (2005). Intelligence tests with higher g-loadings show higher correlations with body symmetry: Evidence for a general fitness factor mediated by developmental stability. *Intelligence*, *33*, 203–213.

Rosenberg, J., & Tunney, R. J. (2008). Human vocabulary use as display. *Evolutionary Psychology*, *6*, 538–549.

Schmitt, D. P. (2005). Fundamentals of human mating strategies. In D. M. Buss (Ed.), *The handbook of evolutionary psychology*. NewYork: Wiley.

Silventoinen, K., Posthuma, D., van Beijsterveldt, T., Bartels, M., & Boomsma, D. I. (2006). Genetic contributions to the association between height and intelligence: Evidence from Dutch twin data from childhood to middle age. *Genes, Brain and Behavior*, *5*, 585–595.

Simpson, J. A., & Gangstead, S.W. (1991). Individual differences in sociosexuality: Evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology*, *60*, 870–883.

Śmieja, M., & Stolarski, M. (2018). Assortative mating for emotional intelligence. *Current Psychology*, *37*(1), 180–187.

Springer, J. (2009). Evolution's Match.com. Psy-chCRITIQUES, 53.

Sternberg, R. J. (1996). *Successful intelligence*. New York: Simon & Schuster.

Sternberg, R. J. (1998). *Handbook of creativity*. Cambridge: Cambridge University Press.

Sundet, J. M., Tambs, K., Harris, J. R., Magnus, P., & Torjussen, T. M. (2005). Resolving the genetic and environmental sources of the correlation between height and intelligence: A study of nearly 2600 Norwegian male twin pairs. *Twin Research and Human Genetics*, *8*, 307–311.

Taylor, M. D., Hart, C. L., Smith, G. D., Whal-ley, L. J., Hole, D. J., Wilson, V., & Deary, I. J. (2005). Childhood IQ and marriage by mid-life: The Scottish Mental Survey 1932 and the Midspan studies. *Personality and Individual Differences*, *38*, 1621–1630.

Tornquist, M., & Chiappe, D. (2015). Effects of humor production, humor receptivity, and physical attractiveness on partner desirability. *Evolutionary Psychology*, *13*(4), 1–13. https://doi.org/10.1177/1474704915608744

Veblen, T. (1899). *The theory of the leisure class*. New York: Macmillan.

Wagstaff, D. L., Sulikowski, D., & Burke, D. (2015). Sex-differences in preference for looking at the face or body in short-term and long-term mating contexts. *Journal of Evolutionary Psychology*, *13*(1), 1–17.

Wilson, D. S. (2007). *Evolution for everyone*. New York: Bantam.

Zahavi, A. (1997). *The handicap principle: A missing piece of Darwin's puzzle*. New York: Oxford University Press.

Zebrowitz, L.A., & Rhodes, G. (2004). Sensitivity to "bad genes" and the anomalous face over-generalization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior*, *28*, 167–185.

36 Consumer and Marketer Intelligence

Harish Sujan and Mita Sujan

The *Encyclopedia Britannica*, a superbly crafted and well-marketed set of reference books that did much to enhance the knowledge of its readers, is now in an age well in the past. Its impact was enormous, yet it pales when contrasted with what Google, the search engine, has achieved in the current age. As a result of Google, picking on one example of its impact, a large number of professionals, such as doctors, expect their clientele to be remarkably well-informed and feel the need to retool their own knowledge and skills in order to continue to stay a relative expert. Google, though at the top of the heap, is only one of many new products and services that, through remarkable marketing, have considerably altered the lives of consumers living in current times.

The marketing of new products and services can, and often does, make our lives better. But there are products and services that have made our lives worse; notable examples are the marketing of opioid-based pain killers and mortgage loans that borrowers have very little chance of being able to pay back. Skillful marketing impacts consumers both positively and negatively, both fulfilling and taking away some of our most cherished dreams.

Two stories illustrate marketing's potential for high impact, good and bad. Both are from the developing world. In 1976, Dr. V. retired from working for the Government of India, at age fifty-eight, and chose to invest his life savings in a small hospital in which a few of his family members and he would perform cataract surgery (Rangan, 2009). He was inspired by Mohandas Gandhi's sense of social justice: "I suggest we are all thieves in a way. If I take anything that I do not need for my immediate use and keep it, I thieve it from somebody else" (Gandhi, 1960, p. 3). Instead of keeping as profit the difference between the revenue from these surgeries and the cost of doing them, he used the excess inflow of money to perform free surgeries. Between then and now, Aravind Eye Hospitals, the enterprise, grew into several hospitals in different locations and have performed 4 million surgeries, 75 percent of them free. Dr. V.'s idea has substantially alleviated blindness among people living at the bottom of the economic pyramid. The number of surgeries that Aravind performs is so large that it amounts to about 60 percent of all eye surgeries carried out in the United Kingdom, which, although less populated than India, is a large country. Even more remarkable than the volume is Aravind's ability to perform surgeries at a very low cost, about 1/1,000th of the cost of an eye surgery in the United Kingdom. They are able to do this, judging by postsurgery complications, while keeping the quality level about the same (Rosenberg, 2013). Aravind is a story of the creation of products and services that have reached a large number of consumers, fulfilling a basic need, and at a price they can afford. Aravind has enhanced the well-being of its clients and has gained their trust. Their consumers' trust has translated into a positive, strong brand reputation that causes others, potential consumers, to seek them out.

At about the same time as Dr. V. started Aravind, the Swiss company Nestlé, after discovering that the market for infant formula in the developing world had become saturated, sought markets for their formula in the less developed part of the world. They started with the Dominican Republic. Their marketing acumen led them to recognize that Bottom of the Pyramid consumers in this country had a deep desire to see their babies as healthy and as well-nourished as those in the wealthy, developed part of the world. They placed pictures of a Western baby next to pictures of their product package, in billboards and other commercials, and they supplied free samples to hospitals; new mothers who tried these samples lost their ability to breastfeed. Since buying enough of the product was beyond the economic means of Bottom of the Pyramid families, mothers frequently either overdiluted the formula or used rice water as a substitute. This led to severe malnutrition among infants in the Dominican Republic. Infant formula, rather than turning babies into well-fed, well-nourished Western look-alike babies, was doing the exact opposite. Not only were these babies malnourished but, in addition, had frequent bouts of diarrhea. The water mothers used to prepare formula was often contaminated and the bottles were often far from sterilized.

Nestlé defended their marketing actions by saying that it would have been unethical for them to deny Bottom of the Pyramid consumers access to their product. The families that used formula instead of breast milk had freely made this choice. Despite their defense, they encountered a boycott, in the Western world, and were threatened by the possibility of a sharp drop in the equity of their highly valued brand. Nestlé's marketing had been outstanding, perhaps even more outstanding than Aravind's – great advertising, great sales promotion. Yet the consequence of their marketing on human welfare was quite the opposite. Aravind had dramatically increased human welfare while Nestlé had dramatically decreased it. And, as it turned out, this decrease in human well-being had the potential for eroding trust in the brand Nestlé, not just in the Dominican Republic but worldwide. Senior management at Nestlé showed maturity and wisdom by backing away from their irresponsible marketing. As a result, they maintained their brand equity. They continue to be on Interbrand's list of best global brands (Interbrand, 2017). More noteworthy is that they feature high on Fortune's list of companies that are changing the world for the better, through making the provision of social benefit an integral part of their corporate agenda (McGirt, 2017).

The examples suggest that marketing strategies that have positive short-term consequences, sales, profit, and customer satisfaction, can differ considerably in their long-term consequences. So, when consumers themselves make the decision not to reject the marketer's offer, they could still experience a reduction in their well-being that they only discover significantly later. Marketing strategies that enable sales, profit, and customer satisfaction are generally considered intelligent, even if there is a significant possibility that long-term customer well-being could be compromised. The driver of long-term customer well-being is something that goes beyond intelligence.

In his research on human intelligence, Sternberg (1998) suggests that intelligence applied to practical pursuits is generally aimed at maximizing the achievement of practical outcomes: consequences that are immediate and of interest primarily to the self. It requires going beyond intelligence, into wisdom, to balance the short run with the long run and self-interests with the interests of others. Good leaders, Sternberg suggests, act not just intelligently but in addition act wisely (see also Sternberg, 2013). Translated to marketing, Sternberg's perspective is that marketing intelligence leads to short-run positive outcomes, sales, profit, and short-run customer satisfaction, while marketing wisdom leads to long-run positive outcomes, notably a strong reputation for the brand.

The perspective on good business that Csikszentmihalyi (2003) provides reinforces this. Csikszentmihalyi separates one-year managers from hundred-year managers. Hundred-year managers (there are many more one-year managers) look deep into the future while making their strategic choices. Dr. V's choices made him a hundred-year manager. Nestlé's management evolved from a one-year focus to a hundred-year focus. Consistent with both Sternberg's and Csikszentmihalyi's characterization of good leadership, we judge marketing intelligence by the effect the strategies chosen have on sales, profit, and immediate customer satisfaction. Feeling the need to provide a broader perspective, we supplement an analysis of marketing intelligence with an analysis of marketing wisdom. We justify this by suggesting that wise marketing enhances brand reputation, an outcome that is often recognized as extremely important.

The balance adopted by wise managers, to focus beyond their corporation's immediate well-being, quite ironically tends to do much for the long-term financial health of their company. Brand equity tends to be a sizeable part of the net worth of corporations. Looking at the 2018 numbers that Forbes provides, for many of the more successful companies, examples are Apple, Amazon, and Walmart, brand equity is a large part of their net worth (Badenhausen, 2018). Apple, Amazon, and Walmart, quite evidently, care for their customers and focus on the long run.

For consumers, intelligent behavior is in the service of acquiring what is currently needed, at a price that is not a rip-off and without having to go through hoops to achieve this. Wise behavior goes beyond this by recognizing the superficiality of some needs, selectively pursuing important needs, and being willing to go through hoops, if necessary, to fulfill these needs. The fact that much consumer acquisition does little to enhance consumers' well-being (Myers, 1992) suggests that consumer behavior tends to be more intelligent than wise. There are domains in which consumer behavior is more likely to be wise, for example in the pursuit of religion and exercise (Mochon, Norton, & Ariely, 2008). Csikszentmihalyi (2003) points out that, although in general consumer behavior tends to be at a low level of complexity, intelligent but not wise, in a few domains consumers wisely pursue distant, challenging goals, driven by a desire to self-actualize.

Marketing Intelligence Through the Lens of Practical Intelligence

Practical intelligence is about managing the external, real world through adapting to it, through selecting exposure to it, and through shaping or altering elements of this external world. It enables immediate outcomes of importance to the self. Adaptive intelligence relates to knowledge and skills that enable actions that are appropriate for the external context. Selection intelligence relates to choosing, if one has this choice, external environments that better enable effective adaptation. Shaping intelligence relates to changing the external environment, if this is possible, to better enable effective adaptation (Sternberg, 1984, 1985, 1998, 2013). This perspective – individuals' adapting to, selecting, and shaping the world around them to achieve important practical outcomes – goes well with the agenda marketing research has been pursuing. It offers a helpful lens by which to view the impact of marketing strategies on sales, profit, and immediate customer satisfaction (see Figure 36.1).

Marketer Adaptation

A seminal paper written by two researchers who considerably shaped the field of marketing research suggested a broadened view of what marketing is (Kotler & Levy, 1969). They asked that marketing not be considered, as it was then, a narrow activity restricted to the "selling of toothpaste, soap and steel" (p. 10). The purview of marketing, they suggested, is considerably broader: encompassing the fulfillment of human needs through services, ideas, and organizations, both business and nonbusiness. An indication of Kotler and Levy's impact is that the title of a popular book on a successful US presidential campaign in 1960 was *The Making of a President*, the title of a similar book in 1968 was *The Selling of a President*, and, in 1992, the title became *The Marketing of a President*. A movement from a production focus to a sales focus to a marketing focus is precisely the trajectory advocated by researchers of the marketing concept (Kohli, Jaworski, & Kumar,

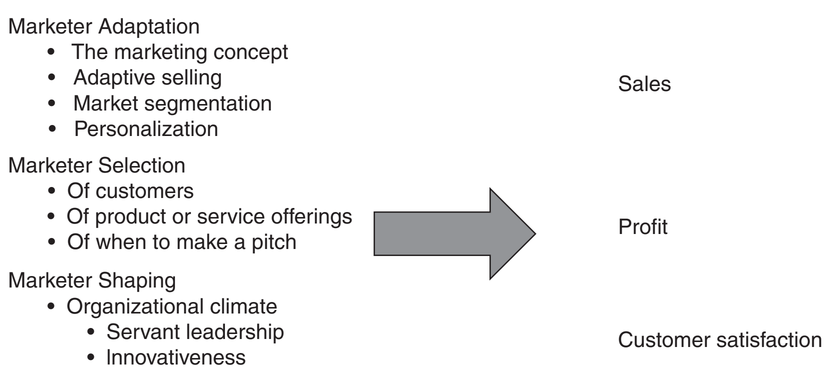

Figure 36.1 *Marketer intelligence.*

1993; Saxe & Weitz, 1982). Under a production focus, attention is directed at engineering a near-perfect product; under a sales focus, attention is directed at what would cause people to buy the product; and, under a marketing focus, attention is directed at what customers need. Notwithstanding the emphasis placed on the marketing concept by researchers, a product and a sales focus have far from disappeared. With rare exceptions, recent encounters with a doctor should remind us that customer-oriented persuasion is not a universal norm. Among doctors, customer *compliance* – a sales focus – has been and continues to be the important yardstick for success.

Progressively, researchers recognized that what the customer needed was influenced by the customer's environment and that the environment also influenced whether or not the customer sought to fulfill this need. The seeking of information on customers' needs and the context affecting these needs was given the term "intelligence generation." Since a group of people within an organization, not just an individual, often decided on marketing actions, intelligence generation was supplemented with "intelligence dissemination" – the sharing of information gathered. The inclusion of a third construct, "responsiveness," choosing marketing actions that were appropriate to the customer's needs and context, brought the marketing concept into the realm of intelligent adaptation (Kohli et al., 1993).

Personal Selling

The subarea of marketing that gave considerable importance to the idea of marketer adaptation was research on personal selling effectiveness. Individual salespeople who were good at reading customer needs and responding with adapted selling behavior tended to be more effective. Since salespeople knew their customers one on one, it was recognized that they had the opportunity to personalize their communications. Advertisers, in contrast, could only adapt to a cluster of consumers. This, it was suggested, allowed salespeople to apply the marketing concept better: selling needed to be recognized as a more empowered form of marketing communication than advertising (Weitz, 1978; Weitz, Sujan, & Sujan, 1986). Following years of research that tested the benefits of adaptive selling in retail sales, business-to-business sales, and even in trade shows, its value became well established. A meta-analysis that spanned 31,000 salespeople demonstrated that adaptive selling enhances not just performance but also salespeople's own job satisfaction (Franke & Park, 2006). Thus, adaptive marketing behavior by salespeople, quite evidently, qualifies as intelligent behavior.

Market Segmentation

The core idea in market segmentation is to divide markets into clusters, based on consumer demographics or psychographics, identify clusters more likely to respond well to the marketer's offering, and formulate ways to present the offering to this segment of the market that makes it more likely to be well received. This form of adaptive marketing has been, and is, widely practiced in advertising. The sophistication of going beyond demographics to psychographics – attitudes, interests, and opinions – to form clusters was seen as a remarkable step forward. Clusters could be based not just on characteristics like household income and geographical location but also on lifestyle characteristics such as "The Ethical Highbrow" and "The High Achiever" (Wells, 1975). Important to segmentation research was being able to identify consumers whose responsiveness to a particular marketing action is nearly identical. This question has been, and continues to be, answered with the help of sophistical mathematical models (Grover & Srinivasan, 1987; Jia, Wang, & Xiong, 2017; Obilo & Alford, 2018). Beyond personalized selling, messages tailored to market segments constitute adaptive intelligence.

Personalization

With the advent of online retailing, the dynamics of gathering information, market intelligence, that enables adaptive marketing behavior, has changed. While in-store adaptation, learning about the customer through conversation and personalizing the appeal, in general, has positive consequences on performance (Mittal & Lassar, 1996), online personalization may or may not have a positive consequence. Offsetting the benefit of tailored appeals are customers' concerns that information on them was gathered covertly. The feeling of vulnerability that the covert gathering of information creates may result in a negative consequence, notwithstanding the customer's appreciation of a personalized appeal (Aguirre et al., 2015). At times, marketers need to adapt not just to their customers' needs but also to their selling context. Effective online personalization, adaptation, requires both a tailored message and transparency of the process by which the message was created.

The adaptive element of Sternberg's theory of contextual intelligence (Sternberg, 1984) has been researched significantly in marketing. Considerable evidence exists for adaptive intelligence enhancing real-world performance. Adaptation, this research clarifies, is not just about tailoring messages to the customer's need but also about sensitivity to the marketing exchange process.

Marketer Selection

The customers marketers choose to approach and build a relationship with, if selected poorly, could cause poor results despite intelligent adaptation. If selected well, the choice would make the task of adaptation considerably less onerous. The offering, the product or service, the marketer chooses to market considerably changes the task of marketing. A product/service proposition seen by customers as having little value could negate the benefit of intelligent adaptation. The choice of when to make a particular persuasive appeal significantly influences the likelihood of this appeal's effectiveness. There are poor times to make a pitch and excellent times to make the very same pitch.

Customer Lifetime Value

A body of research in the marketing literature has been labeled customer lifetime value. It is based on the notion, expressed within research on the marketing concept, that both customers and marketers need to extract value from the exchange; for marketers, unless they are a nonprofit organization, the value includes profit. Marketers compute the investments they expect to make in their relationship with a customer as well as the returns they expect to get, over the lifetime of the relationship, convert these numbers to a present value, and evaluate the profit they will make. For the marketer to value this relationship, the profit needs to be over the threshold they specify (Singh & Jain, 2010). Although this perspective has something in common with the investment model of relationship maintenance (Rusbult, Johnson, & Morrow, 1986), there is an important difference. In the investment model perspective, people will not walk away from a relationship that is less satisfying than an alternative if the difference is less than the psychological investment they have in the existing relationship. Customer lifetime value is about entering a relationship rather than exiting it. It suggests that firms project their investments over time and work through the question of sunk costs before they choose to enter a relationship.

Since it is a precise mathematical idea, customer lifetime value research is concerned with questions of how well the costs and benefits have been projected. Illustratively, a customer who engages in positive word-of-mouth that causes other consumers to become customers and existing customers to buy more provides additional returns that need to be factored in. Not including the revenue impact of positive word-of-mouth may lead to a decision to not engage with this customer even though the profit this customer would generate would be well over the marketer's threshold. Thanks to an abundance of research, the availability of data, complex modeling skills, and careful conceptualization (e.g., factoring the cost of customer acquisition, relationship management, and retention), this research has proved to be very helpful for firms in their making prudent selection decisions on which customers to engage with and which ones not to (Singh & Jain, 2010).

The Value Proposition

Csikszentmihalyi (2003) has argued that human well-being has been deeply impacted in the last few decades through the progress that science and business, working in tandem, has accomplished. The consumption of some goods and services has proved to be invaluable in enhancing human welfare. Implicitly alluding to their offerings enhancing consumer welfare, marketers frequently use the phrase "value proposition." Specifically, this phrase alludes to the benefits being offered by a product or service and the price being charged to acquire it: whether or not a cost-benefit analysis makes the offering a highly, modestly, or marginally valuable proposition (Skålén et al., 2015). For some products, where consumers are price inelastic, the price does little to change the value proposition. For some other products, price is a very important part of the proposition. To illustrate, a proposition that has been very highly valued is the one that Google offers: they more than any other search engine will provide the most relevant answers to browsers' questions; they will do this at no cost but want in return to track individual browsers' interests and behaviors. Google consumers appear to be privacy inelastic in that the amount of personal information being tracked does not alter their consumption of Google. Another illustration is Uber's proposition, highly valued but less so than Google's: convenient and inexpensive travel. Uber consumers are price elastic: consuming less when prices are higher (A. Payne, Frow, & Eggert, 2017).

The greater the extent to which a firm provides products and/or services that customers value highly, the greater is their market performance and the lesser is their need for adaptation. Since firms continually work at innovating their offerings, conduct research and development to identify newer and better services and products, and choose to bring to market very few of their ideas, their value propositions are a form of intelligence related to selection.

The reason driving the choice of what product or service to offer, at what price, is often based on more than how much it will improve the consumer's life. To illustrate, the choice can be driven by whether or not the product, or service, can be made efficiently and reliably, whether or not skills exist within the marketing group and the advertising agency to promote the product or service well, and whether or not likely customers have the financial resources and discipline to pay for the product/service on time (Skålén et al., 2015). The choice can be driven primarily by the interests of customers, managers, or shareholders (A. Payne et al., 2017). The motives guiding the choice could, quite evidently, change the intelligence of the choice.

The process by which new products and services are developed and marketed, spelled out in the marketing literature, focuses more on creative intelligence than it does on the selection aspect of contextual intelligence (Alam & Perry, 2002; Wind & Mahajan, 1997). Nevertheless, identified clearly is the need for good selection early in the development process, to focus on value propositions that potentially are highly valued, and good selection late in the development process to choose what to go to market with and how.

Unilever, a Dutch-British company, markets beauty products under the brand name Dove. In response to a public concern that the pursuit of beauty among women was, too prevalently, leading to an eating disorder, they launched in 2004 a real beauty campaign that pictured everyday women and not models. The purpose was to give women a sense of confidence and self-esteem about their current appearance. Hindustan Unilever is the Indian subsidiary of Unilever. Recognizing a bias against dark shades of skin color among Indians, they marketed a cream called Fair and Lovely that lightens skin color (Deshpande & Chaturvedi, 2017). The market they achieved was huge. Based on the criteria of sales, profit, and customer satisfaction, this choice of product has been an enormous success. They have displayed selection intelligence. In 2009, an organization called Women of Worth launched a Dark is Beautiful campaign that, consistent with the Dove campaign, was designed to help Indian women feel that their unaltered appearance was beautiful. Consistent with research on authentic versus hubristic pride, Women of Worth felt that attempting to be fairer than you naturally are creates a hubris that prompts a search for grandiosity. In contrast, being proud of being the skin shade you are enables a quiet sense of authentic pride. Hubristic pride does not provide comfort while authentic pride does (Tracy et al., 2009). Since Hindustan Unilever continues to sell large amounts of Fair and Lovely cream, despite the Women of Worth campaign, and achieve profit and customer satisfaction, from the perspective we have taken on practical intelligence they have demonstrated selection intelligence. Should the Women of Worth campaign create a backlash this intelligence will prove to have been without wisdom!

Philips, a Dutch company, develops prototypes of a large number of products, only a few of which they take to market. A product they had developed well before the launch of Facebook was an electronic button that stored a large number of personal likes and dislikes. When two people wearing their buttons came physically close to one another the buttons communicated using something that resembled modern-day Bluetooth and beeped if the shared likes crossed a threshold. Philips never brought this product to market. The success of Facebook, a product considerably different and more elaborate than Philips' beeping button, suggests that they may have shown in this choice a lack of selection intelligence.

PreSuasion: Cialdini's Extension of His Principles of Influence

Marketing research has been significantly influenced by Cialdini's six principles of influence (Cialdini, 1988). Each of these principles has been recognized by marketing researchers as an effective way to achieve persuasion. For example, while trying to persuade hotel guests to reuse their towels, telling them that many other current guests are reusing their towel (consensus) persuades (Goldstein, Cialdini, & Griskevicius, 2008). Extending this perspective on persuasion, in a recent book Cialdini (2016) points out that there are "privileged moments, identifiable points of time when an individual is particularly receptive to a communicator's message" (p. 14): times when the principles he enunciated work better. Illustratively, if, as a part of a convention, the movie *An Inconvenient Sequel: Truth to Power* is shown (this is the sequel *to An Inconvenient Truth*, in which Al Gore explains the evidence for climate change), making a pitch to reuse towels in their hotel rooms, because others are going to be doing this too, is likely to be many times more effective than in the absence of this movie. Cialdini's (2016) book, *PreSuasion*, highlights that an important way for marketers to be contextually intelligent is for them to time selectively when they make their persuasion pitches.

Shaping Intelligence: Best Achieved Through Good Leadership

Although marketing personnel, say, advertising copywriters, can shape their working environment, more so when they act collectively than individually, leaders, notably CEOs, are far more empowered to shape climate and through it the delivery of marketing-related offerings. An advertising story published in the 1970s illustrates this. A creative copywriter acquires a new boss who believes copy should be in the service of dry, unimaginative market research data. The copywriter, creative to the extreme, fought against this new culture only to lose his job – and, to add drama to this story, his wife too (Dillon, 1972).

Schneider and his research colleagues (e.g., Ehrhart et al., 2011) have suggested that customers receive significantly better service when those serving them, in turn, receive great service internally – they are treated well and supported by their organization. The notion of putting employees first stems back to Robert Greenleaf's (1972) conception of servant leadership. This leadership style is characterized by placing the good of those led over and above the self-interests of the leader. Good leadership, Greenleaf suggested, is the development of subordinates rather than the glorification of leaders. Servant leadership causes subordinates to contribute to the organization well beyond the formal job requirements (Walumbwa, Hartnell, & Oke, 2010).

An impressive application of these ideas is in the story of the development of Southwest Airlines. Herb Kelleher, a co-founder in 1971 and the CEO till 2007, led with this philosophy: "Your employees come first. And if you treat your employees right, guess what? Your customers come back, and that makes your shareholders happy. Start with employees and the rest follows from that" (10th quote, Brown, 2016; Freiberg, 1996). Employees, gratified by the internal service they received and the joy they experienced from working at Southwest, delivered better than average service to their customers while being paid less than average (Heskett & Hallowell, 1997). This leadership philosophy caused Southwest to move from being a regional Texas carrier, in peril of being ousted from this market in the first few years of its existence by its competitor Braniff Airlines, to being the second largest airline in the world in terms of passengers carried (Fleming, 2017). Not inconsequentially, it has the lowest rate of passenger complaints among US carriers (Statista, 2019).

A story that complements the Southwest Airline story is one of the heroic actions of the service employees at the Taj Mahal Hotel in Mumbai, in November 2008, when terrorists attacked the hotel (Deshpande, 2011). The management of the hotel, to shape organizational climate, recruited more from rural than urban areas and focused on character, notably the showing of respect and empathy, while making their selections. The incentives at the hotel were designed to reward kindness shown to customers. The work environment created a climate for extreme, not just above-average, customer-centric behavior. So, when a group of guests were kept in a locked room with the doors barricaded and the sounds of gunshots outside, a young hostess walked around the room comforting the guests while offering them alcoholic beverages. Several employees lost their lives that night. Some of them could have escaped but chose to stay to take care of their guests. Driving the work environment that enabled this was Ratan Tata, the head of the Tata Group that owns this hotel.

Leadership matters not just for the climate that enables good and even extreme service but also for the development of innovative new products. Research suggests that the vision of the CEO and the extent to which the CEO inspires subordinates to share their vision improves both the quality and the speed of innovation (Sattayaraksa & Boon-itt, 2016). In his book, Isaacson (2011) suggests that Steve Jobs' vision and management style were central to Apple's innovations. The style Isaacson describes Jobs to have adopted is a far cry from the methods used to shape their organizations, and, consequently, their markets, by the leaders at Southwest and the Taj. With a very different style, Jobs shaped Apple's offerings and the world many of us were living in:

> One of the last times I saw him, after I had finished writing most of the book, I asked him again about his tendency to be rough on people. “Look at the results,” he replied. “These are all smart people I work with, and any of them could get a top job at another place if they were truly feeling brutalized. But they don’t.” Then he paused for a few moments and said, almost wistfully, “And we got some amazing things done.” Indeed, he and Apple had had a string of hits over the past dozen years that was greater than that of any other innovative company in modern times: iMac, iPod, iPod nano, iTunes Store, Apple Stores, MacBook, iPhone, iPad, App Store, OS X Lion – not to mention every Pixar film. And as he battled his final illness, Jobs was surrounded by an intensely loyal cadre of colleagues who had been inspired by him for years and a very loving wife, sister, and four children. (Isaacson, 2012, p. 94)

Jeff Bezos, too, shaped an organization, and through this the world of marketing, by creating an everything store. Stone (2013) describes the working climate at Amazon as being far from easy but being part of innovations that stretch the boundaries of retailing was and continues to be strong motivation for working for Amazon. Although some innovations have flopped, there is little doubt that Amazon has created what could hardly be imagined earlier and has shaped the way we, as consumers, live.

Contrasting with these stories is the inability to innovate. Nokia was the market leader in cell phones before the iPhone was introduced and seems to have been paralyzed into inaction after it was. Barnes and Noble was far and away the retailing giant for books when Amazon entered into the market. They dabbled in online retailing but did not have the mindset to innovate as Amazon did. The contrast emphasizes that shaping is by and large the prerogative of the CEO and that there are differences, stark differences at times, in shaping intelligence.

Consumer Intelligence Through the Lens of Contextual Intelligence

Consumer research too goes well with an adapting to, selecting, and shaping perspective on human intelligence (see Figure 36.2).

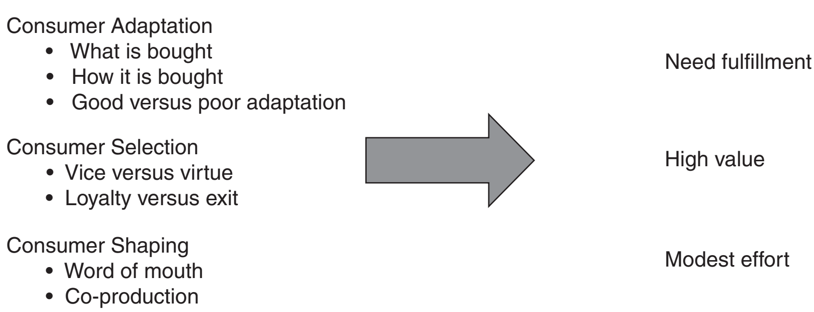

Figure 36.2 *Consumer intelligence.*

Consumer Adaptation

Consumers are adaptive about what they buy. They are also adaptive about how they buy. One form of adaptation while deciding on what to buy is making a trade-off between costs and benefits. A less expensive option may be preferred if, on evaluation, the consumer feels the lower cost outweighs the lower benefit. Consumers adapt to the process of buying by trading off effort with accuracy. Although with more effort they know they can get closer to their optimal choice, they may decide that the incremental effort is not worth the incremental benefit (Bettman, Luce, & Payne, 1998; Luce, Bettman, & Payne, 2001). Baumgartner (2010) suggests that consumer purchase decisions differ on being more or less involved (the outcome matters more or less), more or less deliberative, and more functional versus more psychosocial. Each of these differences alters consumers' decision-making processes. In general, consumers adapt based on their goals and based on the purchasing context. An interesting example of this is a study by Payne, Bettman, and Johnson (1988), in which they found that increasing time pressure caused consumers to adaptively change the heuristic they used to make their choices: it was adaptive in that the change improved accuracy.

One domain in which adaptive decision-making has been studied is consumers' health-related choices. Research suggests that consumers make both adaptive and maladaptive choices. An example of a maladaptive choice is using goal progress as an excuse: when consumers are further along in their goal of achieving weight loss, they choose as a gift a bar of chocolate rather than an apple even though this choice compromises their weight-loss goal (Fishbach & Dhar, 2005). An example of an adaptive choice is preceding an unhealthy dessert with a healthy rather than an unhealthy entrée (Dhar & Simonson, 1999). Research has identified the conditions under which consumers act adaptively and intelligently and the conditions under which they act maladaptively and unintelligently.

Consumer Choice: All About Selection

Consumer-behavior research is dominated by questions relating to consumer choice. While some of the questions are about what consumers choose (brand, retail store, recommender, media channel, and product), more of the questions are about how they go about choosing. An important difference in the process of choosing is whether consumers use preformed attitudes or construct their attitudes while making choices (Bettman et al., 1998). Much of the research on consumer choice does not evaluate the intelligence of the choice, that is, the effect on outcomes of importance to the consumer. Rather the focus in research on consumer choice is on understanding what consumers decide to do or not do. Breaking from the mainstream, two subdomains that associate with intelligence relate to virtue versus vice and loyalty versus exit.

Virtue vs. Vice

Consumer research has separated consumption goals into yielding to temptation and maintaining self-control. Yielding to temptation – an example is the eating of a chocolate cake – is labeled a vice and maintaining self-control – an example is the eating of fruit – is labeled a virtue. Intelligent selection would be evident in actions that skew goal pursuit toward virtue. Actions that skew goal pursuit toward vice would indicate lesser selection intelligence. Illustratively, it has been found that consumers estimate their calorie consumption to be lower when they add a virtue food to a vice food than if they were to eat the same vice food alone (Chernev & Gal, 2010). This is an indication of consumer choice that is less rather than more intelligent. An indication of consumer choices that are more intelligent is from a study of assortment size. Sela, Berger, and Liu (2009) found that, in the face of temptation, consumers make more virtuous choices if they search among larger assortments. This, they reasoned, is because the larger the assortment size the more difficult it becomes for consumers to justify their choice and the more difficult it is to justify a choice the more likely they are to, intelligently, choose virtue. A third illustration relates to price promotions. Promotions cause consumers to buy more. Intelligently, some consumers search for promotions among healthy foods and, less intelligently, other consumers search for promotions among unhealthy foods (Yan et al., 2017).

Loyalty vs. Exit

Hirschman (1970) identified four potential responses of dissatisfied people, consumers, or employees. The responses are exiting the relationship, staying loyal to the relationship, neglecting the relationship, and using one's voice to state an opinion about the relationship. He suggested that using "voice" affords an opportunity for the relationship partner to dispel the cause of dissatisfaction. Exiting the relationship without voice does not afford this opportunity. Explicating this, Rusbult and colleagues (1988) divide the four responses into active versus passive and constructive versus destructive: voice is both constructive and active. Their analysis, complementing Hirschman's exposition, indicates that voice is an intelligent response while passively exiting a relationship without even a protest is not intelligent. If the response to voice is adequate, then exit will not occur and loyalty will prevail. From an investment model of relationship building, investments will not be wasted (Rusbult et al., 1986).

Research suggests that the strength of an existing relationship with a retail store determines whether or not a dissatisfied customer will quietly exit the relationship. If the relationship is strong, voice rather than exit is the more likely response (Haenlein & Kaplan, 2012). Additional research suggests that commitment to a relationship alters the inclination to exit without voice. The more committed a customer is, the more likely they are to complain first and give the service establishment the opportunity to respond (Kim, Lee, & Mattila, 2014). This research suggests that selection intelligence among consumers is associated with the use of voice prior to exit and, as a result, with greater loyalty to retailers, service establishments, and other marketers. This is because exiting a relationship prematurely, just like the pursuit of vice over virtue in consumption, is less intelligent.

Consumer Shaping

Word-of-Mouth

The governance of marketing occurs through a democratic process. Through responses to market research, consumers, collectively, shape marketers' actions. Although cooperatives and word-of-mouth have been part of consumer behavior for long, consumer networking has been modest until the online, Amazon world changed retailing. Going back a few years to when the impact of this new, retailing-world order began to be felt, surveys reported by Floyd and colleagues (2014) indicated that consumers got their information more online than in-store (2013: 12 percent in-store and 52 percent online), trusted online reviews (2012: 70 percent), and changed their behavior based on these reviews (2012: 65 percent of potential customers chose a brand that was not in their original consideration set). There is much to suggest that the reviews consumers post online have shaped and continue to shape the world of retailing and consumer behavior.

Research on online reviews is large and grows rapidly. Among the more fundamental questions asked is the issue of whether volume or valence of the reviews matters more. The answer appears to be that both do, in an interactive manner (Floyd et al., 2014). A question nearly as fundamental is what triggers online reviews. One of the answers is that the interestingness of the product does not generate sustained online conversations as much as the publicness of the product (Berger & Schwartz, 2011). That is, the posting of online reviews is more about shaping public views than it is about private, personal reactions and consumption. A study by Cheema and Kaikati (2010) provides evidence to support the view that consumers are well aware that online reviews change consumer behavior. They found that people who like to be different are unwilling to post positive reviews of products they own, because they recognize that their reviews could lead to broader adoption and take away from their distinctiveness.

The growing research on online reviews offers an understanding of what posts are intelligent in shaping consumption behavior in ways that are consistent with the poster's goals.

Co-production

At times consumers can co-create products with the marketer. An illustration is Ikea furniture, which often requires considerable time, skill, and effort to convert components into a piece of furniture. Co-creation converts consumption from being passive to being active, bolsters consumers' self-worth, and causes them to value the product more (Mochon et al., 2008). If the marketer permits flexibility of creation, then co-creation permits better personalization, and this too adds value to the product (Roggeveen, Tsiros, & Grewal, 2012). Can individual consumers' co-creation shape or alter what other consumers co-create and what other consumers buy?

Tchibo is similar to Starbucks in that it runs a large number of coffee shops but in Germany. The important difference is that it sells a significantly larger range of noncoffee products than Starbucks does. Beyond selling noncoffee products, Tchibo asks its customers to creatively design everyday products that may appeal to other customers too. They, Tchibo, would use the design to make and market these inventions. Illustrations of ideas that Tchibo customers have are a water bottle that has a compartment to keep one's keys (prevents the problem of losing keys in the gym), a chopping board that has a place to store what has been cut (making it significantly more convenient), and a power outlet where inserting large plugs does not crowd out empty sockets (making all sockets usable). Customers have generated as many as 2,000 new product ideas a year that Tchibo has made and marketed. The conversation among Tchibo customers about the products they and others have designed is considerable. It brings Tchibo customers together, often at a Tchibo coffee shop over coffee. It enhances sales and profit as well as customer satisfaction. Highly noteworthy is that the enthusiasm customers feel for Tchibo enhances their loyalty. Tchibo has become, as a result, one of the strongest and most trusted brands in Germany (Petersen & Wathieu, 2012). In 2018, Tchibo was ranked seventh among best product brands, in between Coca-Cola and Apple (Markenranking, 2018).

Beyond Practical Intelligence: Marketer and Consumer Wisdom

The focus of this chapter is on viewing marketer and consumer intelligence through the lens of practical intelligence. The two examples we used to introduce this chapter were chosen to point out that intelligence by itself may not be enough to influence longer-term outcomes, notably consumer well-being and brand reputation. Both Nestlé, creating a market for their infant formula in the developing world, and Aravind, working to fulfill the human need for sight, marketed remarkably intelligently. Nestlé hurt human welfare to an extent that prompted a US congressional hearing while Aravind provided "manna in the wilderness" by giving sight to many who thought their blindness could not be reversed. The difference, we suggest, is that Aravind acted with wisdom and Nestlé did not. Nestlé put the interests of their senior management, and their shareholders, front and center, while Aravind balanced these interests with those of their customers, customers' families, and doctors (for an explanation of the argument that wisdom is a balancing of interests, see Sternberg, 1998). Our examples additionally suggest that, when wisdom supplements intelligence, marketers achieve not only sales, profit, and immediate customer satisfaction but also brand equity.

This suggestion that a balancing of interests enhances brand reputation was spelled out while describing shaping intelligence. We described in this section the evolution of Southwest Airlines. The senior leadership at Southwest, subscribing to the servant leadership philosophy (Walumbwa et al., 2010), put their employees ahead of their customers (Ehrhart et al., 2011) and achieved not just sales, profit, and customer satisfaction but also brand equity. Even today, Southwest is among the world's most admired brands. The story of how employees of the Taj Mahal Hotel in Mumbai countered a terror attack made a similar point. These employees, treated with kindness by management, put their lives on the line for their customers: their behavior caused the hotel's brand equity to grow.

Research on what makes brand equity grow suggests that, beyond an evaluation of liking, customers react emotionally to experiences with brands (Keller, 1993, 2001). This can cause them to feel with intensity affection, attachment, warmth, and other "hot" emotions and move them from simply liking a brand to feeling admiration for it (Park et al., 2010; Thomson, MacInnis, & Park, 2005). It is unlikely that without showing kindness and support for their employees, without the wisdom to focus away from their own and shareholder priorities, senior managements can achieve brand admiration.

Beyond focusing away from the self to other people, focusing away from the present to the future contributes to wisdom. Csikszentmihalyi (2003) separates hundred-year managers from one-year managers to suggest that hundred-year managers are wiser. The wise managements of the Taj and Southwest showed a vision that extended way into the future.

For consumers, wisdom enables well-being. Thinking beyond immediate self-interests and just the present, as it does with marketers, enables wisdom. Research on healthy eating and adopting an active rather than a sedentary lifestyle suggests that consumers' focusing on the future enables their well-being. This research evaluates the effect of a construct labeled "consideration of future consequences" and shows that this consideration enhances self-control (Dassen, Houben, & Jansen, 2015; van Beek, Antonides, & Handgraaf, 2013). When consumers think about the effect of their current behavior on consequences in the distant future they regulate their eating and exercise better. In addition, when they translate distant goals into current plans, they self-regulate better (Wilkowski & Ferguson, 2016).

There is evidence to support the idea that, when consumers move away from focusing on their own needs to prosocial spending, they improve their own happiness (Aknin et al., 2013; Aknin, Dunn, & Norton, 2012). The wisdom to care for others improves consumers' own well-being. The Tchibo story suggested that consumers could derive happiness not just from focusing on others but also from collective activities with other consumers. This suggestion is consistent with the finding, in the literature on close relationships, that happiness comes from taking an "us" rather than a "you" or "me" perspective (Bradbury & Fincham, 1990).

Marketer and consumer wisdom, we suggest, is an important supplement to intelligence derived from adapting to, selecting, and shaping one's environment. It enables the more important goals of brand worth and personal well-being, respectively.

References

Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing*, *91*(1), 34–49.

Aknin, L. B., Barrington-Leigh, C. P., Dunn, E. W., Helliwell, J. F., Burns, J., Biswas-Diener, R., . . . & Norton, M. I. (2013). Prosocial spending and well-being: Cross-cultural evidence for a psychological universal. *Journal of Personality and Social Psychology*, *104*(4), 635–652.

Aknin, L. B., Dunn, E. W., & Norton, M. I. (2012). Happiness runs in a circular motion: Evidence for a positive feedback loop between prosocial spending and happiness. *Journal of Happiness Studies*, *13*(2), 347–355.

Alam, I., & Perry, C. (2002). A customer-oriented new service development process. *The Journal of Services Marketing*, *16*(6), 515–534.

Badenhausen, K. (2018). The world's most valuable brands 2018. *Forbes*, May 23. www.forbes.com/sites/kurtbadenhausen/2018/05/23/the-worlds-most-valuable-brands-2018/#64dc1031610c

Baumgartner, H. (2010). A review of prior classifications of purchase behavior and a proposal for a new typology. In N. K. Malhotra (Ed.), *Review of Marketing Research*, Vol.6 (pp. 3–36). Armonk, NY: M.E. Sharpe.

Berger, J., & Schwartz, E. M. (2011). What drives immediate and ongoing word of mouth? *Journal of Marketing Research*, *48*(5), 869–880.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, *25*(3), 187–217.

Bradbury, T. N., & Fincham, F. D. (1990). Attributions in marriage: Review and critique. *Psychological Bulletin*, *107*(1), 3–33.

Brown, J. (2016). 30 intelligent Herb Kelleher quotes. Addicted2Success. https://addicted2success.com/success-advice/30-intelligent-herb-kelleher-quotes/

Cheema, A., & Kaikati, A. M. (2010). The effect of need for uniqueness on word of mouth. *Journal of Marketing Research*, *47*(3), 553–563.

Chernev, A., & Gal, D. (2010). Categorization effects in value judgments: Averaging bias in evaluating combinations of vices and virtues. *Journal of Marketing Research*, *47*(4), 738–747.

Cialdini, R. B. (1988). *Influence: Science and practice* (2nd ed.). Glenview, IL: Scott, Foresman.

Cialdini, R. B. (2016). *Pre-Suasion: A revolutionary way to influence and persuade* (1st ed.). New York: Simon & Schuster.

Csikszentmihalyi, M. (2003). *Good business: Leadership, flow and the making of meaning*. New York: Viking.

Dassen, F. C. M., Houben, K., & Jansen, A. (2015). Time orientation and eating behavior: Unhealthy eaters consider immediate consequences, while healthy eaters focus on future health. *Appetite*, *91*, 13–19.

Deshpande, R. (2011). *Terror at the Taj Bombay: Customer-centric leadership*. Cambridge, MA: Harvard Business School Publishing.

Deshpande, R., & Chaturvedi, S. (2017). *Fair and lovely versus dark is beautiful*. Cambridge, MA: Harvard Business School Publishing.

Dhar, R., & Simonson, I. (1999). Making complementary choices in consumption episodes: Highlighting versus balancing. *Journal of Marketing Research*, *36*(1), 29.

Dillon, J. (1972). *The advertising man*. New York: Harper's Magazine Foundation.

Ehrhart, K. H., Witt, L. A., Schneider, B., & Perry, S. J. (2011). Service employees give as they get: Internal service as a moderator of the service climate-service outcomes link. *Journal of Applied Psychology*, *96*(2), 423.

Fishbach, A., & Dhar, R. (2005). Goals as excuses or guides: The liberating effect of perceived goal progress on choice. *Journal of Consumer Research*, *32*(3), 370–377.

Fleming, A. (2017). The largest airlines in the world, by passenger count. Tripsavvy. www.tripsavvy.com/the-largest-airlines-in-the-world-53223

Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., & Freling, T. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing*, *90*(2), 217–232.

Franke, G. R., & Park, J.-E. (2006). Salesperson adaptive selling behavior and customer orientation: A meta-analysis. *Journal of Marketing Research*, *43*(4), 693–702.

Freiberg, K. (1996). *Nuts!: Southwest Airlines' crazy recipe for business and personal success* (1st ed.). Austin, TX: Bard Books.

Gandhi, M. K. (1960). *Trusteeship*. Ahmedabad: Navajivan Publishing.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*(3), 472–482.

Greenleaf, R. K. (1972). *The institution as servant.* Cambridge, MA: Center for Applied Studies.

Grover, R., & Srinivasan, V. (1987). A simultaneous approach to market segmentation and market structuring. *Journal of Marketing Research*, *24*(2), 139–153.

Haenlein, M., & Kaplan, A. M. (2012). The impact of unprofitable customer abandonment on current customers' exit, voice, and loyalty intentions: An empirical analysis. *The Journal of Services Marketing*, *26*(6), 458–470.

Heskett, J. L., & Hallowell, R. (1997). *Southwest Airlines – 1993 (A)*. Cambridge, MA: Harvard Business School Publishing.

Hirschman, A. O. (1970). *Exit, voice, and loyalty responses to decline in firms, organizations, and states*. Cambridge MA: Harvard University Press.

Interbrand. (2017). Best global brands. Interbrand. www.interbrand.com/best-brands/best-global-brands/2017/ranking/

Isaacson, W. (2011). *Steve Jobs* (1st ed.). New York: Simon & Schuster.

Isaacson, W. (2012). The real leadership lessons of Steve Jobs. *Harvard Business Review*, 90, 92–102.

Jia, C., Wang, Y., & Xiong, W. (2017). Market segmentation and differential reactions of local and foreign investors to analyst recommendations. *The Review of Financial Studies*, *30*(9), 2972–3008.

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, *57*(1), 1–22.

Keller, K. L. (2001). Building customer-based brand equity. *Marketing Management*, *10*(2), 14–19.

Kim, M. G., Lee, C. H., & Mattila, A. S. (2014). Determinants of customer complaint behavior in a restaurant context: The role of culture, price level, and customer loyalty. *Journal of Hospitality Marketing and Management*, *23*(8), 885–906.

Kohli, A. K., Jaworski, B. J., & Kumar, A. (1993). MARKOR: A measure of market orientation. *Journal of Marketing Research*, *30*(4), 467–477.

Kotler, P., & Levy, S. J. (1969). Broadening the concept of marketing. *Journal of Marketing*, *33*(1), 10–15.

Luce, M. F., Bettman, J. R., & Payne, J. W. (2001). An integrated model of trade-off difficulty and consumer choice. *Journal of Consumer Research*, 1, 11–35.

Das Deutsche Markenranking. (2018). Best product brands Germany 2018. Ranking the Brands.com. www.rankingthebrands.com/The-Brand-Rankings.aspx?rankingID=27&year=1205

McGirt, E. (2017). These fifty companies are changing the world. http://fortune.com/2017/09/07/these-50-companies-are-changing-the-world/

Mittal, B., & Lassar, W. M. (1996). The role of personalization in service encounters. *Journal of Retailing*, *72*(1), 95–109.

Mochon, D., Norton, M. I., & Ariely, D. (2008). Getting off the hedonic treadmill, one step at a time: The impact of regular religious practice and exercise on well-being. *Journal of Economic Psychology*, *29*(5), 632–642.

Myers, D. G. (1992). *The pursuit of happiness: Discovering the pathway to fulfillment, wellbeing and enduring personal joy* (1st ed.). New York, NY: Avon Books.

Obilo, O. O., & Alford, B. L. (2018). Market segmentation via attitudinal functions: A multi-method approach. *Qualitative Market Research*, *21*(1), 63–81.

Park, C. W., MacInnis, D. J., Priester, J., Eisingerich, A. B., & Iacobucci, D. (2010). Brand attachment and brand attitude strength: Conceptual and empirical differentiation of two critical brand equity drivers. *Journal of Marketing*, *74*(6), 1–17.

Payne, A., Frow, P., & Eggert, A. (2017). The customer value proposition: Evolution, development, and application in marketing. *Journal of the Academy of Marketing Science*, *45*(4), 467–489.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Petersen, S., & Wathieu, L. (2012). *Tchibo ideas: Leveraging the creativity of customers*. Berlin: European School of Management and Technology.

Rangan, V. K. (2009). *Aravind eye hospital, Madurai, India: In service for sight*. Cambridge, MA: Harvard Business School Publishing.

Roggeveen, A. L., Tsiros, M., & Grewal, D. (2012). Understanding the co-creation effect: When does collaborating with customers provide a lift to service recovery? *Journal of the Academy of Marketing Science*, *40*(6), 771–790.

Rosenberg, T. (2013). A hospital network with a vision. *The New York Times*, January 16.

Rusbult, C. E., Farrell, D., Rogers, G., & Mainous, A. G., III. (1988). Impact of exchange variables on exit, voice, loyalty, and neglect: An integrative model of responses to declining job satisfaction. *Academy of Management Journal*, *31*(3), 599–627.

Rusbult, C. E., Johnson, D. J., & Morrow, G. D. (1986). Predicting satisfaction and commitment in adult romantic involvements: An assessment of the generalizability of the investment model. *Social Psychology Quarterly*, *49*(1), 81–89.

Sattayaraksa, T., & Boon-itt, S. (2016). CEO transformational leadership and the new product development process. *Leadership and Organization Development Journal*, *37*(6), 730–749.

Saxe, R., & Weitz, B. A. (1982). The SOCO scale: A measure of the customer orientation of salespeople. *Journal of Marketing Research*, *19*(3), 343–351.

Sela, A., Berger, J., & Liu, W. (2009). Variety, vice, and virtue: How assortment size influences option choice. *Journal of Consumer Research*, *35*(6), 941–951.

Singh, S. S., & Jain, D. C. (2010). Measuring customers lifetime value: Models and analysis. In N. K. Malhotra (Ed.), *Review of marketing research* (Vol. 6, pp. 37–62). Armonk, NY: M.E. Sharpe.

Skålén, P., Gummerus, J., von Koskull, C., & Magnusson, P. R. (2015). Exploring value propositions and service innovation: A service-dominant logic study. *Journal of the Academy of Marketing Science*, *43*(2), 137–158.

Statista (2019). U.S. airlines ranked by rate of consumer complaints in 2018 (per 100,000 enplanements). Statista (website). www.statista.com/statistics/539278/customer-complaints-by-us-airlines/

Sternberg, R. J. (1984). A contextualist view of the nature of intelligence. *International Journal of Psychology*, *19*(4–5), 307–334.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge: Cambridge University Press.

Sternberg, R. J. (1998). A balance theory of wisdom. *Review of General Psychology*, *2*(4), 347–365.

Sternberg, R. J. (2013). The WICS model of leadership. In M. G. Rumsey & M. G. Rumsey (Eds.), *The Oxford handbook of leadership*. (pp. 47–62). New York: Oxford University Press.

Stone, B. (2013). *The everything store: Jeff Bezos and the age of Amazon* (1st ed.). New York: Little, Brown and Company.

Thomson, M., MacInnis, D. J., & Park, C. W. (2005). The ties that bind: Measuring the strength of consumers' emotional attachments to brands. *Journal of Consumer Psychology*, *15*(1), 77–91.

Tracy, J. L., Cheng, J. T., Robins, R. W., & Trzesniewski, K. H. (2009). Authentic and hubristic pride: The affective core of self-esteem and narcissism. *Self and Identity*, *8*(2–3), 196–213.

van Beek, J., Antonides, G., & Handgraaf, M. J. J. (2013). Eat now, exercise later: The relation between consideration of immediate and future consequences and healthy behavior. *Personality and Individual Differences*, *54*(6), 785–791.

Walumbwa, F. O., Hartnell, C. A., & Oke, A. (2010). Servant leadership, procedural justice climate, service climate, employee attitudes, and organizational citizenship behavior: A cross-level investigation. *Journal of Applied Psychology*, *95*(3), 517.

Weitz, B. A. (1978). Relationship between salesperson performance and understanding of customer decision making. *Journal of Marketing Research*, *15*(4), 501–516.

Weitz, B. A., Sujan, H., & Sujan, M. (1986). Knowledge, motivation, and adaptive behavior: A framework for improving selling effectiveness. *Journal of Marketing*, *50*(4), 174–191.

Wells, W. D. (1975). Psychographics: A critical review. *Journal of Marketing Research*, *12*(2), 196–213.

Wilkowski, B. M., & Ferguson, E. L. (2016). The steps that can take us miles: Examining the short-term dynamics of long-term daily goal pursuit. *Journal of Experimental Psychology: General*, *145*(4), 516–529.

Wind, J., & Mahajan, V. (1997). Issues and opportunities in new product development: An introduction to the special issue. *Journal of Marketing Research*, *34*(1), 1–12.

Yan, J., Tian, K., Heravi, S., & Morgan, P. (2017). The vices and virtues of consumption choices: Price promotion and consumer decision making. *Marketing Letters*, *28*(3), 461–475.

PART VII

Intelligence and Its Role in Society

37 Intelligence in Worldwide Perspective

A Twenty-First-Century Update

Weihua Niu

For thousands of years of human history, understanding the nature of "intelligence" has been a quest of the utmost importance, attracting many sages and intellects around the world. In ancient Greek culture, Plato (428/427–348/347 BCE) expressed his belief that human beings are born with different levels of intelligence, strength, and courage. In his opinion, those who were not overly bright, strong, or brave were suited to various trades such as farming, blacksmithing, and building, whereas those who were somewhat bright, strong, and especially courageous were suited to defensive and policing professions. Those who were extraordinarily intelligent, virtuous, and brave were suited to run the state itself as part of the aristocracy, a Greek word for "rule by the best" (Plato, 1992).

In ancient Chinese culture, Confucius (551–479 BCE) presented a different view of intelligence than Plato. Using the words "intelligence" (智) and "knowledge" (知) almost interchangeably, Confucius believed that people were varied in their levels of intelligence by how knowledge was acquired and utilized. In the *Doctrine of the Mean*, Confucius (2010) said,

> Some are born with the knowledge of those duties; some know them by study; and some acquire the knowledge after a painful feeling of their ignorance. But the knowledge being possessed, it comes to the same thing. Some practice them with a natural ease; some from a desire for their advantages; and some by strenuous effort. But the achievement being made, it comes to the same thing. (p. 9)

Although acknowledging that some people are born with knowledge or intelligence, Confucius believes that these people are extremely rare and truly exceptional. Confucius would not consider even himself to be one of them. Therefore, he emphasized the importance of learning and self-cultivation in acquiring knowledge or intelligence.

This discrepancy in philosophical views is one of the first pieces of evidence that people from different cultures view intelligence differently. To Plato, intelligence is something that one is born with, whereas, to Confucius, intelligence is something that one can earn and accumulate throughout one's life. Both Plato and Confucius have had a profound impact on the development of great civilizations in the world and their views on intelligence also deeply affect how people across the world currently perceive and attempt to measure intelligence.

The author would like to acknowledge the contribution of Jillian Brass who coauthored the same chapter, which appeared in the first edition of *The Cambridge Handbook of Intelligence*.

Many scholarly works examine the role of culture in understanding and measuring intelligence, including several comprehensive reviews (e.g., Saklofske et al., 2015; Serpell, 2000; Sternberg 2004; Sternberg & Grigorenko, 2006). This chapter first summarizes some of the main points and findings from studies on implicit theories of intelligence, adding new evidence from recent studies, particularly originating in East Asia. It then reviews history and new developments in measures of intelligence in different countries from different continents, followed by a summary of worldwide studies on social intelligence, emotional intelligence, and cultural intelligence. It concludes with a discussion on how culture affects people's conception of intelligence and how globalization in the twenty-first century affects the development of social aspects of intelligence.

Implicit Theories of Intelligence Across Different Cultures

What is intelligence? Many psychologists around the world have proposed theories to elucidate this question. There are probably as many definitions of intelligence as there are experts who study it. As noted by Detterman (1986), there is no definitive definition of intelligence; the concept has evolved and will continue to evolve over time. Many researchers also recognize that intelligence cannot be understood outside a cultural context (Greenfield, 1997; Sternberg 2004). People from different cultures may perceive intelligence differently, depending on what is considered to be important in that culture.

One important approach to studying people's conceptions of intelligence is through investigating the cultural prototype of an intelligent person. This approach is relatively straightforward: Lay people are asked to list characteristics associated with the term "intelligence" or "an intelligent person." Many researchers credit Neisser (1979) for his acknowledgment of the importance of this approach. Sternberg coined the term "implicit theories of intelligence" to describe this approach, in comparison to the other type of approach, based on experts' explicit theories of intelligence. Sternberg and his colleagues conducted a series of empirical studies in the 1980s (Sternberg, 1985; Sternberg et al., 1981), studying laypeople's implicit theories of intelligence. These studies generated wide interest around the world in investigating definitions of intelligence within each specific culture.

Earlier studies on implicit theories of intelligence were conducted in the United States. Bruner, Shapiro, and Tagiuri (1958) asked laypeople to identify character traits that are associated with intelligent people and found that the traits include clever, deliberate, efficient, and energetic. People tended *not* to associate social aspects such as "dishonest," "apathetic," and "unreliable" with intelligence. Similarly, Neisser (1979) also found that "the ability to think logically," "verbal fluency," "wide general knowledge and common sense," "openness to experiences," and "sensitivity to one's own limitations" were important in the conception of intelligence. In other words, both of these earlier studies indicated that intelligence was mostly associated with cognitive abilities and some personality traits.

Sternberg and colleagues (1981) had the general public list behaviors that characterize intelligence, academic intelligence, everyday intelligence, and unintelligence. They then asked another group of people from varying backgrounds to indicate the importance and characteristics of each behavior associated with their ideal concepts of intelligence, academic intelligence, and everyday intelligence. Findings from this study suggested that intelligence consisted of at least three common components: problem-solving abilities, verbal abilities, and social competence. Importantly, such views on the core components of intelligence were found to be shared by both laypeople and experts who study intelligence. The difference between laypeople's and experts' evaluations of intelligence is that laypeople did not consider motivation to be an important ingredient of "academic" intelligence, whereas the experts did believe it to be so. Additionally, the laypeople placed somewhat greater emphasis on practical intelligence than did the experts.

However, this conception of intelligence is not consistently shared by people from other parts of world, especially Asia and Africa where social and emotional competence and even moral characters are important in people's implicit theories of intelligence.

Asia

In Asia, many studies have been conducted to investigate people's implicit theories of intelligence using samples from mainland China, Taiwan, Hong Kong, Japan, Korea, India, and Malaysia.

Chinese Culture. The literal translation of the Chinese phrase for "intelligence" (聰明) is "to have sharp hearing and clear vision," or "to have a clear understanding (of a situation)." The phrase itself reflects the Chinese view of intelligence, which has historically emphasized the correctness of one's perception and comprehension. The implied meaning is that, with a clear perception and understanding of a situation, one can act properly.

Overall, Chinese hold similar views to Western conceptions of intelligence, which are a strong emphasis on cognitive abilities such as curiosity, knowledge, memory, imagination, and problem-solving as well as reasoning skills (Cai & Jiang, 1995; Wan, Li, & Jing, 1997; Zhang & Wu, 1994). However, four unique characteristics are included in Chinese conceptions of intelligence, which are somewhat different than Western implicit theories of intelligence. First, Chinese people value memory skills more than their Western counterparts, a possible result of different instructional practices and values in Chinese and Western schools (Chen 1994; Fang & Keats, 1987). Second, there is a greater emphasis on diligence and malleability of human potential among Chinese students, which explains why Chinese students tend to hold a higher level of achievement motivation in comparison to their Western counterparts (Bai, Liu, & Hu, 2007; Chen, Lee, & Stevenson, 1996; Fwu et al., 2017; Hau & Ho, 2012). In other words, the Confucian view of intelligence quoted at the very beginning of the chapter has a profound influence on how contemporary Chinese people view intelligence, which may suggest that, overall, Chinese hold more incremental theories of intelligence in comparison to their Western counterparts. Third, social aspects of intelligence are emphasized more in the Chinese implicit theories of intelligence. For examples, studies have shown that qualities such as benevolence, filial piety, retaining appropriate behaviors and conversation in a social context, and seeking harmony between humanity and nature, health and longevity, as well as action through inaction are included in Chinese notions of intelligence (Chen & Wong, 2014; Yang & Sternberg, 1997a). Lastly, Chinese people also place more value on one's knowledge to themselves or intrapersonal skills, such as knowing how to express one's self appropriately in a social context, a high level of self-knowledge, and being perceptive and responsive to changes in immediate circumstances (Yang & Sternberg, 1997b). All these four characters suggest that cultural values, societal, family, and school practice can affect lay people's view of intelligence.

Japan. Also influenced by Confucianism, Japanese conceptions of intelligence seem to also emphasize social competence, diligence, and modesty. Moreover, Japanese children were found to emphasize more on classroom behavior such as "can remember well what has been learned before," "having his or her own way of thinking," and "good in mathematics" but, as they grew older, Japanese students focused more on organization, management, planning, and social factors such as responsibility and sociability in their conception of intelligence. Interestingly, all age groups of Japanese students disassociated arrogance and selfishness from intelligence. Similar to Chinese students, Japanese students also consistently rated memory and good concentration skills as being important to the concept of intelligence (Azuma & Kashiwagi, 1987; Furnham & Mkhize, 2004; Ueda, 1989).

When asked the question, "How can people become more intelligent?" Japanese students across all age groups placed great emphasis on effort-related descriptions, such as "engaging in everything seriously," "making an effort (try harder)," and "trying everything without giving up." In other words, Japanese students believe working hard makes people more intelligent.

Emphasizing the importance of effort in conceptions of intelligence and related concepts in Japanese culture was also found in many cross-cultural studies examining attribution theories. Overall, Japanese students placed greater emphasis on effort whereas American students placed greater emphasis on one's innate ability (for a review, see Holloway, 1988).

Korea. Interestingly, cultural influence on people's implicit theories of intelligence was not found to be salient in studies from South Korea. Lim, Plucker, and Im (2002) replicated Sternberg and colleagues' (1981) study using a sample of both Korean college students and the general public, who were approached at the railway station. They found that Korean implicit theories of intelligence were only slightly different from those of Americans. Similar to findings from studies of the Chinese and Japanese, they also found that Korean participants emphasized social competence in their conception of intelligence. However, when Korean participants were asked to evaluate other people's intelligence, they emphasized problem-solving ability over all other factors, an evaluation that shows much similarity with the views of American counterparts.

India. Although also geographically in Asia, Indian societies represent a different culture from the East Asian culture. They, too, have a long history of cultural tradition that still deeply affects the lives of modern Indian people and their ways of thinking. In studying Indian implicit theories of intelligence, researchers have found that the Indian culture emphasizes greatly the social aspects of intelligence. Only one-third of the attributes referred to the cognitive domain. The rest are to social and emotional aspects of intelligence. In other words, according to the Indian culture, an intelligent person knows how to speak and behave in a context-sensitive manner and is able to value options and make wise generalizations and discriminations. Effective communication in India refers to someone who speaks only when necessary, who can make one's mind clear using minimum words, and who uses hidden meaning in speech and remains focused on the problem under discussion (Srivistava & Misra, 2001).

Malaysia. Malaysia represents another type of Asian culture, in which Islam is the official and most widespread religion. Gill and Keats (1980) studied Malay University students' views of intellectual competence, in comparison with those of Australians. They found that, whereas Australian students rated academic skills more highly and stressed the ability to adapt to new events, Malays placed great emphasis on social and practical skills along with speed and creativity.

Swami et al. (2008) asked 235 college students in Malaysia, along with 347 college students from Britain and 137 college students from the United States, to indicate their agreement with thirty statements about what intelligence is, the source and stability of between group differences in intelligence, and the practical relevance as well as social implications of intelligence. Most of the statements were derived from a summary of a psychological study asking fifty Western experts in intelligence and applied fields about their views of intelligence. Similar to the findings of Gill and Keats (1980), this study also demonstrated that Malaysians place more emphasis than do their Western counterparts on social competence and the practical aspects of intelligence.

Africa

Not only people from Asia (typically viewed as the East) view intelligence differently from the West; people from Africa also have different conceptions than Westerners. According to Sternberg (2004), African conceptions are more consistent with Eastern than with Western views. In a review examining the relationship between personality and intelligence in a cultural context, Ruzgis and Grigorenko (1994) argued that the implicit theories of Africans revolve largely around skills that help to facilitate harmonious and stable intergroup relationships. Such a view is supported by many empirical studies from Africa.

Using semantic-differential scales, Wober (1974) studied conceptions of intelligence among members of different tribes in Uganda as well as within various subgroups of the tribes. In a result surprising to many Westerners, it was revealed that traditional Ugandans associated intelligence with slowness, gradualness, and taking one's time, whereas Western-educated Ugandans and Indians in Uganda associated it with speed. There is also a difference in conceptions of intelligence both within and between tribes. Beganda tribespeople associated intelligence with words such as persistent and hard-working, whereas the Batoro thought of it as soft, obedient, and yielding.

Serpell (1974) asked Chewa adults in rural eastern Zambia to rate village children on how well they could perform tasks requiring adaptation in the everyday world (practical and social intelligence). He found that the ratings did not relate to children's cognitive IQ test scores, which had been assessed by the investigators. The results suggested that Chewa criteria for judgments of intelligence were not the same as Western notions of intelligence. In many places in Africa, the games people play, such as "kala," encourage the development of numerical ability (Gardner, 1983). In a series of experimental studies, Cole, Gay, and Glick (1967) found that Kpelle adults in Liberia succeeded far more than American adults in estimating the quantity of a group of objects.

Grigorenko and colleagues (2001) conducted a study investigating the implicit theories of intelligence in a Kenyan village. They found that, in rural Kenya, intelligence consists of four different concepts: knowledge and skills, respect, comprehension of how to handle real-life problems, and taking the initiative. Of these four skills, only the first relates to cognitive skills, while the other three fall into the social domain.

South America and East Europe

Implicit theories in South America and Eastern Europe fall somewhere in between the views of the East and the West. In Chile, for example, Garcia-Cepero and McCoach (2009) surveyed 372 school teachers and college professors with regard to their implicit theories of intelligence. Using both Sternberg's theory of successful intelligence and Gardner's theory of multiple intelligences as their framework to design questionnaires, they asked participants whether they agreed with views relating to these two theories. They found that Chilean educators acknowledge the importance of practical, analytical, and creative attributes in their prototypes of an intelligence person. However, participants were fairly neutral about whether interpersonal and intrapersonal attributes characterized intelligent people.

In Eastern Europe, Kopic, Vranic, and Zarevski (2009) asked 330 eighth-graders from Croatia to list attributes associated with an intelligent person, which revealed five meaningful factors attributed to an intelligent person: (1) cognitive abilities, (2) practical intelligence, (3) interpersonal characteristics, (4) motivation, and (5) "academic" intelligence and verbal abilities. All five characteristics had been included in previous studies using Western samples (such as studies of Sternberg and colleagues); however, the importance of interpersonal characteristics and practical intelligence seem to be recognized more in the Croatian culture than in Western culture.

In summary, studies of implicit theories of intelligence in different parts of the world suggest intelligence may not mean the same thing in different cultures. In Western Europe and North America, where many modern intelligence theories and measurements have been generated, intelligence is largely related to one's cognitive abilities, whereas the rest of the world seems to view other aspects of intelligence such as social acuity, emotional intelligence, and morality to be more important than did their Western counterparts. Even within the domain of cognitive functioning, some areas are emphasized more in some cultures (such as memory skills in China) or may mean different things (such as the meaning of sensitivity to information having much more comprehensive implications in Indian culture). However, this does not mean that social, emotional, and moral components of intelligence are entirely excluded from the Western notion of intelligence; nor does it mean that cognitive functioning is valued less in other parts of the world. In fact, despite the differences in components of intelligence, people around the world share some core views in their conceptions of intelligence, including cognitive competence (both verbal and nonverbal) and social-emotional competence. Most attempts at the measurement of intelligence have been focused on the former (cognitive competence), even though there is an increasing amount of effort in recent years to develop scales to measure the latter. The next section primarily focuses on examining measures of intelligence in the former area (cognitive competence).

Measurements of Intelligence Around the World

As noted in the previous section of the chapter, ideas about intelligence vary across cultures and sometimes even within cultures. Just as definitions of intelligence exist throughout the world, instruments used to try to measure and quantify intelligence are used worldwide. However, unlike findings from implicit theories of intelligence, people around the world tend to adopt the same or similar measurements of intelligence (Oakland, Douglas, & Kane, 2016). In other words, many different cultures are actually measuring the same constructs despite differences in ideology and conceptions regarding intelligence.

The process of translating tests is never simple. Van De Vijver (2003) argues that, when tests are translated into other languages, there are several different routes test constructors can take. An application refers to a close translation of the original test, while an adaptation makes changes to the instrument (for instance, substituting words for more appropriate ones or task materials for ones more familiar to the target audience) to emphasize measuring the same underlying constructs. Oftentimes, a literal translation will be inappropriate in a different language or culture. Assembly refers to the construction of an entirely new instrument. Test constructors must decide how an instrument would best fit the population of their country and work accordingly, trying as much as possible to reduce cultural bias stemming from the fact that a test was originally developed for use in a different culture.

One major question hotly debated by psychologists is whether intelligence tests should be measuring the same processes cross-culturally. Are the abilities and skills measured by intelligence tests equally relevant in all parts of the world? Are underlying cognitive processes valued in the same way in a small town in Africa and in Akron, Ohio? For that matter, do people think in the same ways in these different areas? One school of thought is that tests designed by a certain culture primarily measure skills and abilities most valued by that culture that are not as applicable elsewhere. On the other hand, the globalization of tests such as the Wechsler Intelligence Scale for Children (WISC) comes with a certain implication that it is appropriate and even useful to measure the same processes valued in the United States in a multitude of other geographic areas with different values and cultures. Although there is no apparent resolution to bridge these viewpoints, it is evident that, just as with definitions of intelligence worldwide, there will be some discrepancies as well as some similarities in what different cultures want to measure in quantifying or even qualifying intelligence.

As illustrated in the previous section of this chapter, most cultural differences in people's implicit theories of intelligence reflect their cultural value systems. One example lies in Asian cultures seeing effort as being a part of intelligence. However, most intelligence tests developed in the United States and Europe do not measure this factor, as these cultures tend to see intelligence as inherent or based on ability than rather than as a result of hard work.

While the philosophical questions of the degree to which intelligence tests should be specific to the culture in which they are used continue to be studied, it is clear that certain tests such as the Wechsler tests, the Stanford-Binet, and the Kaufman Assessment Battery for Children (KABC) have been exported and are now used in many countries around the world (Lautrey & de Ribaupierre, 2004; Oakland et al., 2016; Sato et al., 2004). Using the same test cross-culturally often leads to the temptation to make comparisons between intelligence test scores in different geographic regions. How intelligent we are relative to other cultures and to people from different geographical locations has become a question of great interest and, at times, of national importance. Years ago, worries about falling behind relative to other countries sparked renewed interest in programs such as gifted education. In modern times, we have the instruments necessary to screen and document intelligence test scores of populations. However, there are major problems with making cross-cultural comparisons of intelligence, the largest and most important of which is inaccuracy.

The validity of making comparisons across different tests, or even the same test adapted and normed for a different population, is questionable. Cross-cultural comparisons that look specifically at numbers are inherently based on the idea that, when we are measuring intelligence, we are all measuring the same thing. The problem is that, more often than not, what we are measuring is quite different.

Even when using the same test major differences can exist in the equivalence of the test across cultures. WISC in its fifth edition in the United States (the WISC-V), where it originated, and has been adapted and re-normed all over the world (Georgas et al., 2003; Oakland, 2009; Oakland et al., 2016). In a survey of European countries, Muñiz and colleagues (2001) asked what the most frequently used psychological test were in each country; WISC and the Wechsler Adult Intelligence Scale (WAIS) were in the top ten for each country surveyed. Muñiz and colleagues (1999) also found that the Wechsler scales rank in the top-ten tests used in Spanish- and Portuguese-speaking countries, including Spain, Portugal, and fourteen countries in Latin America. In a study of the most frequently used psychological tests in schools from sixty-four countries, Oakland and colleagues (2016) again found that WISC was ranked as number one, followed by Raven's Progressive Matrices and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI). The Kaufman Assessment Battery for Children (KABC), Stanford-Binet, and WAIS were also included in the top ten. These studies provide evidence as to the popularity of the US-originated Intelligence instruments such as Wechsler and Raven's across countries, languages, and continents. Many different countries now have their own versions of the WISC, though not all of them have been readapted based on the most recent US edition.

Although many countries are using the same tests, significant issues in translation and adaptation, as well as appropriateness to a new population, affect whether cross-cultural comparisons of scores on this instrument reflect true cross-cultural differences. How relevant and accurate cross-cultural comparisons are seems to depend on the factor being examined.

Psychologists involved in cross-cultural analysis of the WISC have noted that the performance subtests in particular are easily adaptable to other cultures, as the skills they measure – analysis of visual material, pattern completion, and visual-motor integration, for example – are practical across cultures and have a universal feel to them; there are probably very few cultures in which abilities like visual-motor integration or visual analysis, which are generally adaptive skills from an evolutionary standpoint, are irrelevant (Georgas et al., 2003).

However, looking at the verbal subtests opens a host of larger problems. Evidence suggests that verbal thinking is not necessarily the same cross-culturally and therefore a test measuring verbal abilities in the United States may not be as relevant elsewhere. For instance, a study by Peng and Nisbett (1999) suggested that people in China think differently than do those in the United States. When Chinese people were presented with a seemingly contradictory statement, they tended to try to resolve the two sides and find a compromise between them, which the authors termed "dialectical thinking." When presented with the same contradiction, people in the United States tended to polarize their views by picking the half of the apparent contradiction they felt was more accurate and rejecting the other half, a process termed "differentiation of thinking" by the authors. These results seemed to suggest that cognitive processes are different between Chinese individuals and those from the United States.

Problems with subtest translation are not limited to underlying conceptual issues – they also involve the more practical elements of test adaptation. One such problem is the vocabulary subtest of the WISC, which asks children to define words. In adapting this subtest, many countries have found that not all the vocabulary words are directly translatable and that, if they are, the same word in a different language might not have an equivalent "difficulty" level – the word might be more or less common than its English counterpart, which in turn changes the difficulty of the entire subtest. Substituting a more appropriate word for equivalent difficulty would change the content of the subtest; both solutions compromise the integrity of cross-cultural comparisons of ability on this task (Georgas et al., 2003).

As an example, Beller and Gafni (1995) note that, when a test originally written in Hebrew was translated into Russian, Russians did more poorly on a specific analogy where the answer involved understanding the relationship between a dictionary and a definition. The authors noted that, in the Russian language, dictionaries are used to translate, not to define, which led to Russians not being able to recognize this relationship as the correct response for "Telephone book: telephone number." However, Russians performed better on a different analogy, "plow: furrows," as "furrows" in its appropriate translation is a more common word in Russian than it is in Hebrew (Beller & Gafni, 1995).

Another task on the WISC, designed to measure working memory, requires children to repeat a series of numbers first forward and then, later, backward. This task was designed in the United States, where the numbers used in the task have fairly distinct, one-syllable names. Countries that do not have similar ways of naming their numbers might have difficulty constructing an equivalent of this subtest that would measure precisely the same process (Georgas et al., 2003; Kwak, 2003).

Another important issue in cross-cultural comparisons is the familiarity that the test audience has with both the modality of an intelligence test – the methods of administration and materials used – and the information or experiential bases necessary to succeed. While doing research with children from Tanzania, Sternberg and colleagues noted that a short intervention could raise test scores. This result suggested that familiarity and training play key roles in scores and that giving an unfamiliar test to a group of children is likely not to be an accurate measure of cognitive ability alone (Sternberg et al., 2002).

Serpell and Jere-Folotiya (2008) noted that, in Zambia, pencils and paper are rare play things for children before entering school. They found in studies that children from England performed superiorly to children from Zambia in a pencil-and-paper task. However, when the same task was presented in the media of small twisted wires, something with which Zambian children are familiar and English children less so, the performance of the Zambian children was superior. This study suggested that the way in which a task is presented affects performance on the task, depending on familiarity with and training in the presented media.

Another issue to take into consideration is whether children who have access to schooling will do better on cognitive tests, which would suggest that pure, untrained cognitive ability is not the underlying construct being measured. Even if children have a history of schooling, Sternberg and colleagues (2002) point out that children in some parts of Africa do not have equal opportunity to take advantage of their schooling, as the environment in which they are schooled, in terms of stressors and opportunities, is not comparable to school environments in the United States or Western Europe.

In the first edition of *The Cambridge Handbook of Intelligence*, we provided a summary of intelligence developed in different parts of the world (Niu & Brass, 2011). The results demonstrated that, although tests around the world contain different tasks than some of the Western tests we are familiar with in this country, very few seem to be based on entirely different models of intelligence. Even countries that incorporate different ideas such as effort or social responsibility into their conceptualizations of intelligence do not frequently incorporate these ideas into tests used to measure intelligence in their citizens. Generally, countries that have constructed their own tests also rely on translations or adaptations of instruments such as the Wechsler scales, Raven's Progressive Matrices, KABC, and Stanford-Binet. While these instruments have proven to be reliable and valid, they do not always match the values of the cultures in which they are being used. Although intelligence is defined differently throughout the world, the testing of intelligence suggests that what we are content to measure as intelligence may remain far more consistent than our definitions across cultures.

Beyond General Intelligence

A critical feature of the twenty-first century is globalization, with more and more people traveling around the world, studying, working together, and living in a multicultural environment. Even without the actual traveling, the fast-growing usage of the Internet and digital resources makes global communication and collaboration much more accessible than ever before. To some, such changes mean new opportunities; yet, to others, they can be completely overwhelmed. To adapt and to succeed in this increasingly more interconnected and interdependent world, one needs not just general intelligence (the cognitive aspect) but also the ability to tune in and manage oneself and other's psychological state (the social aspect). Some scholars believe that the social aspect of intelligence may be as important, if not more important, than the cognitive aspect (Sternberg & Grigorenko, 2006).

There are an increasing number of studies worldwide to examine the noncognitive facet of intelligence such as social intelligence, emotional intelligence, and cultural intelligences.

Social Intelligence

Social intelligence (SI) was first brought up in the 1930s. It was initially defined as the ability to understand and manage people (Thorndike & Stein, 1937) and later described as the ability to interact effectively with others, which includes knowledge about social situations and having the ability to perceive and interpret such situations accurately, leading to successfully behavior in that situation (Crowne, 2009). SI has become more and more focused on two specific abilities, the interpersonal (the ability to read other people's moods, motives, and other mental states) and the intrapersonal (the ability to access one's own feelings and to draw on them to guide behavior), popularized along the concept of multiple intelligences (Gardner, 1983, 1998, 2002).

Being focused on the social aspect of intelligence, SI is believed to be bonded by culture. Willmann, Feldt, and Amelang (1997) conducted a study examining prototypical behavior patterns of SI among Chinese and German participants. The results showed that, when rating frequencies of a list of social behaviors in SI, Chinese participants place more emphasis on the ability to retain harmonious relations with others and on interdependence, confirming and fulfilling one's expected roles, and acting for the well-being of the entire society, reflecting the classical traditions and ideals of Confucianism. Such an emphasis on interdependence with one another in judging SI is missing in the German sample. When directly comparing the responses to the importance of maintaining socially desirable behaviors and social engagement in deciding one's social intelligence, German participants rated these behavior patterns much lower than their Chinese counterparts, which demonstrated that the construct of SI is culturally dependent. Additionally, their results also showed that such culturally dependent characteristics in people's conception of SI are exhibited more strongly among older and female participants in comparison to younger and male participants.

Emotional Intelligence

Emotional intelligence (EI) is defined as people's ability to recognize their own emotions and those of others, discern between different feelings and label them appropriately, use emotional information to guide thinking and behavior, and manage and/or adjust emotions to adapt to environments or achieve one's goal(s) (Salovey & Pizarro, 2003).

Since the seminal works on the concept of EI were first published in the early 1990s (Mayer, DiPaolo, & Salovey, 1990; Salovey & Mayer, 1990), the concept has spread quickly to every corner of the world. It has been widely embraced by, and resonated with, people from different cultural and faith backgrounds as an essential factor, arguably more important than general intelligence, to predict people's success in their everyday lives. EI is especially crucial in the context of globalization in which different cultural values meet, are intermingled, or collide.

Gabel-Shemueli, Dolan, and Cerdin (2005) studied the influence of EI on global managers' cultural adjustment and job performance. They found that interpersonal emotional components such as empathy, social responsibility, and social relations predicted participants' cultural adjustment, which in turn influenced their job satisfaction and job performance, especially when jobs require international collaborations.

Similarly, Gunkel, Schlaegel, and Teras (2016) surveyed more than 1,500 university students from eighty-three nations in ten different cultural clusters, such as Arab, African, Anglo (US and UK), Confucian Asia, East Europe, Far East, Germanic, Latin America, Latin Europe, and Nordic. They sought to examine the relationship between cultural values, EI, and conflict handling styles. The results demonstrated that EI partially mediates the influence of cultural values on conflicting handling style preference, which subsequently influences the economic outcomes that are related to interpersonal conflicts in the business context. In other words, EI is demonstrated as vital in solving complex social problems in global settings.

Like the construct of SI, EI also shows culturally relevant characteristics, which is consistent with the worldview theory regarding collectivism versus individualism (Ekermans & Privaatsak, 2009). However, cultural differences are subtle (Alhashemi, 2014).

Cultural Intelligence

A new concept, cultural intelligence (CI), emerged in the twenty-first century and has gained worldwide attention in the past decade. It is defined as people's ability to adjust and function effectively in situations characterized by cultural diversity (Ang & Van Dyne, 2008). It is widely accepted as a multidimensional construct consisting of four dimensions: metacognitive, cognitive, motivational, and behavioral.

CI and EI overlap yet they are distinct (Crowne, 2009, 2013). Having a high level of CI does not require someone to have a high level of EI (Earley & Mosakowski, 2004). Likewise, an individual with high-level EI can be low in CI.

Different from personality traits, CI is state-like and malleable. It can be predicted by personality traits such as openness to experience (Ang, Van Dyne, & Koh, 2006; Depaula et al., 2016; Lie, Suyasa, & Wijaya, 2016; Moon, Choi, & Jung, 2012, 2013) and agreeableness (Harrison, 2012). It can also be acquired and developed through cross-cultural contact. Intercultural experience, such as traveling abroad for education, employment, and vacation, as well as immigration, is one of the most frequently examined predictors of CI and proven to be positively associated to its development (Crowne, 2008; Fang, Schei, & Selart, 2018; Harrison, 2012).

Besides direct intercultural experience, training in cross-cultural contact through simulation games, role-play, and classroom training, and even completing an online course with a foreign partner or participating in virtual multicultural team projects, can also improve overall CI (Alexandra, 2018; Ko, Boswell, & Yoon, 2015; MacNab & Worthley, 2012; Rosenblatt, Worthley, & MacNab, 2013).

There are many benefits to obtaining a higher level of cultural intelligence (Fang et al., 2018). CI can facilitate in group acceptance for newcomers, help in developing a transformational leadership style, which emphasizes offering followers a vision and inspiring them by acting as a role model, and lead to improved job performance. It can also help minimize the experience of culture shock, facilitate knowledge sharing, and provide an advantage for learning a new language, as well as improving leadership skills, creativity, and innovation (Kim & Van Dyne, 2012; Korzilius, Bücker & Beerlage, 2017). Different aspects of CI, namely cognitive, metacognitive, motivational, and behavioral, moderate the effect of CI on creativity. For example, Chua and Ng (2017) found that there is an inverted U-shaped relationship between the cognitive aspect of CI and creativity, which suggested that, although cultural knowledge benefits creativity, too much knowledge has a detrimental effect because of cognitive overload and entrenchment. However, the detrimental effect of cognitive overload exists only when metacognitive CI is low. When metacognitive CI is high, there is a positive linear relationship between cognitive CI and creativity.

CI was found to have an indirect effect on psychological well-being, interpersonal effectiveness, and performance. Increased CI makes people happier when interacting with those from different cultures (Chen, Liu, & Portnoy, 2012). Overall, it enables people to look beyond their personal and cultural limitations (Earley, 2002) and to be more adaptive, happier, and more successful in cross-cultural settings (Fang et al., 2018).

With only a little more than a decade's history, CI has a long way to go to establish its status in the field of intelligence, whether as a type of intelligence or as a style of cognition. Nevertheless, it is a critical construct that deserves more attention to allow us to understand the nature of intelligence in today's world.

Conclusions

The major quest of this chapter is to investigate how people around the world perceive and measure intelligence. New evidence from the twenty-first century is added to reflect an updated view on intelligence from a worldwide perspective. Three conclusions can be drawn from a review of these studies.

Universality vs. Cultural Specificity

After reviewing studies on implicit theories of intelligence from selected cultures around the world, the overall picture that emerges is that intelligence is not a *universal* concept; rather, it is *culture-specific*. It is defined and perceived differently by people from different parts of the world and these difference are largely reflective of long-standing cultural traditions. As Greenfield (1998) observed, "cultures define intelligence by what is adaptive in their particular niche," reflecting the multifaceted nature of intelligence. Many contemporary experts on theories of intelligences have addressed this multidimensionality of intelligence (Gardner, 1993, 1995; Mayer, Salovey, & Caruso, 2000; Sternberg, 1997), discussing multiple intelligences (Gardner, 1993), successful intelligence (Sternberg, 1997), or simply an inclusion of SI, EI (Mayer, Salovey, & Caruso, 2000; Sternberg, 1997), and CI (Fang et al., 2018). In other words, most people would agree that there are many aspects of intelligence but what is emphasized depends on culture. For example, many studies have documented that the Western notion of intelligence places more emphasis on cognitive competencies such as attention, speed of learning, logical reasoning, and language comprehension as compared with other cultures (Sternberg et al., 1981). In other parts of the world, individuals are often evaluated by how sensitive they are to the social context as well as their possession of qualities such as chivalry, rectitude, and righteousness – cognitive competence only accounts for a small portion of concepts of intelligence. There is a greater emphasis on domains such as social competence, emotional competence, and competence in action. This distinction may reflect the cultural tradition of the West, where behaviors leading to control over the physical environment are highly valued (White, 1959). On the contrary, the majority of Asian and African societies are historically agricultural; therefore, maintaining a harmonious and stable intergroup relationship is more important in terms of survival and adaptation to the society.

Interestingly, although intelligence is perceived differently, similar measures of intelligence, especially cognitive aspects of intelligence, are widely adopted by people across different cultures. Many countries have constructed their own measures of intelligence to suit both their own purposes – such as school admissions or entry into professions –and their own values – such as information-processing tests in Germany. However, these measures often seem to be used in conjunction with measures imported from the United States such as the Wechsler scales, Stanford-Binet, and KABC, as well as Raven's Progressive Metrics. Therefore, while understandings of intelligence throughout the world are multifaceted, nuanced, and vary in terms of underlying intellectual qualities, what is measured as intelligence across many countries is largely consistent. A benefit of using an instrument such as the WISC is its proven reliability and validity in measuring its underlying construct of intelligence, which is solely cognitively based. The mismatch comes when imported tests based only on cognitive ability are used in countries that value social, emotional, or practical everyday aspects in construing one's general intelligence level, as imported tests largely do not meet these purposes.

What causes this discrepancy between the conception of intelligence and the measurement of intelligence? At least four factors can be accountable for this departure. First, although there might be different foci in terms of what constitutes intelligence, people from different cultures all recognize the importance of cognitive components in their conceptions of intelligence. This aspect of intelligence can be viewed as relatively universal and hence can be measured by similar tests. Second, measures of intelligence are primarily used for academic placement, such as for school admissions and tracking. Although many people criticize such a practice, it is still regarded as an effective way of allocating resources and of helping route students into specific areas of the labor market. This is especially the case in many developing societies, where resources are limited and there is a dire need for a quick and relatively objective way to place people. Despite their many limitations, compared to other types of measurements, IQ tests still demonstrate the highest predictive validity of one's academic achievement. Third, studies have consistently shown a moderate to strong correlation between a person's academic achievement and the analytical component of intelligence, measured by traditional IQ tests such as the Cognitive Abilities Tests (CAT) and the WISC-III (Brody, 1992; Frey & Detterman, 2004; Jensen, 1998; Neisser et al., 1996; Sternberg, Grigorenko, & Bundy, 2001; Watkins, Glutting, & Lei, 2007). It is not surprising that both researchers and educators still use traditional types of IQ tests in assessing individuals. Lastly, the creation of a new measurement based on contemporary theories of intelligences with a broader coverage to measure one's true intelligence is extremely difficult. Although there have been several such attempts (Brackett & Mayer, 2003; Gardner, 1993, 1995; Mayer, Salovey, & Caruso, 2000, 2002, 2004; Stemler & Sternberg, 2006; Sternberg, 2003), the road to perfecting these measurements, while also meeting people's practical needs, is still long and rough. It took many decades for the traditional IQ tests to mature and to be accepted by people in just one culture; it may require more intensive work to come up with new measurements to capture the important features of intelligence that will suit each particular society's needs.

Nature vs. Nurture

As illustrated in the chapter, the Western notion of intelligence strongly emphasizes one's innate ability, a value that can be traced back to ancient Western philosophers such as Plato. People from the rest of world have their own distinctive focuses that are different from Western notions. In many Chinese societies, despite the differences in political ideology, economic development, and even ethnic background, most people believe that knowledge and intelligence are closely related to each other. In their conception, one should also have good comprehension skills and good judgment about the immediate surroundings. Therefore, an intelligent person should have good cognitive competence, a curious mind, a thirst for knowledge, a wide range of knowledge, and possess a good memory (that is ready to take in yet more knowledge). These qualities are closely related to the Chinese cultural tradition of Confucianism, from which the ideas regarding intelligence were quoted at the beginning of the chapter. Although the Japanese view of intelligence was also influenced by Confucianism, the concept of effort, which is very important in the Japanese implicit theory of intelligence, is also largely a result of past and present societal values.

The growth feature of intelligence can also be demonstrated through the concept of CI. As examined in the section on cultural intelligence, CI is malleable and can be obtained through cross-cultural contact and training, either in a face-to-face format or through the Internet.

The Impact of Globalization in the Twenty-First Century

With globalization in the twenty-first century, the world has become more and more interconnected and interdependent. The ever-increasing scale of globalization has made an enormous impact on how people around the world view intelligence. An interesting phenomenon is that, just like the expert explicit theories of intelligence, Western laypeople's implicit theories of intelligence have gradually evolved from primarily focusing on cognitive abilities to emphasizing a comprehensive lists of attributes, including social competence and even moral components of intelligence. For example, Chamorro-Premuzic and colleagues (2010) carried out a study examining laypeople's conception of intelligence in British and American college students. They asked participants to rate the extent to which 109 different behaviors, attributes, or accomplishments are signs of intelligence. Results yielded three major underlying dimensions, namely academic IQ, social awareness, and SI.

Using a slightly different approach than the prototype of intelligence, Paulhus and colleagues (2002) had American and Canadian college students list the names of well-known people in history or in current affairs who are ideal examples of intelligent individuals. The results showed that the individuals named can be clustered into five distinctive categories, representing five different types of intelligences, that is, scientific intelligence (e.g., Einstein and Hawking), artistic intelligence (e.g., Mozart and Shakespeare), entrepreneurial intelligence (e.g., Turner, Trump, and Gates), communicative intelligence (e.g., the president, the prime minister, Oprah Winfrey), and moral intelligence (e.g., Gandhi and Martin Luther King, Jr.). In other words, it seems that not only have Western notions of intelligence influenced people's perception and practice in measuring intelligence across the rest of the world but also other cultures and their views of intelligence have helped shape what contemporary Westerners view as intelligence. Conceptions of intelligence are more inclusive than they used to be.

An increasing interest in social aspects of intelligence, such as SI, EI, and CI, is also observed worldwide in the most recent decade. People need to see the world beyond their personal lens and even beyond the cultural boundary. After all, intelligence is meant to be a mental activity directed toward purposive adaptation to and selection and shaping of real-world environments relevant to one's life (Sternberg, 1985). Knowing how people around the world understand intelligence will only enhance our ability to capture the concept better and measure it more accurately.

References

Alexandra, V. (2018). Predicting CQ development in the context of experiential cross-cultural training: The role of social dominance orientation and the propensity to change stereotypes. *Academy of Management Learning and Education*, *17*(1), 62–78. http://dx.doi.org/10.5465/amle.2015.0096

Alhashemi, S. E. (2014). Measuring emotional intelligence of university students: A comparison between China and Bahrain. *International Journal of Social and Organizational Dynamics in IT*, *3*(4), 59–76.

Ang, S., & Van Dyne, L. (2008). Conceptualization of cultural intelligence: Definition, distinctiveness, and nomological network. In S. Ang, & L. Van Dyne (Eds.), *Handbook of cultural intelligence: Theory, measurement, and applications* (pp. 3–15). New York: Sharpe.

Ang, S., Van Dyne, L., & Koh, C. (2006). Personality correlates of the four-factor model of cultural intelligence. *Group and Organization Management*, *31*(1), 100–123.

Azuma, H., & Kashiwagi, K. (1987). Descriptions for an intelligent person: A Japanese study. *Japanese Psychological Research*, *29*, 17–26.

Bai, X., Liu, H., & Hu, X. (2007). Development of high school teachers and students' view on implicit theories of intelligence. [Chinese]. *Studies of Psychology and Behavior*, *5* (2), 81–85.

Beller, M., & Gafni, N. (1995). Equating and validating translated scholastic aptitude tests: The Israeli case. In G. Ben-Shakhar & A. Lieblich (Eds.), *Studies in psychology in honor of Solomon Kugelmass* (pp. 202–219). Jerusalem: Magnes Press.

Brackett, M. A., & Mayer, J. D. (2003). Convergent, discriminate, and incremental validity of competing measures of EI. *Personality and Social Psychology Bulletin*, *29*(9), 1147–1158.

Brody, N. (1992). *Intelligence* (2nd ed.). San Diego: Academic Press.

Bruner, J. S., Shapiro, D., & Tagiuri, R. (1958). The meaning of traits in isolation and in combination. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp.278–288). Stanford: Stanford University Press.

Cai, X., & Jiang, L. (1995). A cross-cultural study of the intellectual concepts of junior and senior middle school students from five nationalities in Southwest China. [Chinese]. *The Science of Psychology*, *18*, 346–350.

Chamorro-Premuzic, T., Dissou, G., Furnham, A., & Bales, A. (2010). Personality traits and lay conceptions of intelligence. In J. P. Villanueva (Ed.), *Personality traits: Classifications, effects and changes (psychology of emotions, motivations, and actions) UK ed Edition* (pp. 109–123). Hauppauge, NY: Nova Science Publishers.

Chen, X. P., Liu, D., & Portnoy, R. (2012). A multilevel investigation of motivational cultural intelligence, organizational diversity climate, and cultural sales: Evidence from US real estate firms. *Journal of Applied Psychology*, *97*(1), 93–106.

Chua, R. Y., & Ng, K. Y. (2017). Not just how much you know: Interactional effect of cultural knowledge and metacognition on creativity in a global context. *Management and Organization Review*, *13*(2), 281–300.

Chen, M. J. (1994). Chinese and Australian concepts of intelligence. *Psychology and Developing Societies*, *6*, 101–117.

Chen, C. S., Lee, S. Y., & Stevenson, H. W. (1996). Academic achievement and motivation of Chinese students: A cross-national perspective. In S. Lau (Ed.). *Growing up the Chinese way: Chinese child and adolescent development* (pp. 69–91). Hong Kong: Chinese University Press.

Chen, W. W., & Wong, Y. L. (2014). What my parents make me believe in learning: The role of filial piety in Hong Kong students' motivation and academic achievement. *International Journal of Psychology*, *49*(4), 249–256.

Cole, M., Gay, J., & Glick, J. (1967). A cross-cultural study of clustering in free recall. *Psychonomic Bulletin*, *1*(2), 18.

Confucius (2010). *The doctrine of the mean* (trans. J. Leggs). Whitefish, MT: Kessinger Publishing. (Original translation published 1893.)

Crowne, K. A. (2008). What leads to cultural intelligence? *Business Horizons*, *51*(5), 391–399. http://dx.doi.org/10.1016/j.bushor.2008.03.010

Crowne, K. A. (2009). The relationships among social intelligence, emotional intelligence and cultural intelligence. *Organization Management Journal*, *6*, 148–163. https://doi.org/10.1057/omj.2009.20

Crowne, K. A. (2013). An empirical analysis of three intelligences. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *45*(2), 105–114. http://dx.doi.org/10.1037/a0029110.

Depaula, P. D., Azzollini, S. C., Cosentino, A. C., & Castillo, S. E. (2016). Personality, character strengths and cultural intelligence: "Extraversion" or "openness" as further factors associated to the cultural skills. *Avances en Psicología Latinoamericana*, *34*(2), 415–436.

Detterman, D. K. (1986). Qualitative integration: The last word? In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence: Contemporary viewpoints on its nature and definition* (pp.163–166). Norwood, NJ: Ablex Publishing.

Earley, P. C. (2002). Redefining interactions across cultures and organizations: Moving forward with cultural intelligence. *Research in Organizational Behavior*, *24*, 271–299. http://dx.doi.org/10.1016/S0191-3085(02)24008-3

Earley, P. C., & Mosakowski, E. (2004). Cultural intelligence. *Harvard Business Review*, *82* (10), 139–146.

Ekermans, G., & Privaatsak, X. (2009). Emotional intelligence across cultures: Theoretical and methodological considerations. In C. Stough, D. H. Saklofske, & J. Parker (Eds.), *Assessing emotional intelligence: Theory, research, and applications* (pp. 259–290). New York: Springer.

Fang, F., & Keats, D. (1987). A cross-cultural study on the conception of intelligence. [Chinese.] *Acta Psychologica Sinica*, *19*(3), 255–262.

Fang, F., Schei, V., & Selart, M. (2018). Hype or hope? A new look at the research on cultural intelligence. *International Journal of Intercultural Relations*, *66*, 148–171. https://doi.org/10.1016/j.ijintrel.2018.04.002

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*(6), 373–378.

Furnham, A., & Mkhize, N. (2004). Indian and Isi-Zulu-speaking South African parents' estimates on their own and their children's intelligence. *South African Journal of Psychology*, *34*, 363–384.

Fwu, B., Chen, S., Wei, C., & Wang, H. H. (2017). The mediating role of self-exertion on the effects of effort on learning virtues and emotional distress in academic failure in a Confucian context. *Frontiers in Psychology*, *7*, 1033–1089.

Gabel-Shemueli, R., Dolan, S., & Cerdin, J. (2005). Emotional Intelligence as predictor of cultural adjustment for success in global assignments. *Career Development International*, *10*, 375–395.

Garcia-Cepero, M. C., & McCoach, D. B. (2009). Educators' implicit theories of intelligence and beliefs about the identification of gifted students. *Universitas Psychologica*, *8* (2), 295–310.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.

Gardner, H. (1995). Reflections on multiple intelligences: Myths and messages. *Phi Delta Kappan*, *77*, 200–203, 206–209.

Gardner, H. (1998). A multiplicity of intelligences. *Scientific American Presents*, *9*, 18–23.

Gardner, H. (2002). On the three faces of intelligence. *Daedalus*, *131*(1), 139–142.

Georgas, J., Van de Vijver, F. J. R., Weiss, L. G., & Saklofske, D. H. (2003). In J. Georgas, L. G. Weiss, F. J. Van de Vijver, & D. H. Saflofske (Eds.), *Culture and children's intelligence: Cross-cultural analysis of the WISC-III* (pp. 227–240). San Diego: Academic Press.

Gill, R., & Keats, D. (1980). Elements of intellectual competence: Judgments by Australian and Malay university students. *Journal of Cross-Cultural Psychology*, *11*, 233–243.

Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*, *52*, 1115–1124.

Greenfield, P. M. (1998). The cultural evolution of IQ. In U. Neisser (Ed.), *The rising curve: Long term gains in IQ and related measures* (pp. 81–123). Washington, DC: American Psychological Association.

Grigorenko, E. L., Geissler, P. W., Prince, R., Okatcha, F. Nokes, C. et al. (2001). The organisation of Luo conceptions of intelligence: A study of implicit theories of Kenya village. *International Journal of Behavioral Development*, *25*(4), 367–378.

Gunkel, M., Schlaegel, C., & Teras, V. (2016). Cultural values, emotional intelligence, and conflict handling styles: A global study. *Journal of World Business*, *51*(4), 568–585.)

Harrison, N. (2012). Investigating the impact of personality and early life experiences on intercultural interaction in internationalised universities. *International Journal of Intercultural Relations*, *36*(2), 224–237. http://dx.doi.org/10.1016/j.ijintrel.2011.03.007

Hau, K. T., & Ho, I. T. (2012). Chinese students' motivation and achievement. In M. H. Bond (Ed.). *The Oxford handbook of Chinese psychology* (pp. 187–204). New York: Oxford University Press.

Holloway, S. D. (1988). Concepts of ability and effort in Japan and the United States. *Review of Educational Research*, *58*(3), 327–345.

Jensen, A. (1998). The suppressed relationship between IQ and the reaction time slope parameter of the Hick function. *Intelligence*, *26*(1), 43–52.

Kim, Y. J., & Van Dyne, L. (2012). Cultural intelligence and international leadership potential: The importance of contact for members of the majority. *Applied Psychology*, *61*(2), 272–294. http://dx.doi.org/10.1111/j.1464-0597.2011.00468.x

Ko, B., Boswell, B., & Yoon, S. (2015). Developing intercultural competence through global link experiences in physical education. *Physical Education and Sport Pedagogy*, *20* (4), 366–380. http://dx.doi.org/10.1080/17408989.2013.837441

Kopic, K., Vranic, A., & Zarevski, P. (2009). Implicit theories of intelligence in elementary school eighth-grade pupils. *Drustvena Istrazivanja*, *18*(3), 503–521.

Korzilius, H., Bücker, J. J., & Beerlage, S. (2017). Multiculturalism and innovative work behavior: The mediating role of cultural intelligence. *International Journal of Intercultural Relations*, *56*, 13–24. http://dx.doi.org/10.1016/j.ijintrel.2016.11.001

Kwak, K. (2003). South Korea. In J. Georgas, L. G. Weiss, F. J. Van de Vijver, & D. H. Saflofske (Eds.), *Culture and children's intelligence: Cross-cultural analysis of the WISC-III* (227–240). San Diego: Academic Press.

Lautrey, J., & de Ribaupierre, A. (2004). Psychology of human intelligence in France and French-Speaking Switzerland. In R. J. Sternberg (Ed.), *International handbook of intelligence* (pp. 104–134). New York: Cambridge University Press.

Lie, D., Suyasa, P. T., & Wijaya, E. (2016). The mediating role of cultural intelligence in the relationship between the openness to experience personality trait and job satisfaction among expatriates. *Makara Human Behaviour Studies in Asia*, *20*(1), 46–56. http://dx.doi.org/10.7454/mssh.v20i1.3486

Lim, W., Plucker, J. A., & Im, K. (2002). We are more alike than we think we are: Implicit theories of intelligence with a Korean sample, *Intelligence*, *30*(2),185–208.

MacNab, B. R., & Worthley, R. (2012). Individual characteristics as predictors of cultural intelligence development: The relevance of self-efficacy. *International Journal of Intercultural Relations*, *36*(1), 62–71. http://dx.doi.org/10.1016/j.ijintrel.2010.12.001.

Mayer, J. D., DiPaolo, M., & Salovey, P. (1990). Perceiving the affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment*, *50*, 772–781.

Mayer, J. D., Salovey, P., & Caruso, D. (2000). Models of emotional intelligence. In R. J. Sternberg (Ed.), *Handbook of Intelligence* (pp. 396–420). New York: Cambridge University Press.

Mayer, J. D., Salovey, P., & Caruso, D. (2002). *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT), Version 2.0*. Toronto: Multi-Health Systems.

Mayer, J. D., Salovey, P., & Caruso, D. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, *15*(3), 197–215.

Moon, H. K., Choi, B. K., & Jung, J. S. (2012). Previous international experience, cross-cultural training, and expatriates' cross-cultural adjustment: Effects of cultural intelligence and goal orientation. *Human Resource Development Quarterly*, *23*(3), 285–330. http://dx.doi.org/10.1002/hrdq.21131

Moon, H. K., Choi, B. K., & Jung, J. S. (2013). Comprehensive examination on antecedents of cultural intelligence: Case of South Korea. *Personnel Review*, *42*(4), 440–465. http://dx.doi.org/10.1108/PR-10-2011-0146

Muñiz, J., Bartram, D., Evers, A., Boben, D., Matesic, K., Glabeke, K., et al. (2001). Testing practices in European countries. *European Journal of Psychological Assessment*, *17* (3), 201–211.

Muñiz, J., Prieto, G., Almeida, L., & Bartram, D. (1999). Test use in Spain, Portugal and Latin American countries. *European Journal of Psychological Assessment*, *15*(2), 151–157. http://dx.doi.org/10.1027//1015-5759.15.2.151

Neisser, U. (1979). The concept of intelligence. In R. J. Sternberg & D. T. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement* (pp. 179–189). Norwood, NJ: Ablex.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A.W., Brody, N., Ceci, S. J. et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*(2), 77–101.

Niu, W., & Brass, J. (2011). Intelligence in worldwide perspective. In S. Kaufman & R. J. Sternberg (Eds.), *Cambridge handbook of intelligence* (pp. 623–646). New York: Cambridge University Press.

Oakland, T. (2009). How universal are test development and use? In E. Grigorenko (Ed.), *Assessment of abilities and competencies in an era of globalization* (pp. 1–40). New York: Springer.

Oakland, T., Douglas, S., & Kane, H. (2016). Top ten standardized tests used internationally with children and youth by school psychologists in 64 countries: A 24 year follow-up study. *Journal of Psychoeducational Assessment*, *34*(2), 166–176.

Paulhus, D. L., Wehr, P., Harms, P. D., & Strasser, D. I. (2002). Use of exemplar surveys to reveal implicit types of intelligence. *Personality and Social Psychology Bulletin*, *28*, 1051–1062.

Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, *54*(9), 741–754.

Plato (1992). *The Republic* (trans. G. M. A. Grube) (2nd ed.). Indianapolis: Hackett Publishing.

Rosenblatt, V., Worthley, R., & MacNab, B. (2013). From contact to development in experiential cultural intelligence education: The mediating influence of expectancy disconfirmation. *Academy of Management Learning and Education*, *12*(3), 356–379. http://dx.doi.org/10.5465/amle.2012.0199

Ruzgis, P. M., & Grigorenko, E. L. (1994). Cultural meaning systems, intelligence and personality. In R. J. Sternberg & P. Ruzgis (Eds.), *Personality and intelligence* (pp. 248–270). New York: Cambridge University Press.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, *9*, 185–211.

Salovey, P., & Pizarro, D. A. (2003). The value of emotional intelligence. In R. J. Sternberg, J. Lautrey, & T. I. Lubart (Eds.), *Models of intelligence: International perspectives* (pp. 263–278). Washington, DC: American Psychological Association.

Saklofske, D. H., Van de Vijver, F. J. R., Oakland, T., Mpofu, E., & Suzuki, L. A. (2015). Intelligence and culture: History and assessment. In S. Goldstein, D. Princiotta, J. Naglieri (Eds.), *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts* (pp. 341–365). New York: Springer. https://doi.org/10.1007/978-1-4939-1562-0_22

Sato, T., Namiki, H., Ando, J., & Hatano, G. (2004). Japanese conception of and research on human intelligence. In R. J. Sternberg (Ed.), *International handbook of intelligence* (pp. 302–324). New York: Cambridge University Press.

Serpell, R. (1974). Aspects of intelligence in a developing country. *African Social Research*, *17*, 576–596.

Serpell, R. (2000). Intelligence and culture. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 549–580). New York: Cambridge University Press.

Serpell, R., & Jere-Folotiya, J. (2008). Developmental assessment, cultural context, gender, and schooling in Zambia. *International Journal of Psychology*, *43*(2), 88–96.

Srivastava, A. K., & Misra, G. (2001). Lay people's understanding and use of intelligence: An Indian perspective. *Psychology and Developing Societies*, *13*(1), 25–49. https://doi.org/10.1177/097133360101300102

Stemler, S. E., & Sternberg, R. J. (2006). Using situational judgment tests to measure practical intelligence. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 107–131). Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, *49*, 607–627.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.

Sternberg, R. J. (2003). Construct validity of the theory of successful intelligence. In R. J. Sternbger, J. Lautrey, & T. I. Lubart (Eds): *Models of intelligence: International perspectives* (pp. 55–77). Washington, DC: American Psychological Association.

Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist*, *59*, 325–338.

Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, *41*(1), 37–55.

Sternberg, R. J., & Grigorenko, E. L. (2006). Cultural intelligence and successful intelligence. *Group and Organization Management*, *31*(1), 27–39.

Sternberg, R. J., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C, et al. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence*, *30*(2), 141–162.

Sternberg, R. J., Nokes, C., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A, et al. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence*, *29(5)*, 401–418.

Swami, V., Furnham, A., Maakip, I., Ahmad, M. S., Naw, N.H.M., Voo, P. S. K. et al. (2008). Beliefs about the meaning and measurement of intelligence: A cross-cultural comparison of American, British, and Malaysian undergraduates. *Applied Cognitive Psychology*, *22*(2), 235–246.

Thorndike, R. L., & Stein, S. (1937). An evaluation of the attempts to measure social intelligence. *Psychological Bulletin*, *34*, 275–285. https://doi.org/10.1037/h0053850

Van de Vijver, F. J. R. (2003). Principles of adaptation of intelligence tests to other cultures. In J. Georgas, L. G. Weiss, F. J. Van de Vijver, & D. H. Saflofske (Eds.), *Culture and children's intelligence: Cross-cultural analysis of the WISC-III* (pp. 255–263). San Diego: Academic Press.

Ueda, N. (1989). Japanese children's personal theories of intelligence: A developmental study. Doctoral dissertation, Cambridge: Harvard University.

Wan, M., Li, N., & Jing, Q. (1997). A cross-cultural study on middle school and high school students' implicit theories of intelligence: Comparison among Han, Tabitan, and Dongxiang Chinese students. [Chinese.] *Psychological Development and Education*. *2*, 1–6.

Watkins, M. W., Glutting, J. J., & Lei, P.-W. (2007). Validity of the full-scale IQ when there is significant variability among WISC-III and WISC-IV factor scores. *Applied Neuropsychology*, *14*(1), 13–20.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*, 297–333.

Willmann, E., Feldt, K., & Amelang, M. (1997). Prototypical behavior patterns of social intelligences: An intercultural comparison between Chinese and German subjects. *International Journal of Psychology*, *32*(5), 329–346.

Wober, M. (1974). Towards an understanding of the Kigranda concept of intelligence. In J. W. Berry & P. R. Dasen (Eds.), *Culture and cognition* (pp. 261–280). London: Methuen.

Yang, S. -Y., & Sternberg, R. J. (1997a). Conceptions of intelligence in ancient Chinese philosophy. *Journal of Theoretical and Philosophical Psychology*, *17*, 101–119.

Yang, S. -Y., & Sternberg, R. J. (1997b). Taiwanese Chinese people's conceptions of intelligence, *Intelligence*, *25*, 21–36.

Zhang, H., & Wu, Z. (1994). People's conceptions of intelligence: A study of Beijing residents' conceptions of intelligence. [Chinese.] *The Science of Psychology*, *17* (2), 65–69, 81.

38 Historical Evolution of Intelligence

Patricia M. Greenfield

Intelligence Shifts Historically from Practical to Abstract

> Cole, Gay, Glick, and Sharp (1971) took an object-sorting task to Liberia, where they presented it to their Kpelle participants. There were 20 objects that divided evenly into the linguistic categories of foods, implements, food containers, and clothing. Instead of doing the taxonomic sorts expected by the researchers, participants persistently made functional pairings (Glick 1968). For example, rather than sorting objects into groups of tools and foods, participants would put a potato and a knife together because "you take the knife and cut the potato" (Cole et al. 1971, p. 79). According to Glick, participants often justified their pairings by stating "that a wise man could only do such and such" (Glick 1968, p. 13). In total exasperation, the researchers "finally said, 'How would a fool do it?' The result was a set of nice linguistically ordered categories – four of them with five items each" (Glick 1968, p. 13). In short, the researchers' criterion for intelligent behavior was the participants' criterion for foolish; the participants' criterion for wise behavior was the researchers' criterion for stupid.
>
> (Quote from Greenfield, 1997, summarized in Glick, 1974)

This quote encapsulates one of the most important trends in the history of intelligence: the shift from considering intelligence to be practical (Sternberg & Grigorenko, 2000) and contextualized to abstract and decontextualized. This quote refers to a cross-cultural difference. What is the justification to also call it a historical trend?

The answer begins with a definition of intelligence. "Successful adaptation to its own niche marks an animal form as intelligent" (Scheibel, 1996). In both Piaget's theory of cognitive development and Wechsler's psychometric approach to IQ testing, adaptation is central to the definition (Dasen, 1984; Greenfield, 1974/1976). However, neither Piaget nor Wechsler considers the possibility that adaptation may differ in different ecocultural settings (Berry, 1974).

By definition, abstraction implies distancing from the immediate context. Nonetheless, practical contextualized intelligence is adaptive in small, isolated, rural, subsistence communities in which members grow their own food, make their own clothes, and build their own shelters; education takes place in the locations where these subsistence skills are carried out and does not take place in school.

In contrast, abstract, decontextualized intelligence is adaptive in large, interconnected, urban, commercial, high-tech societies in which education takes place in school.

Intelligence, as defined in psychology, focuses on abstract, decontextualized intelligence. The overwhelming majority of psychological scientists assume the universality of this definition without understanding that abstraction is a cultural form. Abstraction as an important desideratum of intelligence is common to both Piaget – whose highest stage of cognitive development, formal operations, is also the most abstract – and to Wechsler, author of the most venerated and popular IQ test.

Historically speaking, small face-to-face agricultural villages existed before large urban centers and therefore practical intelligence existed before abstract intelligence. Therefore, cross-cultural differences between members of communities with the former characteristics and communities with the latter characteristics model long-term historical change in the definition of intelligence. The opening quote shows two extremes along the continuum. But note that practical intelligence still exists alongside abstraction in the United States; witness apprenticeship programs for manual skills like those used in construction work. I will return to this issue in my conclusion.

However, there are many intermediate points in the continuum illustrated in my opening quote; and social change can move cognitive adaptation along the continuum. Here, is an example from Luria's classic research of a shift in cognitive processing that implies a historical shift from practical intelligence to abstract intelligence:

The Soviet Union developed formal education in rural areas after the Revolution of 1917. Because of this development, learning environments came to include literacy and schooling. Luria (1976) compared farmers from remote villages in Central Asia without any school experience with participants having 1–2 years of school experience – and therefore basic literacy skills. Farmers with no schooling addressed Luria's problems as concrete practical situations. After a few years of schooling with basic literacy skills, they took a more abstract approach, separating their thought processes from the practical situation. Again, this comparison between schooled and unschooled farmers was a synchronic or cross-sectional model of what happened historically when print literacy and elementary schooling for children were introduced into the oral culture of rural communities.

Here, is a qualitative example of Luria's findings. Participants were shown drawings of a hammer, saw, log, and hatchet. They were asked, "Which ones are alike?" Rakmat was thirty-nine years old, had never been to school, and was illiterate. He groups items by their use in a practical context (Luria, 1976, p. 56):

RAKMAT: "They're all alike. I think all of them have to be here . . ."

LURIA: "But one fellow picked three things – the hammer, saw, and hatchet – and said they were alike."

RAKMAT: "A saw, a hammer, and hatchet all have to work together. But the log has to be here too!"

LURIA: "Why do you think he picked these three things and not the log?"

RAKMAT: "Probably he's got a lot of firewood, but if we'll be left without firewood, we won't be able to do anything."

Thus, Rakmat constructs a practical situation in which all the items are necessary. Similar to Glick's example in Liberia, "every attempt to suggest the possibility of categorical grouping met with protest: 'That's wrong. Some stupid fellow told you that, he doesn't understand anything'" (Luria, 1976, p. 77). For this group, our definition of intelligent was their definition of stupid!

Contrast Rakmat's responses with those of Yadgar, eighteen years old with two years of school experience; he has acquired basic literacy skills. He is shown drawings of a glass, saucepan, spectacles, and bottle and asked, "Which ones are alike?" Yadgar answers, "The glass, spectacles, and bottle all fit together. They're made of glass, but the saucepan is metal." He constructs a superordinate category – material – that is removed from the practical context of use. Luria's quantitative analysis showed that Rakmat was typical of the group with no school or literacy experience, whereas Yadgar was typical of young people with one to two years of school and print literacy experience. In sum, the historical change of introducing literacy and schooling into a formerly illiterate environment with all education in situ produces decontextualized and abstract concept formation (Greenfield, 2019).

Computers continue this trend toward abstraction as technology advances. OLPC (One Laptop Per Child) is a US-based nonprofit organization that provides the world's poorest children with laptops and software designed for independent learning. A field experiment in Ethiopia explored the effect of computers distributed by OLPC on the abstract reasoning of ten- to fifteen-year-olds (Hansen et al., 2012). In four schools, 202 children in Grades 5, 6, and 7 were given laptops (all the children in three schools, half the children in a fourth); these children were compared with 210 fifth, sixth, and seventh graders who were not given laptops (all the children in three sociodemographically matched schools and the other half of the children in the fourth school). Laptops were mainly used at home and during breaks at school; they were hardly used for teaching purposes in class. The most frequent computer activities were writing, reading, and gaming.

Students with laptops in the two older groups significantly outperformed their peers without laptops on two abstract reasoning tests – one tested analogies, the other categories. Examples of the two tests are shown in Figure 38.1. Note that the categorization test is a child version of Luria's categorization problems, described earlier in the section. Interestingly, the effects of laptops did not improve school performance; instead, the laptop effects were specific to cognitive abstraction. Hence, the progression begun by print literacy and schooling, as shown in Luria's studies, was continued by computer technology in the form of laptops (Greenfield, 2019). Because computers are a technology recently added to a preexisting ecology, this field experiment also models historical change in intelligence.

Why do schooling, print literacy, and computers stimulate the same shift in valued cognitive processes, that is, intelligence, from practical to abstract thinking? The answer lies in a theory of social change and human development.

Theory of Social Change and Human Development

This theory is interdisciplinary, integrating concepts from sociology, anthropology, and psychology. It is also multilevel, positing causal relations among the levels (Greenfield, 2009b, 2016, 2018). It incorporates sociodemographic variables at the top

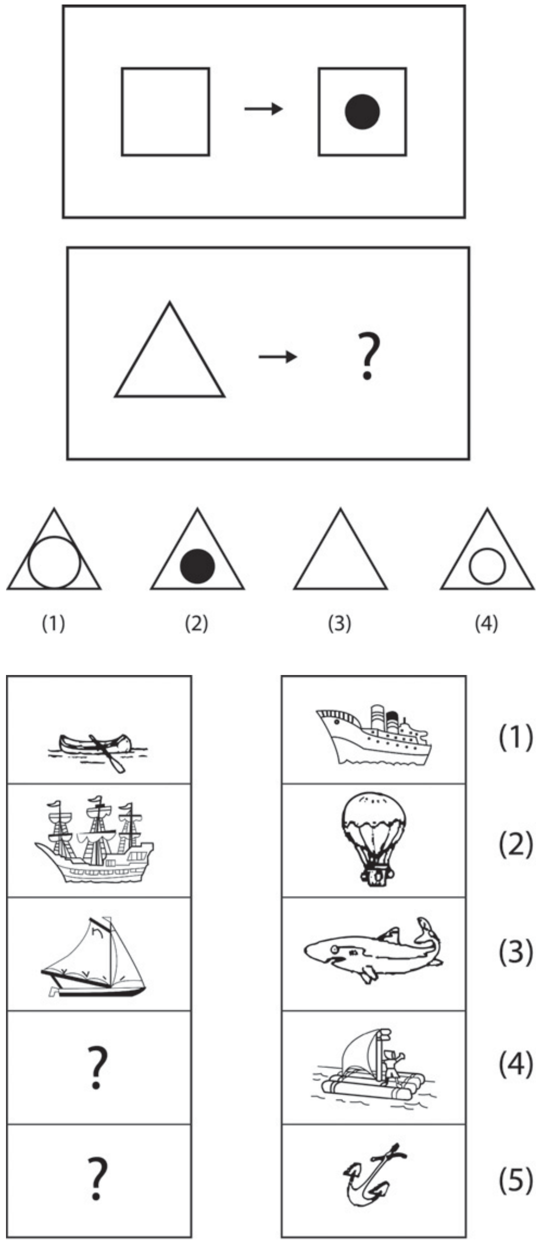

Figure 38.1 *Example of Analogies test and Categories test from the participant's point of view (from Hansen et al., 2012).*
Top: Example of Analogies test from the participant's point of view (the correct solution is alternative 2). By one or more changes to the figure to the left of the arrow in the top row, the figure to the right of the arrow is created. Children are asked which of the four figures in the bottom row should replace the question mark. The answer is given by applying the transformation in the top row to the middle row.
Bottom: Example of Categories test from the participant's point of view (the correct solution is 1 and 4). The child is asked which of the pictures 1–5 are from the same category as the three pictures on the left. In order to answer this question, the respondent has to discover the concept underlying the three pictures and apply them to novel images.

of the causal chain (with nineteenth-century roots in the German sociologist Tönnies, 1887/1957), cultural values at the next level down, learning environment at the next level, and individual development at the bottom (Figure 38.2).

Gemeinschaft (Community) and *Gesellschaft* (Society) summarize the features that anchor each end of the sociodemographic dimension (top level of Figure 38.2). Gemeinschaft denotes a small-scale social entity with all social relations based on close personal and lifelong ties – for example, extended family relations in a rural village; in contrast, Gesellschaft denotes a large-scale social entity with many relationships that are impersonal and transitory – for example, store clerks in an urban city. Each term, Gemeinschaft and Gesellschaft, summarizes a complex of sociodemographic elements. These features of Gemeinschaft and Gesellschaft provide anchors or endpoints for specific dimensions, listed on the sociodemographic level (top rectangle of Figure 38.2). All of the dimensions in the sociodemographic rectangle of Figure 38.2 tend to covary and shift together (Greenfield, 2018). The top horizontal arrow in Figure 38.2 denotes the dominant direction of globalized social change – from Gemeinschaft to Gesellschaft along multiple dimensions exemplified in the top rectangle.

We can think of Gemeinschaft features as being close to the environment in which human beings evolved. However, we have almost no "pure" Gemeinschafts left in the world. Most actual environments are somewhere in between the extreme endpoints on the various dimensions. The horizontal change arrows in Figure 38.2 therefore denote a *direction of movement*, not absolute locations on various scales.

Most important, the sociodemographic level (top rectangle, Figure 38.2) is at the top of the causal chain, influencing each lower level (vertical arrows from the Sociodemographic level to Cultural Values, Learning Environment, and Individual Development – lower three rectangles, Figure 38.2). Each lower level is influenced by and adapted to the ones above it (see the vertical arrows from the Cultural Value level to the Learning Environment, which includes socialization, and from the Learning Environment to Individual Development and Behavior).

So, when there is a shift on the top, Sociodemographic level from Community/ Gemeinschaft features in the direction of Society/Gesellschaft features, then there are correlated shifts on the lower levels of Cultural Values, Learning Environment, and Individual Development in the same direction; these shifts are denoted by the horizontal arrows in Figure 38.2. These are the historical changes. Note that all the historical shifts that are diagrammed did not take place simultaneously. A sense of timing will be given as the chapter progresses. The reader will also see that many historical changes in Cultural Values concerning intelligence are ongoing.

Each shift on a lower level is a theoretically driven prediction. Driven by sociodemographic change, these correlated changes on multiple levels constitute the heart of the theory of social change and human development.

Figure 38.2 is a guide to what follows in the remainder of the chapter; it lists the particular shifts on the levels of Cultural Values, Learning Environment, and Individual Development (bottom three rectangles of Figure 38.2) brought about by the global rise of cities, commerce, formal education, wealth, and communication technologies (right side of top rectangle in Figure 38.2). On the level of Cultural Values, the focus will be on values relating to intelligence (second level, Figure 38.2).

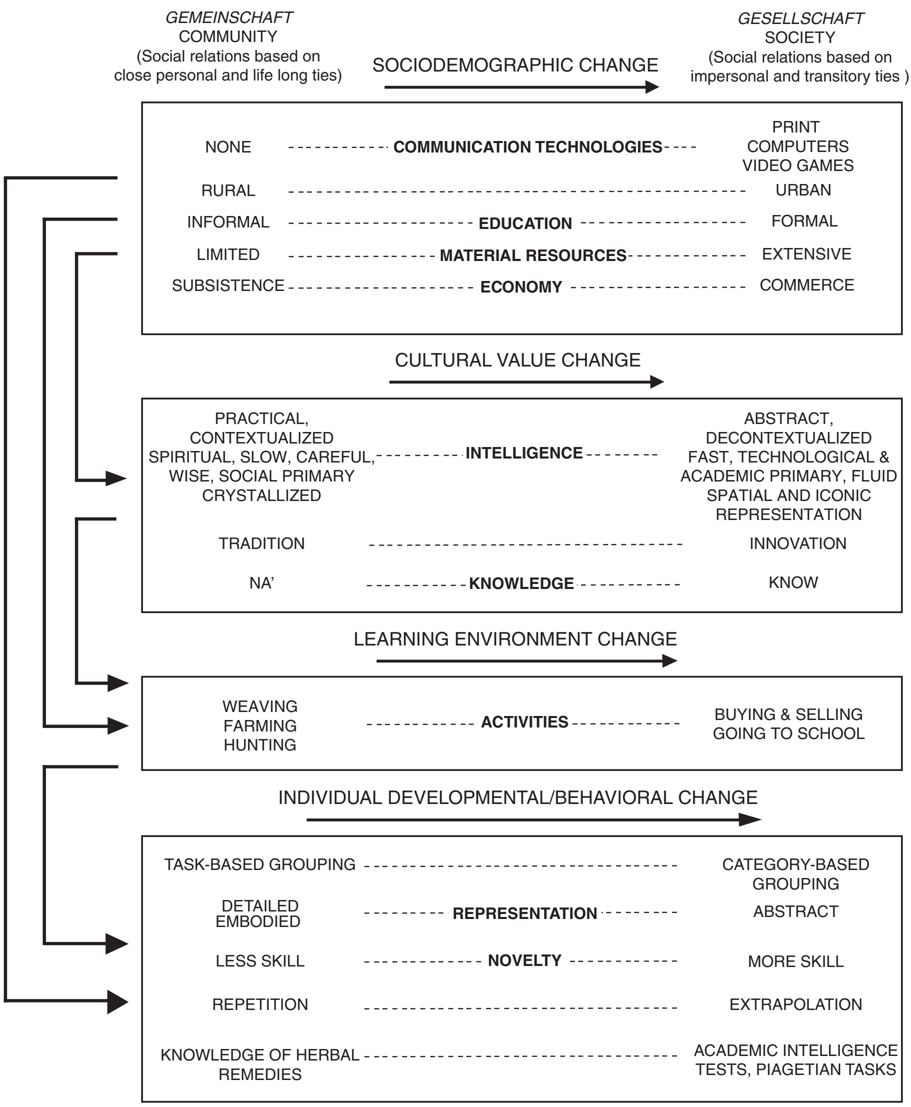

Figure 38.2 *Model of social change, cultural value change, and individual developmental change.*
Relationships for which there is empirical evidence, described in the text, have been selected for inclusion. While the horizontal arrows represent the dominant direction of social change in the world, sociodemographic change can go in the opposite direction. In that case, all the horizontal arrows would be reversed. The vertical arrows represent directions of influence, causal relations.

Both Sociodemographic shifts (top rectangle in Figure 38.2) and shifts in Cultural Values (second rectangle down, Figure 38.2) produce Learning Environment Change (third rectangle down, Figure 38.2), which in turn produces Individual Developmental Change (fourth rectangle down, Figure 38.2).

One relevant tenet of the theory is the equipotentiality of all of the sociodemographic factors in the top rectangle. It is this equipotentiality that explains why both schooling and computers stimulate the same direction of valued cognitive processes, that is, intelligence, toward abstract thinking. Equipotentiality means that when any sociodemographic variable moves from a more Gemeinschaft to a more Gesellschaft value, it will trigger the identical direction of change on a lower level – in this case from practical intelligence to abstraction. A corollary tenet is that whatever sociodemographic factor or factors is/are changing most rapidly at a given time or place will be the major factor(s) driving cultural and psychological change at that time or place. The shifts on all four levels depicted in Figure 38.2 furnish a guide to and summary of this chapter.

I begin with an overview of the sociodemographic level. The top level of Figure 38.2 depicts a global sociodemographic cluster that has been moving over time away from social relations based on close personal and lifelong ties (Gemeinschaft/Community) toward social relations based on more impersonal and transitory ties (Gesellschaft/Society) (Tönnies, 1887/1957). While moving historically in a common direction, the sociodemographic factors defining Gemeinschaft/Community (left side of top rectangle, Figure 38.2) cluster together and the sociodemographic factors defining Gesellschaft/Society (right side of top rectangle, Figure 38.2) cluster together (Greenfield, 2019; Santos, Varnum, & Grossmann, 2017). These patterns of intercorrelation and differentiation are all posited by the theory of social change and human development (Greenfield, 2009b, 2016, 2018).

From Tradition and Task-Relevant Detail to Innovation and Abstract Representation

The gold standard for drawing historical conclusions is diachronic evidence – data collected at different historical periods (Greenfield, 2018). So far, this picture of the shift in defining intelligence from task-relevant cognition to abstraction is based on historical inferences from synchronic evidence – comparative data collected in the same chronological period. This method provides indirect evidence. In contrast, historical studies that provide diachronic data provide direct evidence. A quasi-experimental study of pattern representation in three generations of children and adolescents in a Maya community in Chiapas, Mexico provides just such diachronic evidence concerning the shift from detailed, contextually relevant cognition to abstract representation and greater skill in solving novel problems (Maynard, Greenfield, & Childs, 2015) (see bottom rectangle of Figure 38.2).

The task was to place colored sticks in a frame to represent culturally central woven patterns (Figures 38.3 and 38.4). The first generation was assessed in 1969 and 1970 when the community economy was agriculture based. The second generation was tested in 1991, when the community economy had transitioned to commerce. This movement from a Gemeinschaft to a Gesellschaft ecology continued into the third generation, tested in 2012. The main sociodemographic change between Generation 2 and Generation 3 was an increase in school-based education (see top rectangle Figure 38.2).

Figure 38.3 *A nine-year-old girl places sticks in a frame to represent a red-and-white striped woven pattern.*
Photograph © Lauren Greenfield; original in color.

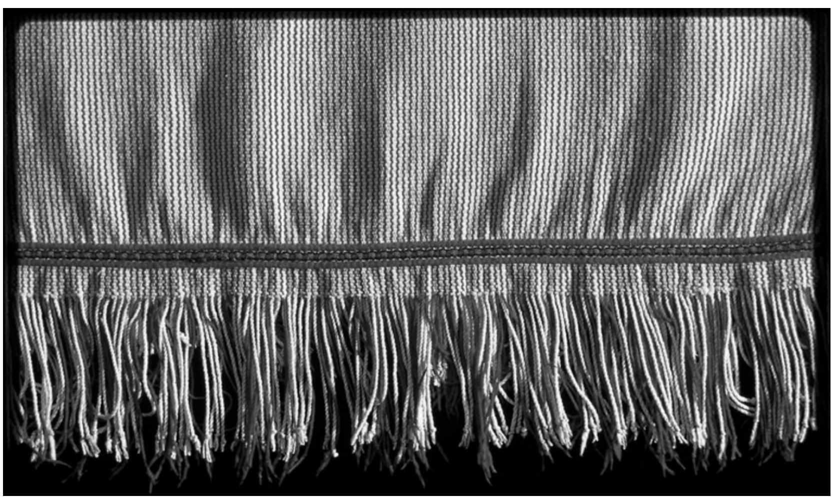

Figure 38.4 *Pattern of a Zinacantec man's red-and-white striped woven poncho (above) and a woman's red-and-white striped woven shawl (next page).*
Note that the shawl's wide red stripes are composed of three thin red lines separated by two thin white lines.
Photographs © Lauren Greenfield; originals in color.

Figure 38.4 *(cont.)*

The Role of Commerce

The historical shift from subsistence to commerce affected cognitive development, making it more abstract. From Generation 1 to Generation 2, there was a shift from detailed, thread-by-thread representation of woven patterns – the kind of representation required to actually create the woven textile patterns (Figure 38.5) – to more abstract representations (Figure 38.6) (diagrammed in bottom rectangle of Figure 38.2).

In addition to representing woven patterns, participants were asked to use the same apparatus to continue novel patterns, striped patterns that were unknown in the community. In addition to more abstract representation of familiar woven patterns in Generation 2, there was also a shift toward a greater ability to represent the novel patterns in the commercial period (diagrammed in the bottom rectangle of Figure 38.2). Note that novelty was a negative in the earlier period of subsistence, represented by Generation 1. If one made a textile that was "different," *t'oso* in Tzotzil, it was considered a bad thing. T'oso had the negative connotation of deviating from a norm. However, by 1993, with commerce established in the community, being different, that is, novel, in textile design was considered a positive (Greenfield, 2004). This value was much more in line with the importance of novelty in tests of fluid intelligence.

Structural-equation modeling showed that both these shifts – toward abstraction and skill in representing novel patterns – were driven by the participation of children and their parents in commercial activities (Greenfield, Maynard, & Childs, 2003). Commercial activities were distinguished from the older subsistence activities, such as males growing corn, females using the corn to make tortillas and weaving clothes for the family. (This shift is diagrammed in Figure 38.2; the links from a commercial economy to a Learning Environment that features buying and selling to abstraction

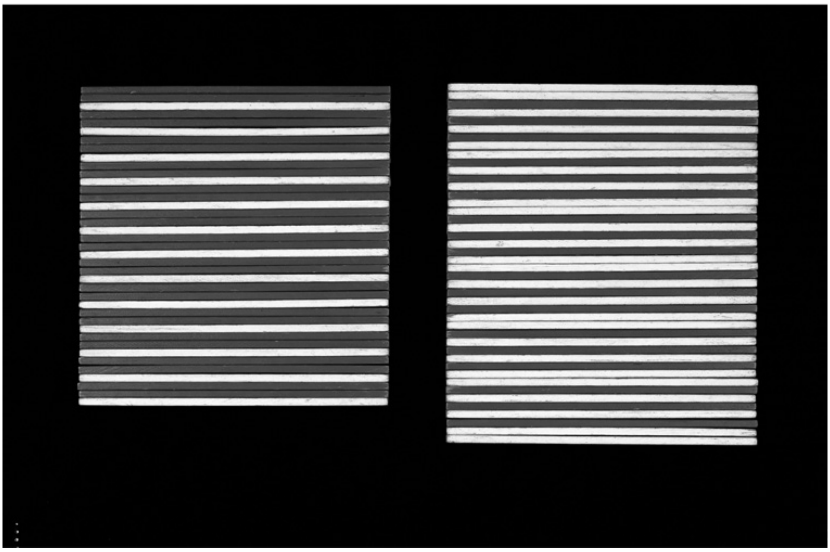

Figure 38.5 *A detailed, "thread-by-thread" representation of two Zinacantec woven patterns: a red-and-white striped poncho (left) and a red-and-white striped shawl (right).*
Photograph by Don Cole; original in color.

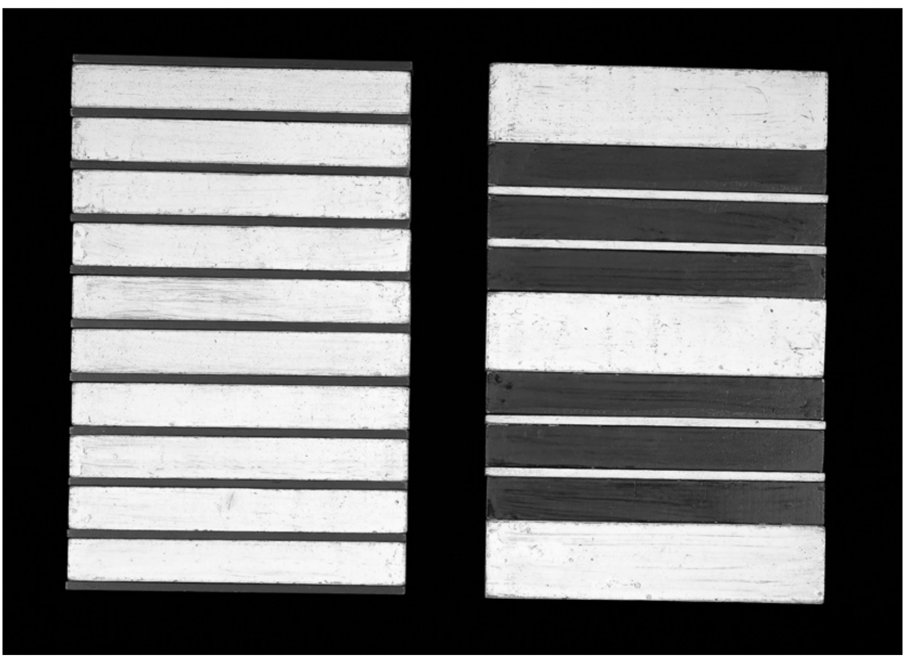

Figure 38.6 *An abstract representation of two Zinacantec woven patterns: a red-and-white striped poncho (left) and a red-and-white striped shawl (right).*
Photograph by Don Cole; original in color.

and skill with novelty is represented by the vertical arrow from the Sociodemographic rectangle to the Learning Environment rectangle to the Individual Development rectangle.) Commercial activities included both participation as a consumer – for example, having a television in the home – and participation as an entrepreneur – for example, a father buying and selling goods for a living. Innovation is valued in a commercial, entrepreneurial economy. Indeed, innovative (vs. traditional) pattern design had entered Zinacantec textile design in the twenty-one years between Generation 1 and Generation 2 (Greenfield, 1999, 2004) (diagrammed in the Cultural Value Change rectangle, Figure 38.2).

One novel pattern, the "growing pattern" (Figure 38.7), deserves description in more detail. Because it is so much like an IQ subtest, this pattern is of special interest as an example of the historical augmentation of fluid intelligence and "going beyond the information given" (Bruner, Goodnow, & Austin, 1956). Figure 38.7 shows the pattern and the three "correct" ways to continue it. The third way, making it grow (far right), would be assumed in our culture and, indeed, I was surprised in 1969 when almost no participants utilized this strategy. Instead, most participants who were old enough to deal with this, our most complex pattern, repeated the pattern (far left); the second most popular correct strategy was to make a mirror image (middle pattern), a strategy probably based on mirror-image borders on each side of a weaving (Greenfield & Childs, 1977; Greenfield et al., 2003). However, as commerce grew, the percentage of children using the growing strategy, compared with the other two correct strategies, increased from 18 percent to 62 percent (Maynard et al., 2015). More children were extrapolating, "going beyond the information given." This is a shift from repetition to extrapolation (diagrammed in the bottom, Individual Development rectangle of Figure 38.2). The connection of this change to a commercial economy is represented by the vertical arrow from the Sociodemographic rectangle to the Individual Development rectangle in Figure 38.2. To use the vocabulary of intelligence test development, Zinacantec children were moving, implicitly, in the direction of higher fluid intelligence; they were treating the pattern as a novel problem rather than something that was already part of their knowledge base (crystallized intelligence). The connection between the shift to a commercial economy and the shift toward fluid intelligence is represented by the vertical arrow from the Sociodemographic rectangle to the Cultural Value rectangle (Figure 38.2).

The Role of Formal Education

From Generation 2 to Generation 3, the increase in educational opportunity drove further increases in abstract representation of woven patterns and skill in continuing novel patterns (Maynard et al., 2015). At the same time, schooling came to replace weaving in the Learning Environment of Zinacantec girls (diagrammed in the third level of Figure 38.2); correlatively, detailed thread-by-thread representation of textile patterns was most frequent in girls with more varied weaving experience (diagrammed in the fourth level of Figure 38.2). This link is represented by the vertical arrow from the Learning Environment rectangle to the Individual Development

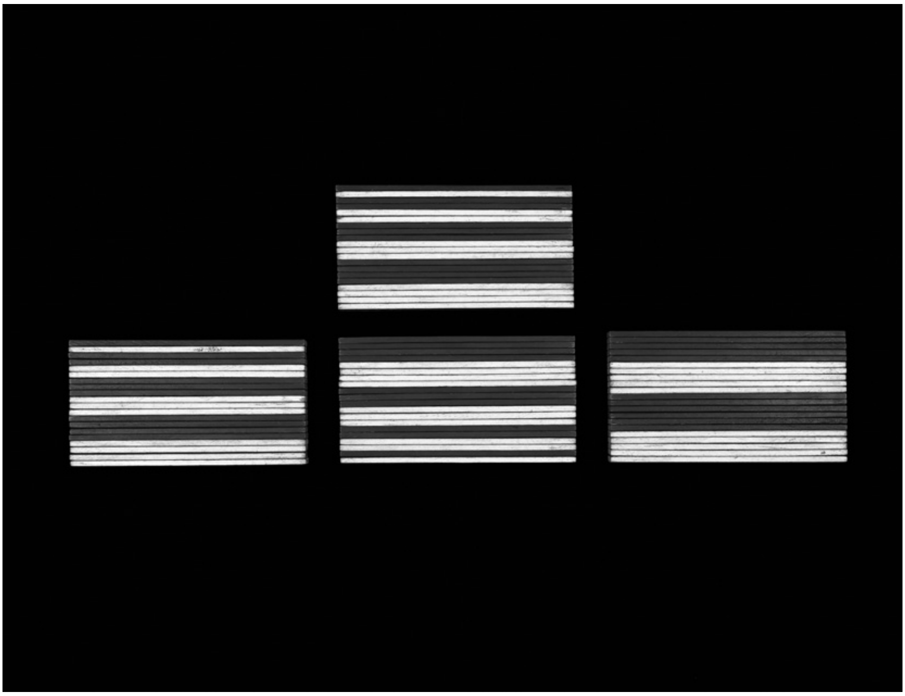

Figure 38.7 *Model for the growing pattern and three possible continuations. The experimenter's model is at the top and three possible "correct" responses appear at the bottom: repetition (left), mirror image (center), and progression (right). All stripes are red and white.*
Photograph by Don Cole; original in color.

rectangle. The type of intelligence required to weave, a historically ancient activity, was distinct from the type of intelligence that was useful at school or in commercial activities, historically newer activities. Across the generations, we see historical change in children's learning environments and in the way intelligence is tacitly defined. Through structural equation modeling, we were able to link learning environment change to tacit expressions of changing intelligence (Maynard et al., 2015).

From Embodiment to Abstraction in the Oksapmin Number System: The Role of Commerce

Saxe (1999) used different groups of adults studied at the same chronological time to link changing sociodemographics to changing learning environment to changing cognition. In the 1970s, the Oksapmin lived in a subsistence ecology of hunting and agriculture. Their number system was a totally embodied one. Body parts were used as numbers. For example, the same word was used for thumb on the right hand and for the number one. “Number cognition was tied to a specific context, the body; it was never abstracted from this context” (Greenfield, 2009, p. 408).

However, wage work on distant plantations with trade stores led to the introduction of trade stores into the Oksapmin environment. With trade stores came the introduction of money and commerce into the subsistence world of the Oksapmin, formerly based on farming and hunting.

> To adapt to the trade stores, Oksapmin people had to add and subtract for the first time. In this commercial environment, the contextualized system of using body-part names for numbers broke down. In adaptation, Oksapmin who were involved with trade stores as sellers or buyers started developing a slightly more abstract system that was usable for addition and subtraction; in this more decontextualized or abstract system, counting words were dissociated from the counter’s actual body parts. (Greenfield, 2009b, p. 408)

(This relation between Learning Environment Change and Developmental Change is represented by a vertical arrow between the two rectangles in Figure 38.2). As the historical niche changed from subsistence to commerce (top rectangle, Figure 38.2), successful adaptation, our definition of intelligence, changed from contextualized to more abstract thinking (Cultural Value rectangle, Figure 38.2).

Tacit Knowledge and Practical Intelligence

So far, I have not dealt with explicit conceptualizations of intelligence; instead, I have dealt with tacit knowledge. But Sternberg and Grigorenko (2000) make the point that tacit knowledge is an aspect of practical intelligence. The definition of intelligence as successful adaptation to a niche means that all intelligent behavior is practical insofar as it is adaptive in a particular niche. And I have shown that, as ecological niches changed, so did the nature of practical intelligence and tacit knowledge. However, I will now switch to explicit conceptualizations of intelligence.

Inferring Historical Change in the Kiganda Conception of Intelligence

In Uganda, Wober (1974) explored the Kiganda concept of intelligence in a number of Baganda participants differing in sociodemographic characteristics. (Kiganda is the adjective; Baganda, the people.) By comparing explicit conceptions of intelligence in groups with different sociodemographic characteristics, we can infer historical change in the concept of intelligence.

Obugezi was the Kiganda language’s word for intelligence. Village adults with limited primary schooling asserted that intelligent people are slow and careful; they also saw intelligence as spiritual, friendly, and public (Wober, 1974). Village teachers were a group with considerably more formal education. In contrast to other villagers, they saw intelligence as significantly more hurried, hasty, and unfriendly.

(This relationship between the Sociodemographic level and the Cultural Value level is diagrammed in Figure 38.2).

These data came from a small agricultural community where, before schooling was introduced by a colonial power, most, if not all, education was in context: learning to do tasks at home or in the fields, learning one's kinship network. So I would make the case that being slow and careful is a component of intelligence in Gemeinschaft worlds more generally. In contrast, the measurement of intelligence in the Gesellschaft world (e.g., IQ tests) puts a premium on speed. I would interpret the difference in friendliness between the two groups as reflecting a greater emphasis on social qualities in a more Gemeinschaft world.

An urban group in Kampala, consisting mainly of teachers, saw intelligence as less spiritual, less careful, and more unfriendly than the other two groups. This group, with the most educational qualifications and living in the city, have gone the furthest in inhabiting a Gesellschaft ecology (top level, right side, Figure 38.2) and their definition of intelligence is adapted to this ecology (second level, right side, Figure 38.2). Because urbanization and formal education were later cultural steps for the Baganda, introduced through colonization, we can infer that the villagers, inhabiting a rural setting and having little formal education, represent an earlier point in time and that the influence of schooling and urbanization has led to historical shifts in the Kiganda concept of intelligence. (This causal relationship is represented by the vertical arrow from the Sociodemographic rectangle to the Cultural Value rectangle, Figure 38.2).

N'glouèlê: Integrating Social Intelligence with Technological Intelligence

Mundy-Castle (1974/1975) distinguished between social intelligence valued in Africa and technological intelligence valued in the West. He emphasized that social intelligence incorporated technical skills insofar as they contributed to the community. Similarly, Dasen, studying a Baoulé village in Ivory Coast, emphasizes that the Baoulé concept of intelligence, *n'glouèlê,* integrates cognitive and social skills, as do many other African concepts of intelligence (Dasen, 2011). Indeed, the most central (in the sense of agreed on) attribute for intelligent children listed by illiterate Baoulé farmers was "readiness to carry out tasks in the service of the family and the community," a social quality (Dasen, 1984, p. 426). For adults, "speaking well in public, knowing how to use proverbs . . . and wisdom . . . are also part of intelligence, but they cannot really be expected of children" (Dasen, 1984, p. 427).

"The more technological skills have to be integrated with the social skills. The child's abilities are useless unless they are applied for the good and well-being of the social group. It is in this integration of social and cognitive attributes that the Baoulé definition of intelligence is most at variance with the Western and psychometric definitions" (Dasen, 1984, p. 427). "What is particularly valued is the initiative in carrying out a needed task without being asked. The second most important attribute is . . . respect of elders, politeness, and compliance" (Dasen, 1984, pp. 426–427). These top child attributes are all social.

So social attributes are more important in Gemeinschaft-adapted intelligence than in our Gesellschaft world. Nsamenang (2003, 2006) and Serpell (2011) make the point that social responsibility is a dimension of intelligence in village settings throughout Africa. Serpell (2017) succinctly notes that, in Zambia, the Chewa concept of *nzelu* (glossed as intelligence) combines smartness with social responsibility. (The primacy of social intelligence in Gemeinschaft ecologies is shown on the left side of the Cultural Value rectangle, Figure 38.2.) I would make the case that these forms of intelligence also have a historical dimension: that the integration of social and technological intelligence is valued in Gemeinschaft environments and that these environments historically preceded Gesellschaft environments, the inhabitants of which place greater value on technological intelligence. It is significant that, in Dasen's (1984) research among the Baoulé, standardized Piagetian tasks of concrete operations, a measure of developing intelligence, created in Switzerland, correlated very strongly with school grades (in a sample of seven- to nine-year-old children) and with an IQ measure developed in the West (a subtest of the Queensland test) but not at all with Baoulé parents' or Baoulé assistants' assessment of children's n'glouèlê. Again, given that schooling, a Gesellschaft influence, was overlaid through European colonization on a prior oral culture with informal education at home, that is, a Gemeinschaft ecology, we can conclude that the qualities of n'glouèlê preceded in historical sequence the qualities of intelligence developed in the school and assessed by Piagetian tasks and IQ tests (shown on the right side of the Cultural Value rectangle, Figure 38.2).

The Historical Introduction of Schooling Develops Concrete Operations, the Piagetian Definition of Intelligence, in Middle Childhood

In line with Dasen's findings, between eleven and thirteen years of age, only about half the unschooled Wolof children in a Senegalese village were successful in solving Piaget's concrete operational tasks; these children were receiving informal education in home and community settings. In sharp contrast, all the schoolchildren in this age group and from the same village succeeded at the tasks. This study was done shortly after independence from France, when formal education was just starting to develop in bush villages and many children still did not attend school. However, schoolchildren in the village performed at the same level as schoolchildren in Geneva, where the tasks were developed, or the United States (Greenfield, 1966). Again, if these Piagetian tasks are taken as measures of developing technological intelligence, we must conclude that success by young children in solving them reflects the historically later introduction of formal schooling through the colonization of Senegal by the French. Recall, too, that formal education is an important component of a Gesellschaft ecology, whereas informal education at home and in the community is a characteristic of a Gemeinschaft ecology. Hence, the sociodemographic shift to the availability of a school in the community (right side, Sociodemographic rectangle, Figure 38.2) availed some children the opportunity to go to school (right side, Learning Environment rectangle, Figure 38.2), leading them to success on Piaget's tests of concrete operations (right side, Individual Development rectangle, Figure 38.2). (These causal relations are represented by the vertical arrows from the Sociodemographic rectangle to the Learning Environment rectangle and from the Learning Environment rectangle to the Individual Development rectangle, Figure 38.2.)

Serpell (2011) is very direct about the war between indigenous village-based conceptions of intelligence as "an amalgam of cognitive alacrity and social responsibility" (p. 128) and the school's exclusive emphasis on the latter:

> Since contemporary Western-style schooling in many African countries tends to assess children's educational progress almost entirely in terms of cognitive skills and knowledge acquisition, these findings have been interpreted as reflecting a serious credibility gap for public basic education with respect to the values and aspirations of parents in rural African communities. (p. 128)

Looking at this war from a historical perspective, we can once again note that school, at its outset, was superimposed on African village life and that, therefore, its valued type of intelligence, deleting the social component, is a more recent phenomenon.

Signs That Wealth and Education Shift the Definition of Intelligence Away from Social Qualities

Grigorenko and colleagues (2001) provide evidence for this thesis. They found that intelligence for Luo living in an agricultural village in Western Kenya was defined in terms of two main components: smartness or knowledge (*rieko*) and social qualities such as respect and care for others, obedience, and diligence (*luoro*). Rieko was considered positive only if luoro was also present, so social goals were dominant in the definition of intelligence. (The primacy of social intelligence is shown on the left side of the Cultural Value rectangle, Figure 38.2.) Without luoro, it was considered that a child could use their rieko for selfish reasons and even against the interests of others. However, there were hints in the qualitative component of the study that education and wealth, two attributes of a Gesellschaft environment, were shifting the defining attributes of intelligence in the cognitive direction and away from the social: The only two participants ranking cognitive rieko higher than social luoro were outsiders to the local community who, unlike the villagers, had attained postsecondary education and were also much wealthier than the villagers. Looking at the situation from a historical perspective, both postsecondary education and wealth were introduced later in time; the Luo's original ecology featured informal education at home and in the community along with a subsistence economy. (These sociodemographic shifts are diagrammed in the top rectangle; the shift to technological intelligence as primary is diagrammed in the Cultural Value Change rectangle; and the two levels are linked with a vertical causal arrow, Figure 38.2.)

A second study (Sternberg et al., 2001) explored the relationship between highly adaptive and contextualized knowledge in the same rural Luo village and tests of academic intelligence. The domain of adaptive knowledge was herbal medicines used to fight illness, necessary in an environment in which most children had parasitic infections at any one point in time. Scores on the test of knowledge of parasites were negatively correlated with all tests of academic intelligence. Hence, they are on opposite sides of the Individual Development rectangle (Figure 38.2). Wealth also had a significant negative correlation with the knowledge of herbal medicines. Again, given that schooling and wealth were overlaid on a culture that already had knowledge of herbal medicine, we can infer that the influence of schooling and wealth will continue to reduce detailed practical knowledge of herbal medicines. Indeed, World Bank (2018a, 2018b) data indicate that both educational opportunity and wealth are continuing to rise in Kenya. These two sociodemographic shifts are shown in the top rectangle of Figure 38.2. A vertical arrow shows the hypothesized causal relationship between these shifts and the shift from knowledge of herbal remedies to academic intelligence in the bottom rectangle.

Crystallized Intelligence: The Case of Know and Na'

Crystallized intelligence is the ability to use skills, knowledge, and experience (Cattell, 1941). Knowledge is therefore a component of intelligence measurement; it is at the heart of the contrast between *know* and *na'*, the Tzotzil word for "know" in the Maya community of Mitontik in Chiapas Mexico (Zambrano & Greenfield, 2004).

> Although na' clearly glosses as 'know' (Laughlin, 1975), and even overlaps with it, its core meanings are surprisingly different. Na is much more demanding in key respects, such as in its reference to practice. However, in a world in which cultures have been in close contact – through involuntary processes such as conquest, voluntary processes such as immigration, and systemic processes such as economic globalization – different ethnoepistemologies can also come into contact. And this is exactly what has happened to na' and know in the Tzotzil-speaking community of Mitontik. Na' . . . epitomizes indigenous values concerning knowledge, whereas know is highly valued in the school, an institution that has been imposed on Maya communities from outside . . . Whereas to 'know' in English always involves the mind, na' often involves the heart and soul . . . Whereas 'knowing' connotes factual knowledge, theoretical understanding, or know-how, na' also connotes knowledge of practice that is habitual and characteristic of a given person; it is very much akin to character. The former type of knowledge is more important in a culture placing a value on social character. Both forms of knowledge coexist in San Miguel Mitontik; however, na', a Tzotzil word, originates in the indigenous Maya culture and is traditionally valued at home. 'Know' (or saber in Spanish) originates in the school, imposed on Mayan communities by the Spanish-speaking Mexican state, the institutional inheritance of the Spanish conquest. (Zambrano & Greenfield, 2004, pp. 252–253)

"Know" is therefore historically more recent.

The word "know" is what the fields of intelligence testing and cognitive psychology care about: the solving of a novel problem once. The word *na'* requires that a problem be solved habitually and repeatedly implemented. “The concept of *na* . . . embraces a broader conception of knowledge (and thus intelligence) that presses us to admit that the academic world produces but a small amount of the knowledge and intelligence in the world” (Zambrano & Greenfield, 2004, pp. 268–269). The historical shift from *na'* to “know” with formal education is diagrammed in the Sociodemographic and Cultural Value rectangles (Figure 38.2), with a vertical arrow representing their causal link.

Historical Change in Cultural Tools: Implications for the Raven’s, a Test of Fluid Intelligence

I now move from crystallized to fluid intelligence and an analysis of the connection between social change and the Raven’s Progressive Matrices test. I again start with Nabenchauk, the Zinacantec Maya community in Chiapas, Mexico. As noted in the section “From Tradition and Task-Relevant Detail to Innovation and Abstract Representation,” over a period of two decades, Nabenchauk shifted from an agrarian, subsistence ecology, where almost all education took place at home and in the community, to a commercial ecology in which schools became increasingly important. The IQ test, the Raven’s Progressive Matrices, depends on understanding the structure of a matrix; this is cultural knowledge (Greenfield, 1998). To solve matrix problems, such as those presented on the Raven’s, one needs to understand that a matrix is organized in rows and columns. One must also understand that there is an ordinal relationship among the columns and rows, as well as what mental operations are relevant to perform on the test matrix. All of this is culture-specific knowledge; there is nothing in the matrix figures themselves that specifies what mental operations to perform.

As a subsistence community whose residents did not read and write their spoken language of Tzotzil, Nabenchauk had no matrices in its environment. Matrices in the form of cross-stitch patterns for embroidery laid out on graph paper were introduced into the agrarian community by school teachers. Clearly the skills measured by the Raven’s could not have been part of the definition of intelligence in Nabenchauk before any matrices existed in the environment. Our research indicated an association between the use of these patterns and schooling: Zinacantec women who had a few years of schooling were more likely to use these patterns for embroidery or weaving than women who had never been to school (Maynard & Greenfield, 2008). These patterns are a very simple form of matrix; unlike the Raven’s, they do not involve any ordinal relationship among columns or rows. However, they form a foundation that could serve as a basis for understanding ordered matrices and for defining intelligence in terms of the mental manipulation of matrices – that is, the use of the Raven’s as a measure of intelligence.

Fast forward to the United States and other countries with widespread and highly developed computer technologies. As part and parcel of the technology, the use of matrices has become increasingly diffused in the population. An example is the popular spreadsheet program Microsoft Excel; Excel provides blank matrices, organized in columns and rows, to be filled in by the user. “Clearly, such a program requires users to represent their data mentally in matrix form, while providing practice in the use of this representational format” (Greenfield, 1998, p. 110).

Thus, it has become increasingly relevant to US culture to utilize the Raven’s as an intelligence measure; this increasing relevance could have produced a part of the Flynn effect, the historical increase in IQ scores that shows up in the historical increase in the Raven’s (for more about the Flynn effect, see Flynn, Chapter 39 in this volume). As a hypothetical, one can imagine that, as formal education and computers continue to develop in Nabenchauk and other Zinacantec communities, trends that are currently taking place (Manago & Pacheco, 2019; Maynard et al., 2015), manipulation of matrices could come to be one measure of intelligence in the Nabenchauk of the future, as well as in other communities around the world in which the expansion of formal education and diffusion of computer technologies are moving environments ever more in the Gesellschaft direction. There is a close relationship between a culture’s technologies and its definition of intelligence (Maynard, Subrahmanyam, & Greenfield, 2005). Hence, as technology develops, we can expect parallel developments in the definition of intelligence, specifically a link between computer technology (top right side of the Sociodemographic rectangle) and this test of fluid intelligence (right side of Cultural Value rectangle). (This causal connection is depicted by the vertical arrow linking the Sociodemographic rectangle with the Cultural Value rectangle, Figure 38.2.)

Historical Change in How Intelligence Is Defined in a Leading IQ Test

Intelligence in our society has often been defined as “what the tests measure.” Apart from academic dissections of the nature of intelligence, we can make a case that popular intelligence tests measure what the society thinks is important. It is therefore instructive to look at how the most venerable children’s intelligence test, the WISC or Wechsler Intelligence Scale for Children, has been changing over recent decades. We will see that, even in our society, Gesellschaft factors are moving our functional definition of intelligence in the direction of decreasing emphasis on social understanding, decreasing emphasis on detailed knowledge, increasing emphasis on abstraction, and increasing emphasis on dealing with novelty.

In line with the tenet of the theory of social change and human development – that the sociodemographic factor changing most rapidly in a community or society will drive changing patterns of human development – technology (top level of Figure 38.2) has created new patterns of cognitive development, strengthening skills that are important in adapting to the spread and enhancement of technologically mediated communication (Greenfield, 1993, 2009a). For example, media such as video games have developed visuospatial skills; and communication has become more abstract (in the sense of removed from the physical world) because of its virtuality (Greenfield, 2019). I am going to make the case that these changes in cognitive skills have also affected our measurement and therefore our societal definition of intelligence.

By 2003, the publication date of the WISC-IV, we were clearly in the midst of a technological revolution, and some very revealing changes took place in the test: Picture arrangement, the only subtest that assessed understanding of social situations and social actions, was dropped from the WISC-IV. Mazes and Object Assembly were also dropped; these tests both involved analysis of a present stimulus situation. In the words of Pearson, the publisher of the WISC:

> Compared to the WISC–III, the WISC–IV FSIQ deemphasizes crystallized knowledge . . . and increases the contribution of fluid reasoning (Matrix Reasoning and Picture Concepts) . . . and Processing Speed (both Coding and Symbol Search). (Pearson, 2010)

Here we see that novelty (labeled as fluid reasoning) and speeded cognition are becoming ever more important in the definition of intelligent qualities to be assessed by the tests (diagrammed on right side of the Cultural Value rectangle, Figure 38.2).

In the following Pearson quote, we see the continuation of another historical trend that we found in Nabenchauk. In the substitution of a Picture Concepts subtest for the older Picture Completion subtest, attention to visual detail is being devalued as a lower-order cognitive ability in comparison with fluid reasoning:

> Picture Completion and other, more traditional measures of perceptual ability, measure visual discrimination and attention to visual detail, which is a lower order cognitive ability than fluid reasoning. (Pearson, 2010)

In other words, in the latest edition of the WISC, understanding of social situations is eliminated as part of assessing intelligence; understanding concrete here-and-now situations is deemphasized; attention to visual detail is considered a lower form of intelligence; while the manipulation of abstractions and processing speed are given greater emphasis. Shifts in defining intelligence that were originally noted when schooling was introduced into Africa have continued and intensified in the United States (Cultural Value level, Figure 38.2).

Recall that, over time, in Zinacantec Maya children's approach to representing woven patterns, the attention to detail required for constructing the woven patterns, gave way to abstract representations that would be useless in creating the patterned cloth. That historical trend was fueled by the development of commerce and formal education (Maynard et al., 2015). These factors are now constants in our US ecology of the last forty years (Huang et al., in prep.); instead, the major change in our ecology is technology. My hypothesis is that technological development and diffusion are now moving our definition of intelligence ever further in the direction of novel and abstract cognition. Indeed, fluid intelligence is defined as the capacity to reason and solve novel problems, independent of any knowledge from the past. It is the ability to analyze novel problems, identify patterns and relationships that underpin these problems, and the extrapolation of these using logic. So an increasing emphasis on fluid intelligence in defining the construct in tested intelligence equates to an increasing emphasis on innovative thought and "going beyond the information given" (Bruner, Goodnow, & Austin, 1956). This is exactly the historical trend we saw in the Zinacantec hamlet of Nabenchauk with the growing pattern and other novel patterns, there fueled by the growth of commerce and formal education. This is an example of the equipotentiality of Gesellschaft features (right side of the Sociodemographic rectangle, Figure 38.2).

The fifth edition of the Wechsler Intelligence Scale for Children, WISC-V, has a new Visual-Spatial Index, which measures a child's ability to reason in nonverbal tasks such as rotating and organizing shapes. These are exactly the skills developed by action video games (Greenfield, 1998). The changes also reflect other cognitive changes that are adaptations to the new technologies: As representation becomes more iconic and less symbolic (Greenfield et al., 1994), Picture Span (short-term memory for visual images) has supplemented Digit Span (short-term memory for numerals). The growing cultural importance of iconic representation is portrayed on the right side of the Cultural Value rectangle, Figure 38.2).

Conclusions

In sum, the history of the nature of intelligence has a globalized direction of change under the influence of the sociodemographic changes shown in Figure 38.2: Ever more technology, urbanization, formal education, wealth, and commercialized economies. The direction of change in valued intelligence, summarized in Figure 38.2, is from the integration of social responsibility, wisdom, and spirituality with cognitive intelligence toward purely cognitive skills; from practical, detailed, and contextualized to abstract, decontextualized cognition; from slow and careful thinking to speeded cognition; from repetition of the known to extrapolation and novelty; from habitual practice to innovation; and, using the language of IQ tests, from crystallized to fluid intelligence.

However, as Sternberg and Grigorenko (2000) have emphasized, practical context-specific intelligence based on tacit knowledge continues to exist alongside more abstract intelligence, even in the United States – for example, Jean Lave and colleagues' work on calculations made while shopping in grocery stores (Lave, Murtaugh, & de la Rocha, 1984) and Silvia Scribner's (1984) research on calculations made by dairy workers. Each setting and function stimulated the development of a specific arithmetic technique that was neither generalizable nor taught in school. In the domain of social intelligence, social responsibility and respect as *desiderata* continue to be important qualities to groups within the United States that are relatively poor and have lacked opportunity for formal education (Greenfield & Quiroz, 2013; Vasquez-Salgado, Greenfield, & Burgos-Cienfuegos, 2014). Social and practical intelligence are basic qualities that were adaptive in human beings' evolutionary ecology. Wisdom is desperately needed to solve today's societal problems (Sternberg, 2018). These qualities are critical to the human condition. However, so long as human technology, wealth, urbanization, commerce, and education continue to expand, these characteristics of human intelligence are in grave danger of receiving ever less respect and development as the future unfolds.

References

Berry, J. W. (1974). Radical cultural relativism and the concept of intelligence. In J. W. Berry & P. R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 225–229). London: Methuen.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin*, *38*, 592.

Cole, M., Gay, J., Glick, J. A., & Sharp, D. W. (1971). *The cultural context of thinking and learning*. New York: Basic Books.

Dasen, P. R. (1984). The cross-cultural study of intelligence: Piaget and the Baoulé. *International Journal of Psychology*, *19*, 407–434.

Dasen, P. R. (2011). Culture, cognition and learning. In A. B. Nsamenang & T. M. S. Tchombe (Eds.), *Handbook of African educational theories and practices* (pp. 159–174). Bamenda: Human Development Resource Center.

Glick, J. (1968). Cognitive style among the Kpelle. Paper presented at the meeting of the American Educational Research Association, Chicago.

Glick, J. (1974). Cognitive development in cross-cultural perspective. In E. M. Hetherington (Ed.), *Review of child development research* (pp. 595–654). Chicago: SRCD.

Greenfield, P. M. (1966). On culture and conservation. In J. S. Bruner, R. R. Olver, P. M. Greenfield, et al., *Studies in cognitive growth* (pp. 225–256). New York: Wiley.

Greenfield, P. M. (1974). Cross-cultural research and Piagetian theory: Paradox and progress. *Dossiers Pedagogiques*, 34–39. English version in K. F. Riegel & J. A. Meacham (Eds.) (1976). *The developing individual in a changing world*, Vol. 1, *Historical and cultural issues* (pp. 322–333). Paris: Mouton.

Greenfield, P. M. (1993). Representational competence in shared symbol systems: Electronic media from radio to video games. In R. R. Cocking & K. A. Renninger (Eds.), *The development and meaning of psychological distance* (pp. 161–183). Hillsdale, NJ: Lawrence Erlbaum Associates.

Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*, *52*, 1115–1124.

Greenfield, P. M. (1998). The cultural evolution of IQ. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 81–123). Washington, DC: American Psychological Association.

Greenfield, P. M. (1999). Cultural change and human development. In E. Turiel (Ed.), *Development and cultural change: Reciprocal processes*. New Directions in Child Development, No. 83 (pp. 37–60). San Francisco: Jossey-Bass.

Greenfield, P. M. (2009a). Technology and informal education: What is taught, what is learned. *Science*, *323*, 69–71.

Greenfield, P. M. (2009b). Linking social change and developmental change: Shifting pathways of human development. *Developmental Psychology*, *45*, 401–418.

Greenfield, P. M. (2016) Social change, cultural evolution, and human development. *Current Opinion in Psychology*, *8*, 84–92.

Greenfield, P. M. (2018). Studying social change, culture, and human development: A theoretical framework and methodological guidelines. *Developmental Review.*

Greenfield, P. M. (2019). Communication technologies and social transformation: Their impact on human development. In R. Parke & G. Elder (Eds.), *Children in changing worlds:*

Sociocultural and temporal perspectives (pp. 235–273). Cambridge: Cambridge University Press.

Greenfield, P. M., & Childs, C. P. (1977). Weaving, color terms and pattern representation: Cultural influences and cognitive development among the Zinacantecos of Southern Mexico. *Inter-American Journal of Psychology*, *11*, 23–48.

Greenfield, P. M., Maynard, A. E., & Childs, C. P. (2003). Historical change, cultural learning, and cognitive representation in Zinacantec Maya children. *Cognitive Development*, 18, 455–487.

Greenfield, P. M., & Quiroz, B. (2013). Context and culture in the socialization and development of personal achievement values: Comparing Latino immigrant, families, European American families, and elementary school teachers. *Journal of Applied Developmental Psychology*, *34*,108–118.

Grigorenko, E. L. et al. (2001). The organization of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavioral Development*, *25*, 367–378.

Hansen, N., Koudenburg, N., Hiersemann, R., Tellegen, P. J., Kocsev, M., & Postmes, T. (2012). Laptop usage affects abstract reasoning of children in the developing world. *Computers in Education*, *59*, 989–1000.

Huang, V. B., Greenfield, P. M., Zhou, C., & Wu, M. (in prep.). Intergenerational differences in child behaviors and parent socialization in United States.

Laughlin, R. (1975). *The great Tzotzil dictionary of San Lorenzo Zinacantán*. Washington, DC: Smithsonian Institution Press

Lave, J., Murtaugh, M., & de la Rocha, O. (1984). The dialectic of arithmetic in grocery shopping. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 67–94). Cambridge, MA: Harvard University Press.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.

Manago, A. M., & Pacheco, P. (2019). Globalization and the transition to adulthood in a Maya community in Mexico: Communication technologies, social networks, and views on gender. *New Directions in Child and Adolescent Development*, *164*, 1–15.

Maynard, A. E., & Greenfield, P. M. (2008). Women's schooling and other ecocultural shifts: A longitudinal study of historical change among the Zinacantec Maya. *Mind, Culture, and Activity*,*15*, 165–175.

Maynard, A. E., Greenfield, P. M., & Childs, C. P. (2015). Developmental effects of economic and educational change: Cognitive representation across 43 years in a Maya community. *International Journal of Psychology*, *50*, 12–19.

Maynard, A. E., Subrahmanyam, K., & Greenfield, P. M. (2005). Technology and the development of intelligence. In R. J. Sternberg & D. Preiss (Eds.), *Intelligence and technology: The impact of tools on the nature and development of human abilities* (pp. 54–97). Mahwah, NJ: Lawrence Erlbaum Associates.

Mundy-Castle, A. C. (1974). Social and technological intelligence in Western and non-Western cultures. *Universitas*, *4*, 46–52. Reprinted in S. Pilowsky (Ed.) (1975). *Cultures in collision*. Adelaide: Australian Association for Mental Health.

Nsamenang, A. B. (2003). Conceptualizing human development and education in subSaharan Africa at the interface of endogenous and exogenous influences. In T. S. Sarasthwathi (Ed.), *Cross-cultural perspectives in human development: Theory, research, and applications* (pp. 213–235). New Delhi: Sage.

Nsamenang, A. B. (2006). Human ontogenesis: An indigenous African view on development and intelligence. *International Journal of Psychology*, *41*, 293–297.

Pearson. (2010). WISC-IV – Frequently asked questions. Pearson Education. http://images.pearsonclinical.com/images/Products/WISC-IV/4326-10_WISC_FAQ_pdf_f.pdf

Santos, H. I., Varnum, M. E. W., & Grossmann I. (2017). Global increases in individualism. *Psychological Science*, *28*, 1228–1239.

Saxe, G. (1999). Cognition, development, and cultural practices. In E. Turiel (Ed.), *Development and cultural change: Reciprocal processes. New Directions for Child and Adolescent Development, no. 83* (pp. 19–35). San Francisco: Jossey-Bass.

Scheibel, A. (1996). Introduction to symposium on the evolution of intelligence. Center for the Study of Evolution and the Origin of Life, University of California, Los Angeles.

Scribner, S. (1984). Studying working intelligence. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context* (pp. 9–40). Cambridge, MA: Harvard University Press.

Serpell, R. (2011). Social responsibility as a dimension of intelligence, and as an educational goal: Insights from programmatic research in an African society. *Child Development Perspectives*, *5*, 126–133.

Serpell, R. (2017). How the study of cognitive growth can benefit from a cultural lens. *Perspectives on Psychological Science*, *12*, 889–899.

Sternberg, R. J. (2018). Why real-world problems go unresolved and what we can do about it: Inferences from a limited resource model of successful intelligence. *Journal of Intelligence*.

Sternberg, R. J., & Grigorenko, E. L. (2000). Practical intelligence and its development. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 215–243). San Francisco: Jossey-Bass.

Sternberg, R. J., Nokes, C., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., Grigorenko, E. L. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence*, *29*, 401–418.

Tönnies, F. (1887/1957). *Community and society (Gemeinschaft und Gesellschaft*) (trans. C. P. Loomis). East Lansing: Michigan State University Press.

Vasquez-Salgado, Y., Greenfield, P. M., & Burgos-Cienfuegos, R. (2014). Exploring home-school value conflicts: Implications for academic achievement and well-being among Latino first-generation college students. *Journal of Adolescent Research*, *30*, 1–35.

Wober, M. (1974). Towards an understanding of the Kiganda concept of intelligence. In J. W. Berry & P. R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 261–280). London: Methuen.

World Bank. (2018a). Data (wealth). World Bank. https://data.worldbank.org/country/kenya

World Bank. (2018b). Data (secondary education). World Bank. https://data.worldbank.org/indicator/SE.SEC.ENRR?locationsKE

Zambrano, I., & Greenfield, P. (2004). Ethnoepistemologies at home and at school. In R. J. Sternberg & E. L. Grigorenko (Eds.), *Culture and competence: Contexts of life success* (pp. 251–272). Washington, DC: American Psychological Association.

39 Secular Changes in Intelligence

The "Flynn Effect"

James R. Flynn

Until 1984, most believed that mean IQ is relatively stable over time, at least in developed nations. There had been reports of IQ gains among seemingly exotic groups – canal children in Britain, Tennessee mountain children, second-generation immigrants in Honolulu – but these seemed to be cases of extreme environments undergoing radical change. In 1948, Tuddenham published the first study using a nationwide sample (Tuddenham, 1948). He compared US military data from World War I and World War II and found about a full standard deviation (SD = 15 IQ points throughout) increase on the Army's mental tests. Because he interpreted these as almost entirely due to extra years of schooling, it seemed to raise no problems about the theory of intelligence. The gains were dismissed as if they had been on academic achievement tests rather than intelligence tests.

In 1982, Richard Lynn found gains on the WISC (Wechsler's Intelligence Scale for Children) in Japan. That these had occurred in a developed nation should have posed the question: Could this be an international phenomenon? But, at that time, there was a perception that the Japanese were unusual and that their gains merely underlined their superiority. However, in 1984, Flynn showed that Americans had also made massive gains: fourteen IQ points over forty-six years (1932 to 1978) (Flynn, 1984).

Practical Implications

Flynn emphasized that, whether or not these were "intelligence" gains, they posed practical problems by acting as a confounding variable. The longer the time between when the test was normed and when it was administered, the more IQ scores were inflated. If you used a test normed on a representative sample from 1932, their general performance (and the norms they set) were much weaker than if you used a test based on a representative sample from 1972. Thanks to twelve IQ points gained over those forty years, a group of subjects merely average (mean 100) on the newer test would have a mean IQ of 112 on the older test.

Obsolete Norms and False Inferences

The inflation of IQs because of obsolete norms led to inflated estimates of the effects of intervention, adoption, and aging; and also misdiagnosis of whether individuals had met IQ cutting lines that affected everything from the administration of the death penalty to who should benefit from special education.

Garber (1988) conducted the famous “Milwaukee Project.” He selected a sample of inner-city black children, the children of low-IQ mothers seemingly destined to wear the label “mentally retarded,” in order to see if radical intervention could boost their IQs. He used proper controls to measure the effects. But others focused on the sheer magnitude of the measured IQs (above 120 in infancy) and were dismayed at a precipitous “IQ decline” at the end of the program. Both of these were artifacts of obsolete norms (corrected in his book). The infant IQs were the product of the Stanford-Binet whose norms from 1932 inflated scores by twelve points thanks to forty years of obsolescence; the decline at age six was the product of WISC norms from 1947 to 1948 that were more recent than those of the Stanford-Binet (by almost sixteen years) and therefore inflated scores less.

The message is that no IQ score is significant unless the test is named, the date of norming given, and the date of administration given, so the number of years of obsolescence can be taken into account. This is done, ideally, by using the rate of IQ gain between the date the test was standardized and the date it was administered. In America since 1932, you should deduct about 0.3 points per year. In some nations for some tests, the rate of gain is unavailable and adjustment of inflated scores is impossible.

Even those who should be most concerned, those who are interested in whether the scores of children put them below the cutting line for mental retardation, are often remiss. Almost 20 years after the 1984 article, Kanaya, Scullin, and Ceci (2003) demonstrated, state by state, how the introduction of more up to date test manuals had radically increased the number of children being diagnosed as suffering from mental retardation. Faced with tougher norms, more of them were scoring at 70 or below. Since the school psychologists did not know what was happening, many simply ignored the later (and more accurate) scores as “deceptive.” As Laid and Whitaker (2011) show, failure to adjust obsolete norms has continued in dozens of articles right into the new millennium.

In addition, there are historically important studies whose results are still cited in the literature as if they could be taken at face value – even though adjustment of the test scores for inflation would radically qualify their results. Flynn (1993) selected the famous adoption study by Skodak and Skeels (1949) as a case in point. When adjusted for the fact that the natural mothers and their adopted children were scored on tests of different obsolescence, the IQ advantage between the adopted child and the natural mother was reduced by a third.

Obsolete Norms and Death

The US Supreme Court has held, in effect, that a capital offender whose IQ on a reliable test places him in the bottom 2.27 percent of the population has a prime facie case of being exempted from the death penalty. That equates with an IQ score of seventy or below, the criterion for mental retardation, a condition that is taken as evidence of diminished responsibility. Ideally, the offender was tested at school prior to the age of eighteen. But, particularly when such scores are not available or when they seem to vary, he is tested while on death row. IQ gains on the Stanford-Binet and Wechsler IQ tests have occurred in America at a steady rate of 0.30 IQ points per year ever since 1947 (Flynn, 2012a). Therefore, if a child took an obsolete test at school, it would inflate his IQ by 0.30 times the number of years between when the test was normed and when the test was administered. This could easily boost his IQ above the seventy cutting line.

When obsolete norms boost a convicted capital offender above the IQ cutting line, this creates an inequity that has nothing to do with whether IQ gains imply that we should set a new standard of mental retardation. Flynn and Widaman (2007) show that, no matter what the standard, failure to correct for obsolescence turns the death penalty into a lottery. You win survival if you were fortunate enough to take a current test and get sixty-eight; and you lose your life if you were unfortunate enough to take an obsolete test and get seventy-five. The remedy, of course, is to reduce obsolete sores by deducting 0.3 IQ points for every year between the norming of the test and its administration. Frumpkin (2003) notes the alacrity of defense lawyers to acknowledge this inequity. It is still resisted by the prosecution but it has slowly gained ground among scholars, such as Gresham and Reschly (2011) and Young (2012). Shalock (2012) recognizes the problem of obsolescence in the AAIDD (American Association for Intellectual and Developmental Disability) User's Guide.

US Federal Courts of Appeal have gradually acknowledged the Flynn effect as "relevant" in capital cases. The major recent cases are Moore v. Quarterman 2006; US v. Parker 2007; Holliday v. Allen 2009; Thomas v. Allen 2010; and Walker v. Kelly 2010. This is despite the fact that revising IQs does not sit easily with prevailing rules of evidence, as noted in Winston v. Kelley (2008). Flynn (2009a) attempts to reconcile the two.

Qualifying for Other Things

When people are scored against obsolete norms, their inflated IQ scores can put them above any IQ cutting line that confers either benefits or penalties. A few may benefit in that they qualify for classes for the gifted or service in the military. The many that suffer are those denied welfare because of mental disability and those denied special education at school. Flynn (1998) discusses the negative side of using IQ scores to label children. Kanaya and colleagues (2003) discuss a whole range of cutting lines relevant to special education, disability benefits, military service, and so forth. Flynn (1985 & 2009a) addresses the separate question of whether IQ gains by those in the bottom 2.27 percent of the population should bring a reappraisal of the cutting line for mental retardation. This has received surprisingly little attention.

Thanks to Baxendale (2010) and Rönnlunda and Nilsson (2009), we now know that other diagnostic instruments such as memory tests are equally compromised by obsolete norms. Memory tests supplement IQ tests to determine whether patients have lost cognitive skills during surgery. Baxendale asks whether IQ tests are not merely the tip of the iceberg – with other diagnostic instruments to follow. Her results for England were replicated in Sweden, so the obsolescence of memory tests is an international phenomenon.

Causes of IQ Gains

There are several positions about causality. First, Brand believes that most (or much) of the gains are due to enhanced test sophistication and therefore lack cognitive significance (Brand, Freshwater, & Dockrell, 1989; Brand, 1990). Second, others believe that, insofar as the gains are real, they must be based on brain physiology. For example, Lynn (1998) emphasizes better nutrition particularly in childhood. Mingroni (2007) argues for less inbreeding as productive of better brains.

Flynn believes that virtually all of the gains have cognitive significance. While acknowledging genetic and biological causes, he believes that the sheer magnitude of the gains suggests cultural factors on various levels. The ultimate (or root) cause is the industrial revolution or modernity. The intermediate causes are the industrial revolution's by-products such as smaller families, parental pressure on cognitive development, enhanced schooling, more cognitively demanding jobs, and more cognitively demanding leisure. The proximate causes are the new "habits of mind" people take with them into the test room, so they can answer more items correctly, such as freeing their minds from the concrete and using logic on abstractions, taking the hypothetical seriously, and classifying particulars. Finally, there are articles that isolate highly specific causes as productive of enhanced performance on particular IQ tests. These positions will be dealt with in turn.

Test Strategy

Brand argues that personality changes are the primary cause of IQ gains. Over time, people have become less responsible and, today, they care less about whether their answers are correct. When they think they know the right answer, they immediately respond rather than spending time pausing to reflect. Therefore, they are less likely to run out of time and leave fewer answers blank. Moreover, when they encounter items about which they are uncertain, they are less inhibited about "intelligent guessing." They select often-correct answers, while the older generation would not tick an answer unless they were sure. He stresses that these differences would affect performance primarily on multiple-choice tests like Raven's particularly if there were a time limit. He nominates Wechsler tests (largely untimed and personally administered) as a better measure and cites what he considers to be minimal gains on the Scottish WISC.

Brand and colleagues (1989), Flynn (1990), and Brand (1990) debated whether or not the Scottish WISC gains were large or small. Flynn argued that Brand's method, which utilized low correlations between WISC items, vastly underestimated gains. He first used the Stanford-Binet method of dividing mental age by chronological age, which put the gains at fifteen to seventeen points over 22.5 years. Converting these into deviation IQs lowered gains to about twelve to thirteen points or a rate of gain at 0.56 points per year, actually larger than most nations. Flynn also emphasized the existence of large Wechsler gains elsewhere.

Two studies from Estonia give conflicting results for and against the role of test sophistication, namely Shiu and colleagues (2013) and Must and Must (2013). The latter argues that, Brand aside, people over time can improve their test-taking strategy simply because tests become more frequent and more important. Pietsching, Tran, and Voracek (2013) shows that this did not occur in Austria. Flynn (2013) remarks that the enormous gains in the Netherlands, twenty IQ points between 1952 and 1982, occurred after their society was saturated with testing. He concludes that test sophistication is more likely to inflate gains in Eastern Europe because testing there only increased gradually since World War II, as distinct from Western European nations (such as Austria) and English-speaking nations, nations where testing has been long established. Fox and Mitchum (2012) showed beyond doubt that Raven's gains are due primarily to how people's minds had altered during the course of the twentieth century, as we shall see shortly.

Enhanced Brains

Anything that has affected the quality of the human brain during the twentieth century bears partial responsibility for IQ trends over time. Woodley, te Nijenhuis, and Murphy (2013) argue that the brain's genetic potential has degenerated since Victorian times based on slower reaction times (how quickly a subject can respond to a visual stimulus), that fewer people are being recognized as geniuses, and that there is dysgenic reproduction (the tendency for the better educated to have fewer children than the less educated). Flynn (2013) noted controversy over whether reaction-time equipment was equally sensitive over time and over the subjectivity of the "genius" classification.

There is little doubt about the negative effect of dysgenic mating and it may have amounted to a loss of five IQ points since Victorian times. This would mean that the environmental factors that caused IQ gains are even more potent than they seem. The twentieth century saw thirty IQ points gained in advanced nations, which means that thirty-five points would have been gained if negative genetic trends were absent. Bratsberg and Rogeberg (2018) used within-family military data from Norway to estimate trends over a period of sixteen years for twenty-year-olds. They were tested during a transitional period (during which gains changed to losses in 1995). The results mimicked IQ trends as normally measured (from one cohort to another). The within-family method eliminates dysgenic trends as possible causes (any such would affect brothers equally) but, as the authors state, such trends might be too small to be statistically detectable. Five points over 150 years would give only 0.53 points over sixteen years.

Environmental factors include those that would have a direct positive effect on the maturing brain such as better nutrition and health care. The only real controversy is whether these factors have been prominent in advanced nations since 1950. As for nutrition, Martorell (1998) stresses enhanced nutrition as an important cause prior to 1950; Lynn (1998) and Storfer (1990) extend it to post-1950. Flynn (2012a) sums up the case for little post-1950 impact. He emphasizes: UK Raven's trends that are not correlated with nutritional trends; international data showing that those with higher IQs did not outgain those with lower IQs (the assumption is that nutrition gains would have affected primarily lower socioeconomic status, or SES, people);

Norwegian data that, while high IQ subjects made lower IQ gains, they also made greater height gains. This last suggests that where low-IQ subjects made greater gains, nutrition was not the answer: they may have made greater *schooling* gains than the upper classes. However, when we assess different IQ gains between age groups, we must say more about the effects of nutrition: improved diet in recent years has had a dramatic impact on the aged.

Jensen (1998) cites better prenatal care and perinatal progress (improved technique at time of birth) as obvious possible causes of small gains. Sir Michael Rutter (2000, p. 223) replies that these improvements have had no net positive effect on mean IQ. For every child who has escaped mental impairment, one or more impaired children have been saved, particularly among those born prematurely.

Steen (2009) offers an encyclopedic account of how much curing US ills would raise IQ, all the way from eliminating poverty, low birth weight, lead poisoning, childhood neglect, untreated illnesses, and fetal alcohol syndrome. He gets a total of five IQ points. Needless to say, America did not eliminate these factors between 1950 and 2000. He has no pre-1950 data and grants that he cannot match health gains to the history of IQ gains.

Mingroni (2007) has been the principal proponent of less inbreeding as productive of better brains. Too much inbreeding matches negative genes during sexual reproduction for all traits including intelligence. Flynn (2012a) argues the case against inbreeding as a significant factor in America all the way back to 1870. He uses a mathematical analysis of US census data and Japanese data on the effects of less inbreeding to show that the latter could explain only a fragment of a point of IQ gains. Woodley (2011) also offers a strong rebuttal.

Society and Its People

Among those who believe social change has been the primary cause of IQ gains, there is general agreement on *ultimate* causes (the industrial revolution and economic progress) and *intermediate* causes (spin-offs such as smaller families, hothouse child-rearing, enhanced schooling, more cognitively demanding jobs, and more cognitively demanding leisure – although there is controversy about the new visual world of the Internet). There is less agreement about *proximate* causes, that is, just what new mental awareness people take with them into the test room so they can answer more items correctly. Indeed, even among those who endorse cultural causes, few speculate about new modes of thinking.

Genovese (2002) is an exception. He goes beyond the fact that we have added more years of schooling to show that modernity has altered how educators shape the minds of schoolchildren. He compared the exams the state of Ohio gave to fourteen-year-old schoolchildren between 1902 and 1913 with those they gave between 1997 and 1999. The former tested for in-depth knowledge of culturally valued information; the latter expected only superficial knowledge of such information and tested for understanding complex relationships between concepts. The former were likely to ask you to name the capitals of the (then) forty-six to forty-eight states. The later exams tended to ask you why the largest city of a state was rarely the state capital (rural members dominated state legislatures, hated the big city, and bestowed the capital on a rural town). Genovese (2002, p. 101) concludes: “These findings suggest that there have been substantial changes in the cognitive skills valued by Ohio educators over the course of the 20th century.”

The Piagetian literature illuminates the kind of mind that modernity engendered. Luria interviewed isolated rural people in Russia in the 1920s. These were people who were like Americans in 1900 with little formal education.

First interview: Fish and crows (Luria, 1976, p. 82)

Q: What do a fish and a crow have in common?

A: A fish lives in water. A crow flies. If the fish just lies on top of the water, the crow could peck at it. A crow can eat a fish but a fish can’t eat a crow.

Q: Could you use one word for them both?

A: If you call them both “animals,” that wouldn’t be right. A fish isn’t an animal and a crow isn’t either. A crow can eat a fish but a fish can’t eat a bird. A person can eat a fish but not a crow.

Note how differently these people classify the world than we do. They exploit the world to their advantage and therefore focus on what differentiates things: The most important thing for them is how different fish and crows are and they are reluctant to lump them together. We have become used to the categories modern science gives us to understand things. We have developed what I call new “habits of mind.” We are ready to ignore differences and consequently lump fish and crows together as animals, dogs and ourselves as mammals, monkeys and ourselves as primates. When asked what dogs and rabbits have in common, we say they are both mammals. They tend to say that you use dogs to hunt rabbits. None of our modern abstract concepts can actually be perceived in the concrete world. We have a whole new pair of spectacles that they lack. The Wechsler IQ subtest called Similarities is all about classification.

Second interview: Camels and German cities (Luria, 1976, p. 122)

Q: There are no camels in Germany; the city of B is in Germany; are there camels there or not?

A: I don’t know, I have never seen German villages. If B is a large city, there should be camels there.

Q: But what if there aren’t any in all of Germany?

A: If B is a village, there is probably no room for camels.

Note how reluctant people once were to use logic to reason about hypothetical situations (what *if* Germany was a country without camels?). Even when the logical conclusion is suggested, the man tries to turn it into something that describes a concrete situation (perhaps the village is too small for camels). In everyday life, whether you have camels is the important question; not using logic on words or symbols that posit a possible situation that has no reference to anything you have encountered in the real world. We have the “habits of mind” that takes the hypothetical seriously no matter how far it is from anything we “know” and we have had plenty of practice in using logic to determine its consequences. Every bit of science or social science you learn at school puts forward hypotheses to be tested in terms of logic and novel experiences. The Raven's IQ test is entirely about using logic to order symbols that are "valued" for their own sake and have no concrete reference. Formal schooling prepares our minds to do IQ tests successfully, while, in 1900, people found their contents and the problems they posed alien.

Beginning in 1990, scholars began to analyze the "rules" one needs to master to perform well on Raven's. Carpenter, Just, and Shell (1990) showed that you could solve Raven's items by applying certain rules of inference. Verguts and colleagues (1999) showed that facility in using these rules was improved by practice. Since Raven's gains occurred long before anyone was practicing for the test, the implication is that modernity altered people's minds in a way *equivalent* to practicing for the test. This culminated in the seminal article by Fox and Mitchum (2012). Using the Advanced Progressive Matrices test, they analyze the different minds people develop when one generation scores higher on Raven's than the last. The following is in my language (reproduced from Flynn, 2012a, pp. 284–286). However, when we met at the University of Richmond, they confirmed that my interpretation is compatible with their analysis.

Some 115 years ago, people just beginning to enjoy modernity were still focused on the concrete objects of the real world. They wanted to manipulate the real world to their advantage and therefore the representational images of objects was primary. If you are hunting, you do not want to shoot a cow rather than a deer; if a bird is camouflaged in a bush, you flush it out so its shape can be clearly seen. Raven's poses a problem that is quite alien to your "habits of mind": You must divine relations that emerge only if you "take liberties" with the images presented. It is really a matter of perceiving analogies hidden behind distracters. I will present a series of analogies (the first three are my own) to illustrate the point.

(1) Dogs are to domestic cats as wolves are to (wild cats). Presented with these representational images people a century ago would have no difficulty.
(2) ▪ is to ◆ as ↑ is to (➔) where the choices are ↑, ➔, ↖, and ↗. Here you must ignore everything about an image except its shape and position. Just as the square has been rotated a half turn, so has the arrow.
(3) ▪ is to ⁄ as ◯ is to (I) where the choices are ⊘, Θ, I, and ⊗. Here you must ignore everything but the number of dimensions: the analogy compares two-dimensional shapes to one-dimensional shapes and all else is irrelevant. Representational images are three-dimensional, so such a contrast requires being well removed from them.
(4) &#B is to B&# as T&T is to ##(left blank) – you must enter a third symbol that fits. This is an item from Fox and Mitchum that illustrates the kind of analogical thinking you must do on the Advanced Raven's Progressive Matrices.

The fact that the right answer to the fourth item has been left blank means that there are no alternatives presented to choose from. You have to deduce that "&" (or ##&) is the correct answer. I got it right, which was reassuring given that I was then seventy-eight years old, by reasoning as follows. In the first half of the analogy, all that has altered is the sequence of symbols: labeling them 1, 2, 3, they have become 3, 1, 2. Applying that to the second half of the analogy, T&T changes to TT&.

Clearly, you are supposed to ignore the fact that the doubled letter (TT) has changed to another doubled symbol (##), so the right answer is ##&. This would really discriminate between the generations. We have moved far away from the "habit of mind" of taking pictorial images at face value; indeed, we are interested only in their sequence and treat images as interchangeable if the logic of the sequence demands it.

The key is this: Anyone fixated on the literal appearance of the image "T," as a utilitarian mind would tend to be, would simply see no logical pattern. Contrast this with Wechsler Vocabulary. The etiology of enhanced scores over time would be quite different. People over time, thanks to the bonus of more education, simply accumulated a larger store of core vocabulary and got no bonus from the shift from utilitarian toward "scientific" thinking – excepting of course for words that labeled abstractions (like species), which now appeared in the new subjects taught.

Fox and Mitchum classify Raven's items in ascending order of "relational abstraction." To quote: "For analogical mapping when relations between objects are unrelated to objects themselves." Once again, in example #4, the relationship can be derived only if one sees that a "T" does not have to retain its identity as a "T." Their core assumption was that "analogical mapping of *dissimilar* objects is more difficult than mapping *similar* objects" (italics mine). I certainly found this to be true. The fact "TT&" had to be translated into "##&" rendered the item much harder to solve. And, if I were my father (born in 1885) and wedded to taking images at face value, I suspect I would have found it insuperable.

Fox and Mitchum analyzed the performance of two samples of young adults tested in 1961 and circa 2006 respectively. They found that, as the degree of deviation toward the abstract increased, certain items became less predictive of performance within the two generations than they were between the two generations. We now know why Raven's scores are so sensitive to environmental change over time. Like our ancestors, we can still use logic to analyze the concrete world. But we have entered a whole new world that allows us to use logic on symbols far removed from the concrete world. It is a world of the hypothetical (if such and such an analogy holds, what are the logical consequences) well removed from the concrete world.

Premodern people see fish as having nothing in common with crows. You can eat one and not the other; one swims, the other flies. We use DNA analysis to divide living creatures into categories that are nonobservable but offer understanding and this language has become that of every person who has been exposed to several years of formal schooling. We know that bacteria differ from one-celled animals, that whales are more akin to land animals than fish, and that the tiny hyrax is more akin to the huge elephant than to the rodents it resembles. We know that stars are different from planets (they look the same in the sky) and, indeed, our whole picture of the universe (and even our approach to explaining human behavior) is based on logic and abstractions. We are exposed to the symbolism of algebra. No one has ever observed an "x."

In other words, using logic on symbols detached from concrete reality has become a habit of mind in no way alien to us. These skills are not merely useful in mathematics and science and computer programming (programmers do very well on Raven's). They help us to create (and comprehend) a nonrepresentational map of the London underground or a map that functionally relates to one another of the tasks a complex business organization performs. We are more ready to engage with Raven's because the rise of modernity has altered our perspective. And the rise of modernity has occurred over only a few generations. Only a test that is sensitive to the new minds that modernity has put into our heads could measure something so malleable. Raven's, more than any other test, *is a barometer of the stages of modernity* and thus continues to play a crucial role in the study of intelligence. It can tell us how far people have gone down the road that enhances our ability to solve the cognitively complex problems of the modern world.

This focus on the altered minds that engendered Raven's gains omits certain twentieth-century developments that caused gains on the whole range of Wechsler subtests. As for US gains from 1950 to 2004 (Flynn, 2009b, 2012a, pp. 21–23):

Vocabulary: 17.80 points for adults (WAIS), 4.40 for children (WISC). The huge gulf for adults is partially because the twentieth century has seen an explosion in university attendance. But children should profit more from the adult gains. The fact that they do not must say something about growing cultural (and therefore linguistic) segregation between adults and children. The adult gains imply that serious writers today have a larger target audience capable of reading their works, although the visual culture of our time drastically limits the number of those willing to do so.

Information: Adults have gained eight points and children only two points. Again, the expansion of tertiary education has created, for the first time, adults most of whom have finished high school and half of whom have some tertiary experience. This has broadened their fund of general information at least superficially.

Comprehension: Large gains by both adults (almost fourteen points) and children (eleven points). This subtest measures the ability to comprehend how the concrete world is organized (why streets are numbered in sequence). The greater complexity of life today poses a challenge that people have met successfully throughout their life span.

Arithmetic: The small gains of 3.50 and 2.30 points respectively reveal the failure of education on any level, from primary to tertiary, to habituate people to the world of numbers. The tertiary failure is unexpected and shocking.

Picture Completion: Eleven to twelve points all ages. We have all entered a world with a far more rich visual culture but this explanation is unsatisfactory until we can be more precise about proximate causes. Just what cognitive shift allows us to better perceive what is missing in a picture (better mapping skills)?

Coding: Very large gains (sixteen to eighteen points) at all ages. This is an information-processing test that utilizes working memory. The modern world has demanded (and gotten) people who can assimilate information at a faster and faster rate.

Block Design: Like Raven's, this subtest signals enhanced ability to solve "hypothetical" problems. The schoolchild gains (almost sixteen points) are significantly greater than the adult gains (ten points). This makes it tempting to assume that the modern school has increased its demands on analytic ability even more than the modern world of work. However, actual Raven's data show just the reverse pattern, namely larger adult than child gains (Flynn, 2012a, p. 57). Although both are tests of fluid intelligence (solving novel problems on the spot), clearly the two make significantly different proximate demands.

Similarities: The huge gains throughout life (twenty points for children, twenty-four for adults) underline Luria's results. They mark the transition from people regarding the world as something to be manipulated for use to classifying the world using concepts removed from the concrete world. Genovese (2002) shows this for schools, Schooler (1998) for the world of work, Greenfield (1998) for the world of leisure.

Specific Causes

Some emphasize a particular cause of IQ gains, not necessarily because they think it dominant but to assess its peculiar potency. Tuddenham (1948) noted the increase in the amount and quality of formal education in America in the first half of the twentieth century. Teasdale and Owen (2008) used enhanced education to explain Danish IQ gains. Ceci (1991) did an overview of the impact of schooling on IQ and Husén and Tuijnman (1991) broadened this to cover its effects on human capital in general. Blair and colleagues (2005) focused on a change in curricula, namely the trend beginning in America in 1990 to use Raven's-type problems as classroom material to promote the ability to reason mathematically.

Zajonc and Mullally (1997) turned to the family: There was conflicting evidence about whether the child's IQ declined as their place in birth order rose (smaller families might mean fewer "disadvantaged" children). Johnson (2005) criticized those who contended that television has acted as a "dumbing-down" influence over time. All of the factors named are valid in the sense that they have a place in the full list of the intermediate factors that were spin-offs from the industrial revolution and economic growth.

The Life History of IQ Gains

British IQ gains began at least as far back as those born in 1872 (Raven, Raven, & Court, 1993). That gains actually began with the onset of the industrial revolution is reinforced by the fact that developing nations show rapid gains either after they began to industrialize (China) or after they began to experience the effects of industrialization when imported from without (education in rural Kenya). Some very advanced nations that have been highly industrialized for more than a century may be showing signs of IQ decline. All of this suggests a life history.

An Expected Decline

The intermediate factors that have caused IQ gains are such that they should be running out of steam, at least in advanced nations. The relevant spin-offs of the industrial revolution and growing prosperity were a more favorable ratio of adults to children in the home, more and better schooling, more cognitively demanding jobs, more cognitively demanding leisure, and vastly better health and conditions of the aged. These same factors can quickly turn from positive to mixed or even negative. The ratio of adults to children in the home has reached a minimum (indeed today there are more solo-parent homes). Middle-class parents may have used up the tricks that make the preschool environment more cognitively enriching. We appear to have reached a limit in terms of enhanced schooling and the number we keep in school into adulthood. The economy is producing fewer cognitively demanding jobs in favor of more service work. Leisure activity may actually be in competition with quality schooling. However, we may be continuing to improve the health of the aged. Social trends differ, of course, from nation to nation.

Table 39.1 presents recent gains and losses for eleven nations. The losses in Scandinavia may look small at about 0.23 points per year but, if they persisted for acentury, they would amount to twenty-three IQ points and undo much of what was gained in the twentieth century (perhaps thirty points). If similar post-1992/1993 data were available from Sweden, it would probably join them, or so Martin Ingvar argues (Kvarnöf, 2018).

The Scandinavian losses have attracted attention, perhaps because the nations are so socially advanced and because of the excellence of the military samples. However, I wish to call attention to the fact that they are for eighteen- to twenty-year-old males only. For the first time, we now have evidence that IQ gains vary by age. Scottish data show that when two cohorts (separated by fifteen years) were compared at age eleven, the latter had gained 3.7 IQ points. When the same cohorts were compared at age seventy-seven, the latter had gained 16.5 IQ points (Staff et al., 2014).

Let us reflect on the reasons for this. When you compare Scottish eleven-year-olds born in 1936 with eleven-year-olds born in 1921, both cohorts have the same number of years of school, which is a leveler because a slightly enhanced *quality* of school is the only thing that counts. Flynn (2016a, pp. 86, 91) uses international data to show that late teens and young adults tend to gain at about twice the rate of schoolchildren because they have many *more* years of school, thanks to the education revolution of the twentieth century. As for the huge gains of the aged (here seventy-seven-year-olds) throughout the twentieth century, the aged above all have enjoyed enhanced quality of life (better health and exercise).

Table 39.2 gives recent trends in the Netherlands. It is often described as a nation that began to suffer IQ losses beginning as early as 1975. However, an analysis of various age trends from about 1968 to 2005 shows a mix: families are furnishing much the same cognitive environment for preschoolers (static IQs), fourteen- to sixteen-year-old school students show significant losses, and adults are still making gains in the world of work (though the data are not robust after 1982). I suspect that the Dutch data offer a better estimate of recent Northern European trends than the eighteen- to twenty-year-old males in Scandinavia. The latter may reflect that teenagers are having increasing difficulty adapting to formal schooling; and the sample are males only. International data show that females are outperforming males at that age (Flynn, 2011).

Table 39.1 *Recent IQ trends in eleven nations (adapted from Flynn & Shayer, 2017: tables 1 and 2).*

GAINS							
Nation	Ages	Test(s)	Years compared	IQ pts. gained	Rate per year	Sample	Reference
United States	1–16	WISC-III & WISC-IV	1989–2002	4.63	0.363	244 took both tests	Flynn, 2012a, pp. 238–239
United States	1–16	WISC-IV & WISC – IV	2002.25–2013.83	3.62	0.313	126 matched subjects	Weiss, Gregoire, & Zhu, 2016
United States	17–89	WAIS-III & WAIS-IV	1995–2007*	3.37	0.281	240 took both tests	Flynn, 2012a, table AII1
United Kingdom	5–15	Raven's	1980–2008	6.23	0.221	Analysis of CPM and SPM	Flynn, 2012a, p. 322
South Korea	5–16	Wechsler tests & KABC	1986–1999 (av. gap = 8.75 years)	6.71	0.767	Small groups took two tests	Te Nijenhuis et al., 2012
German speakers	Median age 42	Vocabulary	1997–2007	3.50	0.350	Meta-analysis 500+ studies	Pietschnig et al., 2010
Netherlands	Several ages	Various	1971–2005	See Table 39.2	See Table 39.2	Varies by age	See Table 39.2
LOSSES							
Nation	Ages	Test(s)	Years compared	IQ pts. lost	Rate per year	Sample	Reference
Norway	18–20	Similarities, Arithmetic, Figures	1993–2002	− 1.98**	− 0.22	Virtually all males	Sundet, Barlaug, & Torjussen, 2004
Finland	18–19	Word, Number, Shapes	1997–2009	− 2.99**	− 0.25	Virtually all males	Dutton & Lynn, 2013

Denmark	18–19	Analogies, Number series,	1998–2003/4	− 1.19**	− 0.215	Virtually all males	Teasdale & Owen, 2008
German speakers	13–42	Spatial	1995–2014	− 9.12	− 0.480	Meta-analysis 96 samples	Pietschnig & Gittler, 2015
France	30–63	WAIS-III & WAIS-IV	1999–2008.5	− 3.80	− 0.400	79 took both tests	Dutton & Lynn, 2015
Australia	6–11.92	Raven's Coloured Matrices	1975–2003	− 1.32	− 0.047	693 & 618 from Victoria	Cotton et al., 2005
Estonia	15–16	Raven's	2001–2012	− 8.68? See text	− 0.723? See text	552, 411, 304 from 3 schools	Korgesaar, 2013

* United States: The reference puts the WAIS-IV norming date at 2006; 2007 is correct.

** Scandinavia: I have calculated using the means and SDs from the original data and, in some cases, the values in terms of IQ points differ slightly from those reported.

Table 39.2 *Recent IQ trends in the Netherlands.*

Ages	Test	Years compared	IQ points gained or lost	Rate per decade	Sample and No.	Reference
ADULTS						
18	Raven's	1972–1981.5	+ 8.67	+ 9.13	All males	Flynn, 1987, table 1
Mean age 40.30	WAIS & WAIS-III	1967.5–1998.5	+ 13.50	+ 4.35	77 took both tests	Wicherts et al., 2014
Bus drivers	GATB (8)	1975.5–1984	+ 0.74	+ 0.87	Applicants: 130 & 1091	Woodley & Meisenberg, 2013
Bus drivers	GATB (8)	1984–1990	+ 3.69	+ 6.15	Applicants: 1091 & 212	Woodley & Meisenberg, 2013
Assessment adults: 13.5 points on WAIS 1967.5–1998.5; other data indicate similar rates of gain over this generation with bus drivers (job with low cognitive demands) variable						
HIGH SCHOOL						
16	GATB	1975–1985	− 3.60	− 3.60	130 & 270	Woodley & Meisenberg, 2013
	GATB	1985–2005	− 1.49	− 0.74	270 & 498	Woodley & Meisenberg, 2013
14	DAT	1984–1994.5	− 2.82 (+)	− 2.69 (+)	857 scored vs. DAT norms	Woodley & Meisenberg, 2013
Assessment ages 14–16: significant loss – about 5 points 1975 to 2005						
PRESCHOOL						
4.92–5.92	RAKIT	1981.5–1992.5	+ 0.81 (−)	+ 0.74 (−)	208 scored vs. RAKIT norms	Woodley & Meisenberg, 2013
Assessment ages 4–6: probably no change						

(+) Next to a loss means the loss should probably be diminished because of sample quality. (−) Next to a gain means it was probably a bit lower than the estimate. See Flynn & Shayer (2017, particularly box 1) for method of computation and description of samples.

On the other hand, I believe that the gains for adults (as in the Netherlands) are under threat. What about the world of work as an important source of cognitive development? That is, what have automation and other trends done to the ratio between cognitively demanding jobs and less demanding service work? As usual economists disagree, but Richard Florida (2014) has provided fascinating US data. From 1990 to 2012 the "working class" share of jobs has steadily declined from 31.4 percent to 20.5 percent; the "creative class" share has been pretty stable, rising from 29.3 to 32.0; the "service class" has risen sharply from 39.3 to 48.5. Note that the ratio between low skill and creative has risen in favor of the former: 1.35 to one in 1990; 1.57 to one in 2012. The message is that any attempt to trace IQ trends without data for various ages and the factors that impact at various ages is incomplete.

Anomalies

Looking back to Table 39.1, it shows anomalies that I cannot explain. The United States is peculiar: IQ gains at the same old rate (about 0.3 points per year) are present in the latest normative samples of children (the WISC of 2014) and adults (the WAIS of 2002). Perhaps it is a slightly less "developed" society than the Northern European nations and is still making up ground in hothousing preschoolers and school quality. On the other hand, if Florida is correct, the next norming of the WAIS (due soon) should show no gain among adults (except for the aged), thanks to the cumulative impact of fewer cognitively demanding jobs.

German speakers are a puzzle: robust vocabulary gains from 1997 to 2007 and equally robust losses on a spatial test from 1995 to 2014. Some other data leave us unsure of trends. The UK gains are over too long a period to measure recent trends and there are disquieting data from Piagetian tests. The losses in France and Australia (minimal) need more evidence and the "losses" in Estonia are suspect because trends by gender differ so much as to cast doubt on the samples (Flynn & Shayer, 2017).

The huge gains in South Korea, 0.767 points per year or more than double the US rate, are not an anomaly. Nations that began to industrialize later began to experience IQ gains later. But they have an advantage: They found industrialization "waiting for them" as compared to nations that had to "grope their way" toward industrialization. Thanks to this, they have faster rates of industrialization and make IQ gains at an accelerated rate. Japan (Lynn, 1982) and China (Raven, Raven, & Court, 1989, p. RS4.8) have made up any IQ deficit they may have suffered early in the twentieth century. South Korea is on their heels. It will be interesting to see whether these nations find that their gains stop now that they have reached parity with the West.

Developing Nations

There is evidence of significant IQ gains in Kenya, Dominica, Saudi Arabia, Brazil, Turkey, and perhaps North Sudan (Flynn, 2012a). Whether they will continue to gain is dependent on their peculiar history over the coming century. In every case, I take the estimates for their mean IQs (compared to Britain at 100) from Lynn and Vanhanan (2006), as roughly accurate at that date.

Daley and colleagues (2003) compared two large studies of seven-year-olds in Embu, Kenya from 1984 and 1998 on the Coloured Progressive Matrices. The Embu District may (emphasis on the "may") be relatively typical of Kenya, although it is a bit more rural: 84 percent as compared to 74 percent for the nation as a whole. The gains were 13.85 IQ points over the fourteen years. Kenya supposedly had at a mean IQ of seventy-two in 2006. If British gains ended in 2006, and Kenyans continued to gain at that rate, they would catch up by 2034 (or seven-year-olds would with other ages unknown). The authors attribute the gains to improved literacy and health. Kenya is fortunate in having an urban workforce with excellent computer skills and profits from an expanding telecommunications sector, so its young children are a good bet. Why adults would probably still lag will be discussed.

Other nations have a worse prognosis. Dominica, an Afro-Caribbean nation of 76,000 people, is a case in point. Meisenberg and colleagues (2005) used a cross-sectional design to compare two adult age cohorts, one born in 1948 and the other in 1983, on Raven's. The gain over the thirty-five years was eighteen IQ points, so, if their 2006 IQ was eighty-two, they should match the United Kingdom by 2041 (in each case, I assume that UK gains are over). This would be a very bad bet. Dominica's economic future and its ability to continue to improve schooling are at the mercy of nature. Hurricane David in 1979 and Hurricane Maria in 2017 had catastrophic effects and this nation has never gone more than ten years without severe storms. What with climate change, its future is dire. Aside from storms, it suffers from volcanic activity, earthquakes, drought, floods, bush fires, and tsunamis.

I think I have said enough to question any study that predicts when the "developing world" will catch up to the "developed world." You would have to study every nation in turn and have better data than we have, and assume that the future will resemble the past. Three more examples preach the same message.

Batterjee (2011) reports results for children aged eight to fifteen between two standardizations of Raven's done in 1977 and 2010, both samples from urban centers in the Makka Province of Saudi Arabia. The gain over the thirty-three years was 11.70 points, so, if their 2006 IQ was eighty-four, their urban schoolchildren should match the United Kingdom by 2051. Again this would be a bad bet. At present, thanks to oil revenue, the public sector manages to employ its native population in unproductive work (Mahdi, 2011). In 2030, when the oil runs out, its population will have doubled.

Colom et al. (2007) compared children aged from seven to eleven who took the Draw-a-Man test in the city of Belo Horizonte in Brazil: In all, 499 were tested in 1930 and 710 in 2002. The gain over the seventy-two years was seventeen points, so, if their 2006 IQ was eighty-seven, young schoolchildren should match the United Kingdom by 2061 (although the Draw-a-Man test is not an ideal predictor). Despite a recent recession, Brazil has generally enjoyed a highly competitive growth rate and has enormous natural resources. It just might progress until 2061, after which all bets are off. The effects of climate change will begin to affect the world economy in ways that are unpredictable, particularly for a nation in the tropics (Flynn, 2017).

Thanks to Kagitcibasi and Biricik (2011), we have Draw-a-Person (the new name for Draw-a-Man) data for Turkey. I had access to unpublished data and my conclusions go beyond those stated by the authors. In 1977, the test was administered to 218 fifth graders in five schools located in the city of Bursa and nearby rural villages; in 2010, to 258 fifth graders in six schools in the same area. This area, like all of Turkey, was in flux in terms of urbanization and growth of the middle class and there is no way of telling whether the local trends were similar to national trends. My guestimate is a gain of 17.32 points over thirty-three years. The test and limited age of the sample and the estimated gain are really too soft for a prediction but for, what it is worth: given a gain at that rate, and given an IQ of ninety in 2006, Turkish schoolchildren should catch the United Kingdom by 2025. Actually, despite all of my qualifications, they might do it, what with Turkey's annual growth rate of almost 6 percent.

Dutton and colleagues (2018) criticize an earlier study by Dutton, Van der Linden, and Lynn (2016) about North Sudan. Both used data from Raven's Coloured Matrices, which was administered in Khartoum (the capital). In the first study, a sample taken when school was voluntary (and pupils largely middle-class) was compared with one taken when schooling had become compulsory (and pupils included many of the poor). The second study used samples both taken after schooling became compulsory. It showed a gain of nine points varying with age (ages six to nine) between 2004 and 2016. Girls gained much more than boys. This makes sense in that, before 2000, their opportunities were far more restricted. Assuming Sudan had a mean IQ of seventy-one in 2006, their young children would match the United Kingdom by 2045.

This prediction is almost surely mistaken. A sample from Khartoum omits the two-thirds of the population who are rural. Its primary school enrolment is 85 percent compared to 62 percent nationwide. Ever since independence in 1956, Sudan has been at war. The first civil war from 1955 to 1972 killed 500,000 people. The second from 1983 to 2005 killed 2 million and 4 million were displaced. There has been an oil boom since 2006 but those employed in agriculture (mainly subsistence farming) have not profited. Since South Sudan became independent in 2011, the carnage in North Sudan (Khartoum is in North Sudan) has been far less, although North Sudan has fought rebel forces near its border and the two Sudans have engaged in periodic clashes.

Fortunately, we have adult data for North Sudan. It is much easier for a developing nation to introduce formal schooling for children than to create cognitively demanding work for adults. Note how much of our data is not only school data but also Raven's data, a test highly sensitive to schooling. Such gains are likely to be much higher than adult gains on a test like on the WAIS, which measures the kind of gains needed to staff a modern economy. Khaleefa, Sulman, and Lynn (2009) report gains on the WAIS-R from two samples, each having a median age of about fifty. Between 1987 and 2007, the gain was 4.05 points for Full Scale IQ. With a mean IQ of seventy-one in 2006, it would be 143 years or 2149 when adult Sudanese matched English adults. This is silly, of course. It really means that the era of massive IQ gains has not really taken off in North Sudan. These are standardization samples and the authors consider them to be representative. Whether this is possible is another question: Even today UNESCO classifies only 53.5 percent of its people as literate.

The pattern of gains shows a very uneven exposure to the modern world. They appear to be in contact with it to some degree, probably through radio, television, the Internet, and some contact with foreigners. They make huge gains on object assembly and a large one on digit symbol (or coding), reflecting modernity's emphasis on spatial skills and speed of information processing. They make moderate gains on picture arrangement and picture completion, reflecting today's visual culture. However, they show a loss on Similarities (rare in the literature), which should rise when schooling helps us put on "scientific spectacles." Perhaps, this is because the "Muslim curriculum" dominates schooling (with its emphasis on the permanence of religious values). The "school basics" subtests like information, arithmetic, and vocabulary show very low gains at only 1.4 points over the twenty years.

The Human Condition

Today, all theories of intelligence must accommodate the causal interaction between the human mind and human society, which has two implications: the fact that "intelligence" breaks down into a multiplicity of cognitive skills that can rise or fall largely independent of one another; and what I shall term the great *elasticity of the human mind*. Before massive IQ gains were documented, scholars tended to assess the intelligence (later called *g* or the general intelligence factor) of the mass of people *at any particular time* and draw pessimistic conclusions. Few pre–World War I intellectuals (Bertrand Russell and the Marxists were exceptions) thought the masses were collectively capable of playing the cognitively demanding social roles that were the prerogative of the upper classes.

Many championed mass education but with a ferocious pessimism. Virginia Woolf refers to the self-taught workingman as someone "we all knew" to be egoistic, insistent, raw, striking, and ultimately nauseating. E. M. Forster has no sympathy with a clerk whose attempts to educate himself are "hopeless." He is simply inferior to most rich people (less intelligent) and typical of urbanized rural laborers. The latter should be stripped of their education and revert to do what they can do well – breed yeomen. D. H. Lawrence, Pound, Yates, H. G. Wells, George Bernard Shaw, T. S. Eliot, Aldous Huxley, Evelyn Waugh, and Graham Greene derided the capacities of the masses (Carey, 1992). Psychologists were appalled that American males tested in 1917 had the mental capacity of a twelve-year-old (note the absurdity of this), which boded ill for the future. Well, they were all wrong: As society made greater cognitive demands on people, at least 30 percent found they could play the professional work roles required. That is really an underestimate of cognitive progress. Compare the farmer of today who really runs a small business enterprise, studying market trends, keeping his books, budgeting and borrowing, using a galaxy of machines and materials, with the "yeoman" of the past.

A similar state of mind persists among those who view falling IQ scores as "the" major problem. During the twentieth century, when society escalated its cognitive demands, average IQ rose. During the twenty-first century, if society reduces its demands, average IQ will fall. Within limits (we cannot all be an Einstein), society will get whatever cognitive skills are needed to perform whatever roles it assigns. Society is lowering IQ because of all sorts of trends that have nothing to do with the limitations of our minds. If economic trends favor mass unemployment and undemanding service work at the expense of cognitively demanding work, if universities educate less and become the playgrounds of hedonistic youth, if teenage boys find formal schooling intolerable, if people read less and kill Martians on the Internet more, if climate change destroys the social fabric, down the scores must go.

During the twentieth century, simply by growing, the market did wonderful things. It alleviated poverty, promoted formal schooling, upgraded jobs, and liberated women; and it raised IQ without our even noticing. But simply catering to the market's demands cannot solve the problems of the twenty-first century. These require more than continued growth and much more than high IQ. They required what Aristotle called "practical wisdom," a collective effort to humanize our societies with critical intelligence and knowledge at a premium. Ignorance cripples every generation, however high their average IQ (Flynn, 2012b, 2016b).

If you want grounds for pessimism, look beyond the IQ tests to other signs of intellectual deterioration. How different contemporary history would be if there were serious efforts to prevent runaway climate change (Flynn, 2017) or if Americans read enough to know something about the Middle East before they went there to "nation-build" and add to the killing. Sadly, all the trends are in the wrong direction. Between 1982 and 2015, the percentage of Americans with a college degree rose from 18 to 32.5 percent (Statista, 2018). During that period, literary reading (more than one hour per week) among adults (eighteen and above) fell from 57 to 43 percent. Among adult males, it stands at 36 percent (NEA, 2016).

The history of the world in the new century will be determined more by *capitalizing* on people's intelligence rather than by whether IQ rises or falls. I acknowledge the fallibility of the prophet. However, I stand by my prediction that market expansion has had its day. We need a revolution in social philosophy.

References

Batterjee, A. A. (2011). Intelligence and education: The Saudi case. *Mankind Quarterly 52*, 133–187.

Baxendale, S. (2010). The Flynn effect and memory function. *Journal of Clinical and Experimental Neuropsychology*, *32*, 699–703.

Blair, C., Gamson, D., Thorne, S., & Baker, D. (2005). Rising mean IQ: Cognitive demand of mathematics education for young children, population exposure to formal schooling, and the neurobiology of the prefrontal cortex. *Intelligence*, *33*, 93–106.

Brand, C. R. (1990). A "gross" underestimate of a "massive" IQ rise? A rejoinder to Flynn. *The Irish Journal of Psychology*, *11*, 52–56.

Brand, C. R., Freshwater, S., & Dockrell, W. B. (1989). Has there been a “massive” rise in IQ levels in the West? Evidence for Scottish children. *The Irish Journal of Psychology*, *11*, 41–51.

Bratsberg, R., & Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences*, *115*, 6674–6678.

Carey, J. (1992). *The intellectuals and the masses: Pride and prejudice among the literary intelligentsia, 1880–1939*. London: Faber.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of processing in the Raven progressive matrices test. *Psychological Review*, *97*, 404–431.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, *27*, 703–722.

Colom, R., & Flores-Mendoz, C. E. (2007). Intelligence predicts scholastic achievement irrespective of SES factors: Evidence from Brazil. Intelligence, *35*, 243–251.

Cotton, S., Kiely, P., Crewther, D., Thompson, B., Laycock, R., & Crewther, S. (2005). A normative and reliability study of the Raven’s Coloured Matrices for primary school aged children from Victoria, Australia. *Personality and Individual Differences*, *39*, 647–659.

Daley, T. C., Whaley, S. E., Sigman, M. D., Espinosa, M. P., & Neumann, C. (2003). IQ on the rise: The Flynn Effect in rural Kenyan children. *Psychological Science*, *14*, 215–219.

Dutton, E., Bakhiet, S. F. A., Osman, H. A., Becker, D., Essa, Y. A. S., Blahmar, T. A. M., Lynn, R., & Hakami, S. A. (2018). A Flynn effect in Khartoum, the Sudanese capital, 2004–2016. *Intelligence*, *68*, 82–86.

Dutton, E., & Lynn, R. (2013). A negative Flynn effect in Finland, 1997–2009. *Intelligence*, *41*, 817–820.

Dutton, E., & Lynn, R. (2015). A negative Flynn effect in France, 1999 to 2008–9. *Intelligence*, *51*, 67–70.

Dutton, E., Van der Linden, D., & Lynn, R. (2016). The negative Flynn effect: A systematic literature review. *Intelligence*, *59*, 163–169.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29–51.

Flynn, J. R. (1985). Wechsler intelligence tests: Do we really have a criterion of mental retardation? *American Journal of Mental Deficiency*, *90*, 236–244.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*, 171–191.

Flynn, J. R. (1990). Massive IQ gains on the Scottish WISC: Evidence against Brand et al.’s hypothesis. *The Irish Journal of Psychology*, *11*, 41–51.

Flynn, J. R. (1993). Skodak and Skeels: The inflated mother-child IQ gap. *Intelligence*, *17*, 557–561.

Flynn, J. R. (1998). The schools: IQ tests, labels, and the word “intelligence.” In J. S. Carlson, J. Kingma, & W. Tomic (Eds.), *Advances in cognition and educational practice*, Vol. *5*, *Conceptual issues in research on intelligence* (pp. 13–42). London: JAI Press.

Flynn, J. R. (2009a). The WAIS-III and WAIS-IV: Daubert motions favor the certainly false over the approximately true. *Applied Neuropsychology*, *16*, 1–7.

Flynn, J. R. (2009b). *What is intelligence? Beyond the Flynn effect*. Cambridge: Cambridge University Press.

Flynn, J. R. (2012a). *Are we getting smarter: Rising IQ in the twenty-first century*. Cambridge: Cambridge University Press.

Flynn, J. R. (2012b). *How to improve your mind: Twenty keys to unlock the modern world*. London: Wiley-Blackwell.

Flynn, J. R. (2013). The "Flynn effect" and Flynn's paradox. *Intelligence*, *41*, 851–857.

Flynn, J. R. (2016a). *Does your family make you smarter? Nature, nurture, and human autonomy*. Cambridge: Cambridge University Press.

Flynn, J. R. (2016b). *The new torchlight list: In search of the best modern authors*. Wellington, New Zealand: AWA Press.

Flynn, J. R. (2017). *No place to hide: Climate change – a short introduction for New Zealanders*. Nelson, New Zealand: Potton & Burton.

Flynn, J. R., & Rossi-Casé, L. (2011). Modern women match men on Raven's Progressive Matrices. *Personality and Individual Differences*, *50*, 799–803.

Flynn, J. R., & Shayer, M. (2017). IQ decline and Piaget: Does the rot start at the top? *Intelligence 66*, 112–121.

Flynn, J. R., & Widaman, K. (2007). The Flynn effect and the shadow of the past: MR and the indefensible and indispensable role of IQ. *International Review of Research in Mental Retardation*, *35*, 123–151.

Florida, R. (2014). Where the good and bad jobs will be, 10 years from now. Vol. 25, City Lab.

Fox, M. C., & Mitchum, A. L. (2012). A knowledge based theory of rising scores on "culture-free" tests. *Journal of Experimental Psychology: General*, *142*, 979–1000.

Garber, H. L. (1988). *The Milwaukee Project: Preventing mental retardation in children at risk*. Washington, DC: American Association on Mental Retardation.

Genovese, J. E. (2002). Cognitive skills valued by educators: Historic content analysis of testing in Ohio. *Journal of Educational Research*, *96*, 101–114.

Greenfield, P. A. (1998). The cultural evolution of IQ. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 81–124). Washington, DC: American Psychological Association.

Gresham, F. M., & Reschly, D. J. (2011). Standard of practice and Flynn effect testimony in death penalty cases. *Intellectual and Developmental Disabilities*, *49*, 131–140.

Husén, T., & Tuijnman, A. (1991). The contribution of formal schooling to the increase in intellectual capital. *Educational Researcher*, *20*, 17–25.

Jensen, A. R. (1998). *The g Factor: The science of mental ability*. Westport, CT: Praeger

Johnson, S. (2005). *Everything bad is good for you: How today's popular culture is actually making us smarter*. New York: Riverhead Books.

Kagitcibasi, C., & Biricik, D. (2011). Generational gains on the Draw-a-Person IQ scores: A three-decade comparison from Turkey. *Intelligence*, *39*, 351–356.

Kanaya, T., Scullin, M. H., & Ceci, S. J. (2003). The Flynn effect and U.S. policies: The impact of rising IQ scores on American society via mental retardation diagnoses. *American Psychologist*, *58*, 778–790.

Khaleefa, O., Sulman, A. & Lynn, R. (2009). An increase of intelligence in Sudan,1987–2007. *Journal of Biosocial Science*, *41*, 279–283.

Korgesaar, M. (2013). *Flynni Efekti esinemine ee sti abiturientide seas Raveni Testi põhjal [Presence of Flynn Effect among Estonian school-leavers on Raven Matrices]*. Tartu: University of Tartu, Institute of Psychology.

Kvarnöf, G. (2018). Interview with Martin Ingvar. Svenska Dagbladet, January 30.

Laird, C., & Whitaker, S. (2011). The use of IQ and descriptions of people with intellectual disabilities in the scientific literature. *British Journal of Developmental Disabilities*, *57*, 175–183.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature*, *297*, 222–223.

Lynn, R. (1998). In support of nutrition theory. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 25–66). Washington, DC: American Psychological Association.

Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality*. Augusta, GA: Washington Summit Publishers.

Mahdi, W. (2011). Oil is no gold mine for jobs in Saudi. *The National*, March 2.

Martorell, R. (1998). Nutrition and the worldwide rise in IQ scores. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 183–206). Washington, DC: American Psychological Association.

Meisenberg, G., Lawless, E., Lambert. E., & Newton, A. (2005). The Flynn effect in the Caribbean: Generational change of cognitive test performance in Dominica. *Mankind Quarterly*, *46*, 29–69.

Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn Effect and other trends. *Psychological Review*, *114*, 806–829.

Must, O., & Must A. (2013). Changes in test-taking patterns over time. *Intelligence*, *41*, 780–790.

NEA (National Endowment for the Arts). (2016). Results from the Annual Arts Basic Survey, August 2016. National Endowment for the Arts.

Pietschnig, J., & Gittler, G. (2015). A reversal of the Flynn effect for spatial perception in German-speaking countries: Evidence from a cross-temporal IRT-based meta-analysis. *PloS One*, *5*, e14406.

Pietsching, J., Tran, U.S., & Voracek, M. (2013). Item-response theory modeling of IQ gains (the Flynn effect) on crystallized intelligence: Rodgers' hypothesis yes, Brand's hypothesis perhaps. *Intelligence*, *41*, 791–801.

Pietschnig, J., Voracek, M., & Formann, A.K. (2010). Pervasiveness of the IQ rise: A cross-temporal meta-analysis (1977–2014). *Intelligence*, *53*, 145–153.

Raven, J., Raven, J. C., & Court, J. H. (1989). *Manual for Raven's Progressive Matrices and Vocabulary Scales* (Research Supplement No. 4). London: H. K. Lewis.

Raven, J., Raven, J. C., & Court, J. H. (1993). *Manual for Raven's Progressive Matrices and Vocabulary Scales* (Section 1). Oxford: Oxford Psychologists Press.

Rönnlunda, M., & Nilsson, L.-G. (2009). Flynn effects on sub-factors of episodic and semantic memory: Parallel gains over time and the same set of determining factors. *Neuropsychologia*, *47*, 2174–2180.

Rutter, J. M. (2000). Comments in discussion on James R. Flynn. In G. R. Bock, J. A. Goode, & K. Webb (Eds.), *The nature of intelligence* (Novartis Foundation Symposium 233). Chichester: Wiley.

Schalock, R. L. (2012). *User's guide. Intellectual Disability: Definition, Classification, and Systems of Support* (11th ed.). Washington, DC: AAIDD Books.

Schooler, C. (1998). Environmental complexity and the Flynn effect. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 67–79). Washington, DC: American Psychological Association.

Shiu, W., Beaujean, A. A., Must, O., te Nijenhuis, J., & Must, A. (2013). An item-level examination of the Flynn effect on the National Intelligence Test in Estonia. *Intelligence*, *41*, 770–779.

Skodak, M., & Skeels, H. M. (1949). A final follow-up study of 100 adopted children. *Journal of Genetic Psychology*, *75*, 85–125.

Staff, R. T., Hogan, M. J., & Whalley, L. J. (2014). Aging trajectories of fluid intelligence in late life: The influence of age, practice and childhood IQ on Raven's Progressive Matrices. *Intelligence*, *47*, 194–201.

Statista (2018). Percentage of the U.S. population who have completed four years of college or more from 1940 to 2017, by gender. Statisca Portal.

Steen, R. G. (2009). *Human intelligence and medical illness: Assessing the Flynn effect*. New York: Springer.

Storfer, M. D. (1990). *Intelligence and giftedness: The contributions of heredity and early environment*. San Francisco: Jossey-Bass.

Sundet, J. M., Barlaug, D. G., & Torjussen, T. M. (2004). The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence*, *32*, 349–362.

Teasdale, T. W, & Owen, D. R. (2008). Secular declines in cognitive test scores: A reversal of the Flynn Effect. *Intelligence*, *36*, 121–126.

te Nijenhuis, J., Cho, S. H., Murphy, R., & Lee, K H. (2012). The Flynn effect in Korea: Large gains. *Personality and Individual Differences*, *53*, 147–151.

Tuddenham, R. (1948). Soldier intelligence in World Wars I and II. *American Psychologist*, *89*, 54–56.

Verguts, T., De Boeck., & Maris, E. (1999). Generation speed in Raven's Progressive Matrices Test. *Intelligence*, *27*, 329–345.

Weiss, L. G., Gregoire, J.,& Zhu, J. (2016). Flaws in Flynn effect research with the Wechsler Scales. *Journal of Psychoeducational Assessment*, *34*, 411–420.

Wicherts, J. W., Dolan, C. V., Oosterveld, P., van Baal, G. C. V., Boomsma, D. I., & Span, M. M. (2014). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, *32*, 509–537.

Woodley, M. A. (2011). Heterosis doesn't cause the Flynn effect: A critical examination of Mingroni (2007). *Psychological Review*, *118*, 689–693.

Woodley, M. A., & Meisenberg, G. (2013). In the Netherlands the anti-Flynn effect is a Jensen effect. *Personality and Individual Differences*, *54*, 872–876.

Woodley, M. A., te Nijenhuis, J., & Murphy. R. (2013). Were the Victorians cleverer than us? The decline in general intelligence estimated from a meta-analysis of the slowing of simple reaction time. *Intelligence*, *41*, 843–850.

Young, G. W. (2012). Understanding Flynn and honoring Atkins: The Flynn effect should apply in capital determinations of mental retardation or intellectual disability. *Vanderbilt Law Review*, March.

Zajonc, R. B., & Mullally, P. R. (1997). Birth order: Reconciling conflicting effects. *American Psychologist*, *52*, 685–699.

40 Society and Intelligence

Susan M. Barnett, Heiner Rindermann, Wendy M. Williams, and Stephen J. Ceci

There are large between-country differences in measures of economic well-being and noneconomic well-being (democracy, rule of law, human rights, health) – but why? Many researchers from different disciplines view increasing the stock of human capital as the key to raising economic development, promoting democratization, and improving health, and hence improving overall societal well-being. The single most studied aspect of human capital concerns cognitive competence (synonymously cognitive ability) – the capacity to assess and solve problems by the use of thinking (intelligence), to acquire, to possess, and to use knowledge. Some have suggested that differences in population cognitive competence might explain these societal differences (e.g., Hanushek & Woessmann, 2015; Hart, 2007; Jones, 2016). At the individual level, cognitive competence is broadly believed to increase productivity and quality in many realms (employment, child-rearing, health and political decisions, to name a few). Substantial correlations between schooling attainment (i.e., highest completed school grade or level) and these societal and individual outcomes have been interpreted to support the proposition that cognitive competence, the best-known measures of which are psychometric intelligence tests, is influenced by schooling (Ceci, 1991) and in turn drives international differences in health, wealth, and modernity. Understanding the processes by which cognitive dimensions of human capital are fostered represents a key issue of our time. Unsurprisingly, many researchers have toiled on this issue in recent years, focusing on the relationship between transnational gaps in cognitive competence and international differences in wealth, longevity, democratization, and so on (e.g., Rindermann & Ceci, 2009).

For example, there are hundreds of empirical studies that are interpreted as showing the impact of cognitive and other skills obtained through education on wages or incomes; the vast majority of them use schooling attainment to represent these skills (see Psacharopoulos & Patrinos, 2004). A small number instead use direct measures of adult cognitive skills (e.g., Alderman et al., 1996; Hanushek et al., 2015; Murnane, Willet, & Levy, 1995). The many empirical studies of the effects of cognitive and other skills on outcomes such as health, nutrition, and fertility almost all use schooling attainment to represent these skills (see Strauss & Thomas, 1998).

What if genetic differences in intelligence of the populations of each country contributed to international gaps in economic growth and health? This hypothesis was advanced in *IQ and the Wealth of Nations* by the British intelligence researcher

Richard Lynn and the Finnish political scientist Tatu Vanhanen (Lynn & Vanhanen, 2002). In it, these authors discussed the relationship between national IQ and national income for a sample of eighty-one countries, concluding that the results imply that, since largely genetically driven IQ differences are the cause of differences in national income, it will be difficult to impossible to eradicate the gap between rich and poor nations and there is little hope for most poor nations ever to catch up with the rich nations (p. 184).

This view of a role played by a genetic factor is partially supported by a subset of other researchers. Rindermann, Becker, and Coyle (2016) conducted a survey of researchers who had published articles on intelligence and related topics. Genes were rated as the second most important determinant – only surpassed by a summed education factor. Interpretation of these data is difficult, given the low response rate to these questions (5 percent): 1,345 researchers were emailed but only 265 individuals responded and only 71 answered the questions on causes of international intelligence differences (the majority being from English-speaking countries), bringing representativeness into question.

And, of course, opinions cannot substitute for empirical research. Hunt (2012) was quite skeptical about the extant database on genetic causes: “It is conceivable that differences in gene pools also contribute to international and, within nations, group differences in cognitive skills, but at present it is impossible to evaluate the extent of genetic influences” (p. 284). However, recent genome-wide association studies (GWAS) and their application to group-level differences (e.g., Piffer, 2015), along with indirect evidence on evolutionary factors (e.g., genetic proximity; Becker & Rindermann, 2016), give further tentative evidence for an impact of genetic factors that does not exclude environmental causes and their possible interplay.

Using a similarly broad swathe of nations, Rindermann (2018) and Rindermann and Ceci (2009) also reported strong relationships between cognitive competence scores that are highly correlated with IQ, which they derived from a variety of international achievement tests (e.g., Trends in International Mathematics and Science Study [TIMSS], Programme for International Student Assessment [PISA], and Progress in International Reading Literacy Study [PIRLS]), and a host of outcomes that include gross domestic product (GDP), health, rule of law, and measures of modernity. However, these authors concluded that the biggest contributor of transnational gaps was country differences in educational attainment. They suggest that changes in national educational policies can be expected to close these international gaps in GDP, health, rule of law, and so on. This does not rule out an impact of concurrent factors such as wealth, political institutions, and the influence of broad and long-term stable (as national differences, not as levels) background factors such as culture (e.g., religion) or evolution.

The causality problem. However, a correlation between cognitive competence and these measures of societal well-being does not necessarily mean causal influence. Indeed, both could be consequences of some other, third factor or causality could be the other way round – that is, societal differences could cause differences in cognitive competence. For example, rich countries can afford better schools and better schools could lead to higher scores on measures of cognitive competence (whether directly school-related, such as achievement test scores, or indirectly school-related, such as measures of abstract reasoning embodied in IQ tests, e.g., Raven's Matrices), without that higher cognitive competence necessarily leading back to greater national wealth. The direction of causality is important if the goal is to change the level of economic and noneconomic well-being of a country. If cognitive competence causes societal differences, then changing cognitive competence might be one solution to alleviating some of the problems some societies are facing. If, on the other hand, causality is the other way around, and cognitive differences are merely a consequence of societal differences, modifying cognitive competence cannot be the solution. Both can be true: For example, it is possible that in rich countries there is a strong impact of cognitive abilities on growth, while for poor countries there is a strong impact of wealth on cognitive increases (in IQ and student achievement). This can be true also for within-country differences: For the rich in a poor country cognitive ability can be much more important for further professional and economic advancement. If cognitive competence is deemed to be a cause of societal differences, the next question is, Can cognitive competence be changed? If cognitive competence is defined as intelligence, as measured by an IQ test, then the issue becomes, Can intelligence be altered? Some have argued that it cannot, at least not as differences between individuals or countries (e.g., Lynn & Vanhanen, 2002), pointing to the substantial heritability of IQ (individual differences) within societies as evidence. Others have pointed to the malleability of IQ (levels) and other measures of cognitive competence as a result of, for example, schooling, to suggest that providing more/better access to education could change cognitive competence and hence broad societal outcomes (Ceci & Williams, 1997) (See Rindermann & Ceci's [2018] analysis of the greater impact of parental education than wealth on their children's intelligence.) This chapter will discuss each of these issues in turn.

International Differences in Cognitive Competence

Mean scores on tests of cognitive competence differ substantially between countries, whether competence is measured by IQ tests or by tests designed to assess school-related achievement. We will discuss each of these types of measure in turn. Lynn and Vanhanen (2002, 2006, 2012) compiled results from many single studies of intelligence throughout the world, measured against a benchmark, the British "Greenwich" IQ of 100. They found wide variability in measures of national IQ. For example, even within Europe, based on the most recent 2012 compilation, national average IQ estimates range from 89 in Serbia to 101 in Switzerland and Iceland. Outside Europe, they found a much larger range. For example, the Hong Kong and Singapore estimates are at 108, while the estimate for India is 82 and for South Africa it is 72. The lowest IQ estimate in their 136-nation sample is Malawi, at 60.[1] These authors note, in particular, the low scores shown by (black) sub-Saharan African samples, which they calculate to have a median score of 69.

[1] The mean of IQ tests is set at 100 for the United Kingdom, with the standard deviation at 15 ("Greenwich IQ").

However, some authors have questioned the validity (both internal and external) of Lynn and Vanhanen's results, particularly pointing to the unrepresentativeness of some of their samples and the meaningfulness of applying generally US/UK-oriented paper-and-pencil tests to people growing up in very different cultures (Barnett & Williams, 2004, 2005; Hunt & Carlson, 2007). Wicherts and colleagues (Wicherts, Dolan, & van der Maas, 2010a; Wicherts, Dolan, Carlson & van der Maas, 2010a) also reviewed evidence of differences in national IQ. Disagreeing with Lynn and Vanhanen's claim that the IQ of sub-Saharan African nations averaged below seventy, their systematic review suggested a figure of approximately eighty IQ points, the discrepancy between the two due mainly to different choices regarding sample inclusion. Wicherts and colleagues also share some of Barnett and Williams' concerns regarding the meaning of these tests for individuals in undeveloped countries.

Wicherts, Dolan, and van der Maas (2010a) reviewed available IQ data on sub-Saharan African IQ for tests other than Raven's Progressive Matrices and Wicherts, Dolan, Carlson, and van der Maas (2010a) reviewed data from studies using Raven's Matrices and argued that Lynn's (2006) estimate of the IQ of sub-Saharan African nations as less than seventy is too low. They established explicit criteria for inclusion and exclusion of samples in their analysis, though they did not require samples to be representative, as they stated that fully representative samples were not available. Lynn and Meisenberg (2010) critiqued Wicherts and colleagues' analyses, suggesting that many of the included studies were based on unrepresentative elite samples. They proceeded to suggest a different subset of the available studies, based on a case-by-case discussion of sample and test characteristics, which confirmed their original IQ estimate. In response, Wicherts, Dolan, and van der Maas (2010b), statistically analyzed the selectivity of their and Lynn and Meisenberg's reviews. They employed independent raters to judge the available samples on criteria such as whether or not they were drawn randomly and whether or not they were considered representative by the original authors. Their analyses showed that Lynn and Meisenberg's sample selection was not predicted by any of the objective criteria but instead was predicted by the IQ found in the sample. That is, Lynn and Meisenberg systematically included low-IQ results and excluded higher IQ results, yielding a wide difference between the means of included and excluded samples. Further, Lynn and Meisenberg did not consistently apply the criteria they discussed to determine inclusion or exclusion of samples. When Wicherts and colleagues applied the same analysis to their own review, they found a smaller difference between the means of included and excluded samples and stronger prediction by the objective criteria. A logistic regression of probability of sample inclusion on mean sample IQ explained more than 50 percent of the variance for Lynn and Meisenberg's review but less than 10 percent for Wicherts and colleagues' (which they explain by the justified exclusion of unhealthy samples and samples in which the test administration was problematic).

This dispute highlights some of the difficulties involved in objectively assessing cognitive competence across different populations, when nationally representative samples are not available, and hence the difficulty in determining the extent to which differences in cognitive competence might play a role in economic and other societal outcomes. Additionally, differences in measures of cognitive competence between countries, whether larger or smaller, do not tell us whether the scores on such tests can be equivalently extrapolated to draw conclusions about competence outside of testing situations in different cultures. In addition, they do not tell us how such differences came about, nor do they tell us whether correlations between these measures and other measures of societal functioning (income, health, rule of law, etc.) represent causal influence from A to B. For the measurement issue, alternative, more widely used sources as student assessments are necessary (e.g., Sandefur, 2016). For competence outside of testing situations, analyses of everyday-life cognition would be helpful (e.g., Hallpike, 1980). For causal analyses, the results of experimental and nonexperimental studies at different levels have to be considered, weighted, and integrated.

What Do International Differences in IQ and Assessment Test Performance Mean?

To make international comparisons meaningful as indicators of some underlying ability, tests must be measuring the same thing – with equal difficulty – in all countries. But intelligence tests were developed in Western countries and, because of this, they are sometimes suspected to measure only an adaptation to a particular culture ("How well can they do our tricks?"; Wober, 1969, p. 488). Intelligence should be defined as thinking ability across cultures but numerous examples can be cited of cultural variability on cognitive tasks, even for very basic perceptual processes involved in spatial cognition (Henrich, Heine, & Norenzayan, 2010). This issue of cross-cultural validity is not a simple matter, owing to differences in language, culture, and knowledge, and it seems fair to say that no test, no matter how "culture-free" it is claimed to be, is impervious to the effects of culture and schooling. Having stated this, it also seems evident that some tests, including knowledge and school-related tasks, are far more influenced by culture than others.

Tests include items of many different types, including explicit tests of vocabulary and figural problems. For example, the Draw-a-Man test (DAM; Goodenough, 1926; Harris, 1963) is a nonverbal intelligence test in which children are required to draw a man. It was often used in African samples, even though it is not generally considered a good indicator of general intelligence (Wicherts, Dolan, & van der Maas, 2010a). Lynn and Vanhanen (2002, 2012) included some samples using the Draw-a-Man test. Wicherts and colleagues suggest that the use of such samples is fraught with difficulties (e.g., in some cases the children completing the test had never used a pencil, had no schooling, and were unfamiliar with two-dimensional pictures). The tests were also being scored according to culturally loaded criteria, including whether or not the children correctly drew Western clothes on their figures, despite being naked themselves. Other culture-dependent tests include the Kaufman Assessment Battery for Children, which includes items that are likely to be unfamiliar to many test-takers in less-developed countries, at least in the past, such as telephones (Wicherts et al., 2010a). Other well-known tests are also culture-dependent – for example, the WISC-III:

> Questions referring to, for example, "advantages of getting news from a newspaper rather than from a television news program" (Wechsler, 1991, WISC-III Manual Comprehension subtest, p. 138), "why it is important for cars to have license plates" (Wechsler, 1991, WISC-III Manual Comprehension subtest, p. 137), "why you should turn off lights when no one is using them" (Wechsler, 1991, WISC-III, Manual Comprehension subtest, p. 134), "what is an umbrella?" (Wechsler, 1991, WISC-III Manual Vocabulary subtest, p. 108), and "in what way are a telephone and a radio alike?" (Wechsler, 1991, WISC-III Manual Similarities subtest, p. 78), would not be equally difficult, even when translated, for individuals from more and less developed countries. (Barnett & Williams, 2004, p. 390)

Wicherts and colleagues noted that small alterations to the WISC-R, to reduce language and other difficulties, made a large difference in scores of Zimbabwean children, which again raises the question of what these tests are measuring.

Performance of different groups on individual test items can be studied to assess potential bias from Differential Item Functioning (Holland & Wainer, 1993), which can result from groups differing for reasons, such as word familiarity in a math problem, that are not related to the traits being measured. As Wicherts and Dolan (2010) explained, diligent researchers may verify that factor loadings are invariant across groups to confirm that tests are meaningful. However, even if factor loadings do not differ between groups, components of tests can show bias against groups of test-takers. Wicherts and Dolan investigated the fairness of a Dutch IQ test for Turkish and Moroccan immigrant groups compared with majority children, using multigroup confirmatory factor analysis to model the relationship between scores on a subtest and factor scores in different groups. They found differences in the intercept such that the same latent ability results in a lower score for members of the minority groups, concluding: "Ignoring intercept differences may lead to the conclusion that bias of IQ tests with respect to minorities is small, while in reality bias is quite severe" (p. 39); they determined that this resulted in an underestimation of IQ for the immigrant group equivalent to about half a standard deviation. This study was not concerned with international IQ differences. However, the findings regarding cultural bias in test items, even when factor loadings are equivalent, brings into question the role of construct invariance in the interpretation of cross-cultural differences in test scores more generally.

Even tests that appear to be less culturally loaded, such as the Raven's Matrices tests, are considered to have questionable psychometric meaning (Wicherts, Dolan, Carlson, & van der Maas, 2010a) due to test-takers' lack of familiarity with stimulus materials (colored geometric shapes, multiple choice format, etc.). Wicherts and his colleagues stated, "Factor analyses show that the *g* loading of the Raven's tests is considerably smaller in African than in western samples" (p. 145) and "it is unclear whether Raven's tests afford an adequate comparison of western and African samples in terms of the construct of *g*" (p. 145).

Some have gone so far as to claim that "intelligence cannot be fully or even meaningfully understood outside its cultural context" (Sternberg, 2004, p. 325). Sternberg uses the term "successful intelligence" to refer to the practical utility of understanding behaviors within the individual's own particular environment and suggests that, if tests are used cross-culturally, "the psychological meanings to be assigned to the scores will differ from one culture to another" (p. 327). The successful-intelligence approach is based on the idea that "components of intelligence and the mental representations on which they act are universal" (p. 327) but "the mental contents (i.e., types and items of knowledge) to which processes such as these are applied and the judgments as to what are considered 'intelligent' applications of the processes to these contents" (p. 327) vary across cultures. Aspects of a test that are familiar in one situation or culture might be less familiar, and therefore potentially more difficult, in another situation or culture, both for individuals from different cultures in the same test situation and for the same individual in different situations (at home in a village while tracking livestock versus sitting at a desk in a school building surrounded by strangers).

The latter is an example of the context or domain specificity of expertise, knowledge, and understanding. An extensive body of research over the 100 years has shown that learning does not always readily transfer to novel contexts (for an overview, see Barnett & Ceci, 2002). An individual may behave intelligently in a familiar context but not successfully apply that intelligence to an unfamiliar context.

Thus, even if an intelligence test is capable of making meaningful distinctions between individuals who have similar life experiences (whether that distinction is phrased in terms of a latent construct such as "*g*," or in terms of motivational or other causes of differential learning from the same experiences, or in terms of attentional or other constraints on demonstrated performance), it may not have the same meaning when comparing individuals with different life experiences. For example, if individuals in one group have spent several hours a day for several years sitting at a desk in a school listening to a teacher and working with paper and pencil on writing and mathematics and another group has never set foot in such a place and never worked with a paper and pencil, any difference in performance is a confound of what that difference would have been had they had the same experience, and the differences caused by the differential experience.

Similarly, performance may be influenced by cultural priorities such as fast performance versus error-free performance. Roivainen (2010) compared scores on nonverbal, and therefore supposedly culture-free, IQ subtests between representative European and US samples and found a *relative* difference in scores between the Perceptual Organization Index and the Processing Speed Index, which can be calculated from scores on various subtests. This r*elative* difference cannot be explained by *overall* intelligence differences. Roivainen suggests it is caused by cultural differences, such as the priority placed on fast performance in the United States compared with Europe. Therefore, if such differences exist between very similar, Western, industrialized nations, it is likely that significant differences exist between these countries and non-Western, nonindustrialized countries. To the extent that this is the case, it would render the interpretation of cross-national scores questionable. Having stated this caution, recent analyses (Warne & Burningham, 2019) of ninety-seven samples from thirty-one non-Western, nonindustrialized countries revealed a strong general factor in the score matrices of 73 percent of these countries, regardless of the factor-analytic methods used, with equal percentages of variance in the first factor. *g* accounted for about half the variance in non-Western samples, similar to what has been observed in Western industrialized samples. Thus, the suggestion has been made that the *g*-factor of cognitive ability tests, standing for general intelligence, appears to be a universal human trait rather than a cultural artifact, although it is possible that the general factor in a matrix of correlations is something different from general intelligence. The mean level of a trait is sometimes separate from the variance associated with it. It is possible to increase the mean without increasing the shared variance across variables.

So, in light of this, what do international differences in IQ test performance mean?[2] Researchers do not want to unjustifiably disparage the abilities of people from other cultures (Ceci & Williams, 2009). Culture has a strong impact on forms of education, on the esteem a given culture assigns to abstract thinking and knowledge, on diligence and effort (Flynn, 2007), and on thinking styles and worldviews. However, this acknowledgment does not obviate the possibility of making cross-cultural comparisons and researchers from third world nations routinely employ measures that are highly correlated with Western measures, if not the same ones. They seem to believe these measures are relevant in their own cultures. The fact that half of the variance in performance on cognitive batteries is the result of non-*g* sources suggests that behaviors and traits other than those reflected by *g*-based ones are important in understanding cross-cultural differences. Cross-cultural research provides a means of identifying both large background factors and the many small ideological, institutional, and behavioral mechanisms through which the worldviews of cultures work to shape cognitive competencies.

Although some (e.g., Lynn & Vanhanen, 2002) would argue that differences are indicative of underlying general intelligence, the latent construct "*g*," the foregoing suggests they are not error-free measures. The relative magnitude of the signal (*g*) and noise (experientially driven differences) is open to debate. Resolving this debate rests, in part, on the issue of malleability (sensitivity to education and other experiential differences) of IQ, which we discuss in the "Malleability of Ability" section. However, even if they do not measure pure "*g*," IQ tests measure something and, if that "something" can be used to make useful predictions, it may be worth understanding. For example, if national IQ measurements (from appropriately representative samples, etc.) are an indicator of national absorption of formal education and, if the effect of widespread formal education is beneficial for society, then the factors that boost national IQ may be worth investment.

An alternative way to measure the effects of formal education is to do so directly, with tests of academic achievement. Using more knowledge-based student achievement or student assessment tests, which had been applied in a few sub-Saharan countries (where IQ scores are also low, e.g., South Africa, Botswana, and Ghana; Sandefur, 2016), Altinok, Angrist and Patrinos (2018), Lynn and Vanhanen (2012), and Rindermann (2018) have demonstrated averages of around 304, 312, and 319 in student assessment scales ($M = 500$, SD = 100) not corrected for age or youth not

[2] There is considerable debate about the meaning of intelligence and whether IQ tests really measure it (Ceci, 1996). However, we will not discuss this wider debate here, except to address issues particular to the interpretation of international comparisons of IQ.

attending school, in IQ metric representing 71, 73, and 73 IQ points for these countries. Measures of cognitive competence other than IQ show large ranges similar to less knowledge-based figural tests such as mazes (e.g., Coloured Progressive Matrices, Standard Progressive Matrices, and Advanced Progressive Matrices).[3] For example, the 2015 Trends in International Mathematics and Science Study (TIMSS), a series of international assessments carried out in about sixty countries around the world to assess mathematics and science learning in the fourth and eighth grades, found large differences in mathematical performance at both grade levels (Mullis et al., 2016). In the eighth-grade sample, Singapore and South Korea recorded the highest average scaled scores, at 621 and 606, respectively, while Argentina and Saudi Arabia scored the lowest, at 349 and 368, respectively. (The mean is 500, the standard deviation, 100.) Findings were similar in the fourth grade: Singapore and Hong Kong recorded the highest average scaled scores, at 618 and 615, respectively, while Kuwait and South Africa scored the lowest, at 353 and 376, respectively. In summary, the well-known large-scale student assessment studies also demonstrate very large transnational differences in cognitive competence, broadly defined as the capacity to assess and solve problems by the use of thinking (intelligence) to acquire, possess, and use knowledge.

The relationship between these two kinds of measures of cognitive competence – intelligence and achievement – is a contentious topic. Some psychometricians argue that intelligence tests, particularly those assessing fluid intelligence (Cattell, 1987), are tapping an innate ability driven by brain differences related to neuronal processing and working memory capacity and, as such, are measuring something completely different from more knowledge-based performance on school-related assessment tests, although the latter is influenced by such neural processing (for a review, see Neisser et al., 1996). However, content and cognitive task analyses of student assessment items (Rindermann & Baumeister, 2015) and the high correlations between aptitude and achievement test scores in intranational samples (Ceci, 1991), coupled with similar cognitive demands and very high correlations at the between-country level (Rindermann, 2018), lead to the conclusion that the various measures of cognitive competence are largely tapping the same characteristic. Translating international score differences into an easy to understand metric, "years-behind-at-school," suggests that the larger transnational gaps are equivalent to about five to ten years of schooling among children, adolescents, and young adults between ages ten and thirty.

Student assessment tests can validate measures of national cognitive competence levels based on psychometric IQ tests. Of course, using this approach suggests educational causes of cross-country differences in cognitive competence; and using another test-based approach does not answer whether the pattern of cognitive

[3] CPM, SPM, and APM – psychometric paper-and-pencil tests using only abstract figures (similar nonverbal-figural scales of CogAT) – are less overtly related to explicitly, school-taught knowledge than intelligence tests using verbal and math tasks or student assessment tests (using verbal and math tasks and knowledge questions). But performance on these tests and intelligence underlying the performance on them are not independent of school attendance and instructional quality (e.g., Becker et al., 2007).

results generalizes outside the test and school context to everyday life. There are important historical analyses of thinking based on the Piagetian cognitive development approach, for example, of average people by Christopher Hallpike (1980) and Alexander Luria (1976), of certain societal institutions and norms by Georg Oesterdiekhoff (2014), or of intellectual elites by Charles Radding (1985). For instance, using books, manuscripts, and letters written during medieval times from Augustine to Abelard, Radding (1985) showed an important historical cognitive development. Such analyses can be done also for currently living people, average people, political elites, and intellectuals, analyzing behavior and "sediments" of past cognitive behavior (e.g., texts, technology, and institutions). For instance, what convictions do people and elites have regarding AIDS, its causes, and how best to deal with it? Especially revealing are indicators of behavioral and epistemic rationality, for example behavior regarding beliefs about witchcraft and science, poverty and wealth, illness and health, failure and success in life. However, as Rindermann (2018) tried to show, applying such an approach will never supplant results based on large samples and objective methods.

Cognitive Competence and Societal Measures

Many have noted that cognitive competence appears to be related to societal measures of economic and noneconomic well-being. Lynn and Vanhanen (2002) assessed the correlation between national IQ estimates and national per capita income (gross domestic product, or GDP, per capita) and found a correlation of $r = 0.62$, for 1998, with higher IQ countries showing higher per capita income. Whetzel and McDaniel (2006) reached a similar conclusion using updated data. They avoided some of the methodological issues raised concerning Lynn and Vanhanen's study by truncating all IQ scores below ninety to equal ninety; the relationship between IQ and GDP remained strong. Other researchers using student achievement studies or further control variables and different statistical methods found supporting positive relationships (Hanushek & Woessmann, 2008; Jones, 2016; Weede & Kämpf, 2002).[4] According to some, cognitive measures appear to be not just related to overall income levels but also to income distribution. Meisenberg (2012), using IQ data from Lynn and colleagues (Lynn & Vanhanen, 2001, 2006; Lynn, 2010; Lynn & Meisenberg, 2010), observed that "IQ is more potent than education, GDP, and other development indicators in predicting an egalitarian income distribution" (p. 106).

Additionally, there are positive correlations between measures of cognitive abilities and noneconomic aspects of national well-being such as democracy, the rule of law, and political liberty. For example, within the United States, voters have a higher IQ than nonvoters, by more than half a standard deviation (Meisenberg, 2015). Internationally, Glaeser, Ponzetto, and Shleifer (2007) have argued that the causal

[4] Describing the positive impact of one variable on the other does not imply that other variables have no influence. Intelligence is not the only determinant for wealth, for example. There are additional factors aside from intelligence (e.g., mineral resources), behind intelligence (e.g., culture), and between intelligence and positive outcomes (like the quality and functionality of institutions).

path runs from increased education to increased democracy. Positive effects remain when income is controlled (Rindermann, 2018): cognitive ability correlates with democracy (*N* = 187) at *r* = 0.55 (partial correlation with GDP controlled = 0.23); cognitive ability correlates with the rule of law (*N* = 153) at *r* = 0.63 (r_p = 0.18). The level of democracy was measured by two indices: one combining variables such as the fragmentation of the vote between political parties and the level of voter turnout, the second aggregating essential political indicators such as guarantees of civil liberties. The rule of law was measured by indices focusing on protection of property rights and judicial independence. The correlations are not extremely high, thus leaving space for exceptions such as high levels of intelligence and knowledge in Singapore or China and only low or zero levels of democracy. At the individual data level (Cunha et al., 2006; Ellis & Walsh, 2003; Thomson, 1937), cognitive ability is negatively correlated with levels of violent crime. This also holds between regions within a country, such as Japan (Kura, 2013). In addition, Rushton and Templer (2009) report noneconomic national well-being correlates, using Lynn and Vanhanen's national IQ data: "Cross-national differences in rate of violent crime (murder, rape, and serious assault) were significantly correlated with a country's IQ scores (mean r = –.25, such that the higher the IQ, the lower the rate of crime)" (p. 345). The relationship remains robust, excluding sub-Saharan African countries for which IQ estimates may be less valid (*r* = –0.35). Burhan and colleagues (2014) analyzed cross-national IQ (using cognitive ability data from Rindermann, Sailer, & Thompson, 2009) and crime measures, after controlling for other societal variables such as drug and alcohol abuse, education, age distribution, income per capita, and urbanization. They examined the relationship between IQ and various types of crime (homicide, burglary, etc.) at the bottom, middle, and top of the IQ distribution. IQ was found to negatively correlate with homicide across the IQ range and with robbery, though only at the top of the IQ range and using controls – the results were not very robust.

Rushton and Templer (2009) also investigated the relationship between national IQ and health measures, reporting correlations between IQ and the rate of HIV/AIDS (*r* = –0.52), infant mortality (*r* = –0.67), and life expectancy (*r* = 0.74). Thus, measures of cognitive competence and indicators of economic and noneconomic national well-being have been shown to be significantly correlated. Even if these cognitive measures are not assessing potential but merely some form of realized potential in academic-style tasks, their relationship with measures of national well-being merits further investigation.

However, interpretation of societal and political correlates of intelligence is complicated. For example, within the United States, the relationship between party affiliation (Democrat vs. Republican) and intelligence suggests a cohort effect (Ganzach, 2017), such that, for individuals born in the 1920s, the relationship between intelligence and Democratic affiliation is negative, that is, more intelligent individuals are less likely to be Democrats, whereas, for individuals born in the 1960s, the relationship is positive, that is, more intelligent individuals are more likely to be Democrats. This opens up the intriguing possibility of cohort effects for other measures of the impact of intelligence on societal variables.

Direction of Causality

Given a correlation between higher national cognitive competence and positive societal outcomes, the question remains: Does higher cognitive competence (howsoever derived) cause the positive outcomes (i.e., smarter people make better decisions and end up richer and healthier), do the positive "outcomes" cause higher cognitive competence scores (i.e., rich, healthy people have time and energy to devote to learning and so end up smarter), or could the relationship go in both directions? It may be easier to study, learn, and score high on cognitive tests if you are healthy and live in a law-abiding democracy that allows all children to attend, and afford, good schools, and studying and learning may lead to better lifestyle decisions. It is also possible that some of the correlations mentioned above are not causal in either direction but are both the consequence of some other factor, such as culture.

Although random-assignment, experimental studies are impractical, individual, within-country, quasi-experimental data do provide some evidence for a causal link between education and earnings. For example, Angrist and Krueger (1991) investigated the way that compulsory schooling age rules affect the amount of education children receive – depending on whether they are born earlier or later compared to the age cutoff – and the subsequent effect this exerts on earnings. Those students "who are compelled to attend school longer by compulsory schooling laws earn higher wages as a result of their extra schooling" (p. 1010). Unfortunately, investigation of the relationship between education and earnings *between* countries is even more difficult due to the many potential confounded variables.

One way to examine such relationships is to look at the correlation between potentially causal factors at some point in history with potential dependent variables at a later time, controlling for the level of likely confounds. Rindermann (2018) adopted this approach. A longitudinal cross-lagged analysis from 1970 to 2010 in a sample of seventeen (student assessment test data) and forty-seven (years of school education) nations was used to assess the possible direction of causality between cognitive human capital and national income. Longitudinally, the standardized path coefficient for the impact of cognitive human capital on GDP per capita (GDP/c) was 0.31 while the coefficient for the impact of GDP on cognitive abilities was 0.05. Using the logarithm of GDP/c, which stresses increases for lower levels of GDP (wealth increases in poorer countries are more emphasized), the reverse effects were stronger, $ß_{CC1 \rightarrow GDP2} = 0.19$ and $ß_{GDP1 \rightarrow CC2} = 0.25$. While wealth is important for poor countries to improve their cognitive levels, for rich countries, cognitive ability becomes more and more important for wealth production. Wealth effects on cognitive ability peter out. As economies modernize, cognitive abilities grow in importance. Cognitive human capital works also indirectly via its positive impact on economic freedom, which itself supports economic growth. In liberal countries, cognitive ability again becomes more important: economic freedom boosts cognitive effects (Coyle, Rindermann, & Hancock, 2016).

The impact of pure economic factors has also been found to be relatively weak at the individual data level, if the socioeconomic status (SES) variable is divided into two of its components: educational attainment and wealth (Rindermann & Ceci, 2018). Using datasets from Austria, Germany, the United States, Costa Rica, Ecuador (indigenous people), and Vietnam, the educational level of parents was more important for explaining (at least statistically) the cognitive ability level of children than the parental level of financial affluence explained cognitive ability. (Similar findings have been reported by Melhuish et al., 2008.) Rindermann and Ceci (2009) suggested that income at the national level could be more important indirectly, depending on the distribution and use of wealth within a country. Economic resources spent for sufficient and high-quality nutrition (proteins, vitamins, minerals; Eysenck & Schoenthaler, 1997; Lynn, 2009) and health care (from pregnancy on to anti-worm treatment and to vaccinations such as against measles; Glewwe & Kremer, 2006) reaching the whole population (including the poor, orphans, and children of poorly educated parents) provide a basis for a healthy cognitive (and physical) development.

There is some evidence that measures of noneconomic well-being can also be affected by cognitive competence. Within-country evidence shows a statistical relationship between individual differences in childhood cognitive ability and adult health, even after controlling for SES (Gottfredson & Deary, 2004). Although these researchers' methodology was not experimental, the longitudinal nature of their study suggests that cognitive ability differences may be causal. However, in the absence of intervention studies, evaluating causality from between-country cognitive competence differences to between-country health differences is more difficult due to the necessity of more extensive controls for other variables, such as access to health care. Nevertheless, different authors using different data sources (educational or competence measures) have come to the conclusion that human capital is more important than wealth even for health factors such as a reduction in the spread of HIV (Lakhanpal & Ram, 2008; Rindermann, 2018). However, the more "political" attributes of societies become, for example, democracy compared to GDP or human rights compared to achievement in science, technology, engineering, and math, the more important cultural factors seem to be. For instance, for understanding international differences in human rights, culture (as indicated by religion) is more important than cognitive ability, as would be supposed by a Piagetian-Kohlbergian view on moral development (Rindermann & Carl, 2018).

As mentioned in the previous section, correlational analyses also found statistical relationships between measures of cognitive competence and democracy. Within-country longitudinal evidence, which supports a causal interpretation, also exists for a relationship between childhood cognitive ability and adult voter turnout, after controlling for various personality and social variables (Denny & Doyle, 2008). Voting – engagement in the political process – could be viewed as an indicator of democratization in general. The same is true for attitudes of tolerance and liberty (Deary, Batty, & Gale, 2008).

Thus, cognitive competence and education may help improve societal well-being, including wealth, and evidence suggests a link between education and wealth, not purely a consequence of wealth buying education. However, generalizing from quasi-experimental data requires caution.

Malleability of Ability

Even if there is a causal relationship between cognitive competence and desirable societal outcomes, there may be nothing that can be done to promote these desirable outcomes unless cognitive competence is malleable. Some have claimed that cognitive competence, as measured by IQ (individual differences), is heavily influenced by genetics and thus is not very malleable in response to policy interventions (see, e.g., Lynn & Vanhanen's comments regarding the impossibility of eradicating the difference between poor and rich countries, mentioned in the introduction to this chapter). High heritability within a population does not, however, necessarily imply (or preclude) equivalent heritability for differences between populations nor does it imply that means cannot be raised. Given the obvious difficulty of conducting behavioral-genetic twin and adoption studies between populations and countries (take two US identical twins separated at birth, send one to live in a village in sub-Saharan Africa and one to live in Pittsburgh, then take two African identical twins separated at birth and . . .), Rushton and colleagues (2007) attempted to address these questions by comparing the patterns of item difficulty and heritability for IQ test items across populations. They used the Raven's Progressive Matrices test, which is often considered one of the least culture-bound tests, and compared groups from Canada, the United States, Serbia, and South Africa. Within the South African sample, they also compared different ethnic/racial groups. They found that population differences on item scores correlated with item heritability within the Canadian and US twin samples, leading them to suggest that IQ differences *between* populations, as well as individual differences *within* populations, are highly genetically driven and hence nonmalleable. These data are also open to alternative explanations. For example, if heritability was driven by attention differences, with more heritable items being those requiring the most careful concentration, international differences due to lack of experience with schooling and sit-down, paper-and-pencil tests might also correlate with this but for environmental rather than genetic reasons. That is, test-takers in a less developed country, where they did not have so much experience with concentrating for long periods of time on written materials, might do poorly on items requiring such careful concentration, compared to test-takers in a more developed country where they have much more experience with such tasks. Admittedly, this is speculative, and perhaps even far-fetched, but it illustrates the difficulty of making transnational inferences based on within-country heritability estimates obtained in developed nations.

Moreover, there is also considerable evidence that IQ levels, and other measures of cognitive competence, can be changed by education (see, e.g., Ceci, 1991; Hansen, Heckman, & Mullen, 2004; Nisbett, 2009), despite strong genetic effects (Neisser et al., 1996). It has been suggested that schooling and school-related activities foster the development of cognitive competencies that promote performance on most intelligence tests (Cahan & Cohen, 1989). Perfectly controlled experiments are impossible to conduct – children cannot be randomly assigned to be deprived of an education in the name of research – but researchers have provided several sources of evidence to support this claim. Some analyses are correlational, such as analyses of the relationship between IQ and number of years in school. However, many come from natural experiments. Ceci (1991) reviewed studies in which IQ has been shown to decline during summer vacations and among those who have been unable to reliably attend school due to their parents' occupation or the unavailability of schools. For example, children living in remote "hollows" in mountains west of Washington, DC, early in the twentieth century, had reduced exposure to school compared to those in less remote areas, presumably independent from genetic background. IQ scores were found to vary with availability of schooling. Further studies found that delayed onset of schooling depresses IQ scores, whether the delay was owing to war, unavailability of teachers, closure due to racial desegregation, or school entry cutoff dates (Cahan & Cohen, 1989; Ceci, 1991; Stelzl et al., 1995). School-age cutoffs were used by Cahan and Cohen (1989) in their quasi-experimental study of the effect of the amount of schooling on fifth- and sixth-graders' scores on various verbal and nonverbal intelligence tests, including the Cognitive Abilities Test and Raven's Matrices. They concluded, "The results unambiguously point to schooling as the major factor underlying the increase in intelligence test scores as a function of age" (p. 1239). Similar results were found by Stelzl and colleagues (1995). They also used a quasi-experimental design to separate schooling from age effects on intelligence test scores of ten-year-old children. Their results showed considerable schooling effects on all tests, including the tests of fluid intelligence.

Academic activities such as training on reasoning have been shown to enhance so-called culture-reduced tests of fluid intelligence without math and verbal tasks similar to Raven's Matrices. For example, Klauer and Phye (2008) have shown in a meta-analysis of seventy-three studies with seventy-nine comparisons a mean effect of cognitive training on intelligence (mainly measures of fluid intelligence, using Cattell's Culture Fair Test) of $d = 0.52$.

Thus, at least within countries, there is considerable evidence that IQ is malleable and that education can lead to changes in cognitive competence, as assessed by measures such as IQ tests. Between-country evidence also shows a correlation between schooling and IQ.

In assessing the benefits of education, it is important to distinguish between the benefits in terms of increases in cognitive competence and the benefits in terms of gaining credentials the world might interpret as a signal of increased cognitive competence (or other related skills), whether actual or not. The latter has been termed the *signal theory* of educational effects (Caplan, 2017; Spence, 1973). Signal theory argues that educational attainments *only* serve to signal the competence level of individuals, that is, attainments achievements do not reflect actual cognitive competence but rather serve as signals for such personal attributes as stick-with-it-ness and conformity. For example, college education would not enhance cognitive competence but merely signals competence to a prospective employer or graduate school admission committee; persons intelligent enough to get through college and to receive a degree are assumed to possess a minimum level of intelligence and beneficial personality traits (e.g., conscientiousness) but college attendance or school education themselves do not increase abilities (e.g., Caplan, 2017;

Charlton, 2009; Murray, 2008). Signal theory is, of course, controversial and is not compatible with the results of much empirical research: Many quasi-experimental studies have shown that the quantity of education alters cognitive competence (academic achievement and IQ; e.g., Cahan & Cohen, 1989; Stelzl et al., 1995). Thus, whether or not there may also be a signaling effect of educational credentials, signal theory cannot explain all of the benefits of education.

At the transnational level, signal theory is irrelevant – Why should the overall economy develop better if people are absent from the labor market to spend their time on "learning" if it brings no real benefit? It seems unlikely that international investors or importers would invest in or buy from a country purely because of the educational credentials of its population.

Further, Wicherts, Borsboom, and Dolan (2010) note that national IQs are not stable and appear to be quite malleable. They point out that "the socio-cultural achievements of the peoples of Mesopotamia and Egypt in 3000 BC stand in stark contrast to the current low level of national IQ of peoples of Iraq and Egypt" (p. 104). Flynn (2018) has written extensively about the changes in IQ over more recent time frames, finding massive gains (two standard deviations) during the twentieth century in many countries, which he attributes to environmental factors such as "more adults per child in the home, more and better schooling, more people at university, more cognitively demanding jobs, and better health and conditions of the aged" (p. 80). He uses the term "cognitive exercise" to describe the consequences of social change as "more schooling, more cognitively demanding work, and more cognitively demanding leisure" (p. 76). Flynn points out that these gains over time do not correlate with the *g* loadings of subtests, neither do deficits caused by such assaults as prenatal cocaine or alcohol exposure and iodine deficiency. Thus, non-*g* gains appear to be causally potent. You may or may not choose to call these gains intelligence but they are very real nonetheless. If changes over time in one country are similar to differences between countries at the same time, then the same arguments apply. However, it has to be noted that levels may change (e.g., the Flynn effect) while patterns (e.g., differences or relative ranks between individuals and peoples) may change less across time.

Policy Implications

If schooling can change cognitive competence, and cognitive competence affects national economic and noneconomic well-being, then investment in raising the national level of schooling might be a good way to alleviate some of society's ills. Reviewing evidence of the interrelationship between schooling, intelligence, and income, several authors concluded, for different countries (including the United States, the United Kingdom, South Africa, Sweden, and Germany), that schooling increases individual income, both directly and via enhancement of intelligence (Bond & Saunders, 1999; Ceci & Williams, 1997). However, variations in individual IQ only explain a small amount of variance in individual income in the intranational samples.

Psacharopoulos and Patrinos (2004) reviewed studies of the return on investment in education in the tradition of the pioneering work of Angrist and Krueger (1991), based on human capital theory. Return on investment is measured by the increase in per capita income for each additional year of schooling. Their review encompasses studies from many countries, each evaluating intranational returns on investment, focusing only on individual income differences but considering both individual and social costs. (Note that the income benefit may include both increases due to improved competences, cognitive and other, and increases due to signaling effects.) Rates of return vary by geographic region and are higher for less well-developed nations. Returns are also higher for primary education than for secondary or higher education, a finding consistent with Heckman and Masterov (2007). Private returns for primary education in sub-Saharan Africa are shown to be very high (37.6 percent), while social returns (including shared, "social" costs) are also high (25.4 percent).

An investigation by Rindermann and Ceci (2009) of the relationships between aspects of national educational systems and cognitive competencies aimed to determine the optimal educational policy choices to efficiently promote cognitive competence. The most important factor seems to be a general high educational level of society (high adult literacy rate, adults who have attended many years of school, adults who completed secondary or at least primary school). Cognitive competence is defined by Rindermann and Ceci as the mean cognitive competence level of students at school (measured using large-scale international student assessments such as TIMSS, PIRLS, and PISA) and the mean intelligence level in society, adapted from Lynn and Vanhanen (2006). Strong, positive relationships were found between kindergarten attendance and subsequent cognitive competence, even after controlling for other factors such as GDP, suggesting that early education provides a basis for subsequent successful ability development. Similar beneficial results of preschool education were found within different countries (e.g., W. S. Barnett & Boocock, 1998; Cunha et al., 2006). Number of instructional hours was also correlated with competence, leading to the conclusion that the more formal education students receive – and the younger they are when they begin to receive it – the higher their achieved cognitive competence levels are (at the individual data level; see also Ceci, 1991). However, just spending more money seems to be ineffective: Although educational expenditures are highly correlated with cognitive outcomes, the relationship disappears when GDP is partialled out.

Large class sizes were found to have a negative effect on cognitive competence, though this can be alleviated by cram school attendance, where available, and good discipline helps promote success, as does the use of achievement tests and central exit exams. Discipline and behavioral education seem to be especially important for pupils from families with low educational background (Woodworth et al., 2008).[5] More time spent on homework has a negative effect on cognitive performance in

[5] We use the term "low educational background" instead of terms such as "poor" or "minority" because the decisive variable seems to be not the status of being poor or an ethnic minority but the educational background of the parents and their values and abilities (Rindermann & Ceci, 2018). Among both poor and minority parents, the offspring of those with higher educational attainment have superior cognitive outcomes than those of wealthier parents with lower educational attainments.

poor school systems (but only at the cross-country level!). Overall, the results of Rindermann and Ceci's study suggest that increased gross and net learning *time* (from kindergarten and early school enrollment to adults' level of education) is important for the development of cognitive competence. However, as Hanushek and Woessmann (2008) note, quality of education is also important: "Knowledge rather than just time in school is what counts . . . School attainment has a positive impact only if it raises the cognitive skills of students – something that does not happen with sufficient regularity in many developing countries" (p. 658). Discipline of students (e.g., attending school regularly, not arriving late, not disturbing lessons), effective classroom management by teachers, and the use of high-stakes tests also lead to more net learning time.

Some countries with large gender inequities do not educate their females as much as their males. If a large part of half the population is exposed to less education, then the benefits of education for cognitive abilities and hence societal benefits will be reduced correspondingly (Hunt, 2012). Also, if other factors such as malnutrition and exposure to environmental toxins, for example lead in paint and gasoline (Canfield et al., 2003), can affect cognitive abilities, then addressing malnutrition and removing these toxins could have a beneficial effect (Hunt, 2012).

Caveats

Education is not an isolated factor. Several studies have shown strong relationships between educational level and attributes of educational systems, on the one hand, and cognitive competence, on the other. The obvious consequence would be to recommend the extension of education and the improvement of educational systems. But the realization as well as outcomes of such reforms could be faced with several problems:

1. Educational attributes of societies do *not* exist *accidentally*. For instance, the existence of a large private school population in the United States and the absence of this sector in Scandinavia have their roots in cultural, historical, and social features of societies that cannot be neglected.
2. The same attributes of educational systems could have *differential impacts* depending on other educational and cultural features of societies. For example, late school enrollment in Finland is not detrimental because traditionally literacy education (at least the beginning of literacy education) occurs in families. Large class sizes in East Asian countries do not impede achievement because the entire culture emphasizes personal effort and discipline and because regular instruction in school is often accompanied by instruction in cram schools outside of regular education. So, in these countries, reforms leading to earlier school onset or smaller classes would likely have rather small effects.
3. Educational attributes such as kindergarten attendance, discipline, central exams, the use of tests, the age at which students are first segregated into more versus less academic tracks, and instructional techniques *cannot be easily manipulated*. Educational traditions react sluggishly to attempts to change their direction.

Additionally, pressure groups could oppose reforms (e.g., in Mexico) and there could be conflicts of interests between parties, trade unions, parental organizations, and media.

4. Educational reforms have *side effects*. For instance, if, in less developed countries, the educational level is raised, traditional aspects of societies from familial cohesion up to the influence of an old religious elite (e.g., mullahs and sangomas – healers in sub-Saharan Africa) may be weakened. A culture might change when educational levels increase. Single modifications such as earlier tracking could increase within-country differences or, like delaying tracking, increase the "bright flight" to private schools where a more tailored academic experience can be offered for those who can afford it.

Reciprocal causation. Neither at the level of individuals nor at the level of nations is education the single determinant of cognitive ability differences or of developmental processes. Numerous other factors (e.g., culture and genetics) have been empirically verified (for a list, see Rindermann, 2018). And, of course, there are reciprocal effects: Education nurtures ability and ability promotes insight into the benefits of education and more generally into the advantages of a stimulating environment and lifestyle. Intelligence and knowledge enhance the ability to understand causal relationships, to anticipate future events, to act in a rational manner, and to modify environments – from their physical aspects to their social and cultural dimensions. So, intelligent people may start with a higher probability of modifying their physical, social, and cultural world and be able to construct this world in a more beneficial and more complex way; and such an environment will have an impact on ability.

Recommendations for Future Research

Psychological and economic researchers have conducted many statistical studies of the possible benefits of cognitive competences and education and why countries differed in economic and cognitive development. In future research, this research should be complemented by in-depth studies of single countries and their educational policies and the possible effects of other social, economic, and cultural conditions supporting or impeding ability development. Such studies should ideally focus on countries at the top of international competence, such as the culturally very different Finland and Singapore. Possibly their experiences could not only increase our knowledge of the determinants of cognitive enhancement but also assist other countries in their educational reforms.

Conclusions

Research on this topic is difficult due to the fact that experimental methods are not appropriate for many questions. Inferences must be derived from nonexperimental, correlational data – cross-sectional, cross-lagged longitudinal, or quasi-experimental.

Conclusions cannot be based on a single, watertight experiment but must be generated by converging weaker evidence from multiple sources. That being said, for some questions, enough nonexperimental data exist to allow tentative conclusions. Evidence suggests that education does build cognitive competence and education and cognitive competence promote better social outcomes, in terms of both economic and noneconomic factors. Cognitive competence here is used to refer to the ability demonstrated in academic-style, paper-and-pencil tasks of the sorts of skills schools seem to build. These studies do not assess practical abilities, creativity, wisdom, and so on. Such skills are certainly useful and may or may not correlate (positively or negatively) with education, GDP, and other societal outcomes. However, within the limited sphere of the cognitive tests discussed here, cognitive competence appears malleable and education both fruitful and beneficial to society.

References

Alderman, H., Behrman, J. R., Ross, D., & Sabot, R. (1996). The returns to endogenous human capital in Pakistan's rural wage labor market. *Oxford Bulletin of Economics and Statistics*, *58*, 29–55.

Altinok, N., Angrist, N., & Patrinos, H. A. (2018). Global data set on education quality (1965–2015). Policy Research Working Paper No. WPS 8314. Washington, DC: World Bank.

Angrist, J., & Krueger, A. B. (1991). Does compulsory schooling affect schooling and earnings? *Quarterly Journal of Economics*, *106*, 979–1014.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*(4), 612–637.

Barnett, S. M., & Williams, W. M. (2004). National intelligence and the emperor's new clothes. *Contemporary Psychology*, *49*, 389–396.

Barnett, S. M., & Williams, W. M. (2005). IQ-income data do not prove poor countries must remain poor. *PsycCRITIQUES*, *50*(13).

Barnett, W. S., & Boocock, S. S. (Eds.). (1998). *Early care and education for children in poverty. Promises, programs, and long-term results*. Albany: State University of New York Press.

Becker, D., & Rindermann, H. (2016). The relationship between cross-national genetic distances and IQ-differences. *Personality and Individual Differences*, *98*, 300–310.

Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2007). The effect of schooling on psychometric intelligence: Does school quality make a difference? Unpublished paper. Berlin: Max-Planck-Institute for Human Development.

Bond, R., & Saunders, P. (1999). Routes of success: Influences on the occupational attainment of young British males. *British Journal of Sociology*, *50*, 217–249.

Burhan, N. A. S., Kurniawan, Y., Sidek, A. H., & Mohamad, M. R. (2014). Crimes and the Bell curve: The role of people with high, average, and low intelligence. *Intelligence*, *47*, 12–22.

Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development*, *60*, 1239–1249.

Canfield, R. L., Henderson, C. R., Jr., Cory-Slechta, D. A., Cox, C., Jusko, T. A., & Lanphear, B. P. (2003). Intellectual impairment in children with blood lead concentrations below 10 microg per deciliter. *New England Journal of Medicine*, *348*, 1517–1526.

Caplan, B. (2017). *The case against education*. Princeton: University of Princeton Press.

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam: Elsevier.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, *27*, 703–722.

Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist*, *52*, 1051–1058.

Ceci, S. J., & Williams, W. M. (2009). Should scientists study race and IQ? Yes: The scientific truth must be pursued. *Nature*, *457*, 788–789.

Charlton, B. G. (2009). Replacing education with psychometrics. *Medical Hypotheses*, *73*, 273–277.

Coyle, T. R., Rindermann, H., & Hancock, D. (2016). Cognitive capitalism: Economic freedom moderates the effects of intellectual and average classes on economic productivity. *Psychological Reports*, *119*, 411–427.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education*, Vol.1 (pp. 697–812). Amsterdam: North-Holland.

Deary, I. J., Batty, G. D., & Gale, C. R. (2008). Bright children become enlightened adults. *Psychological Science*, *19*, 1–6.

Denny, K., & Doyle, O. (2008). Political interest, cognitive ability and personality: Determinants of voter turnout in Britain. *British Journal of Political Science*, *38*, 291–310.

Ellis, L., & Walsh, A. (2003). Crime, delinquency and intelligence: A review of the worldwide literature. In H. Nyborg (Ed.), *The scientific study of general intelligence. Tribute to Arthur R. Jensen* (pp. 343–365). Oxford: Pergamon.

Eysenck, H. J., & Schoenthaler, S. J. (1997). Raising IQ level by vitamin and mineral supplementation. In R. J. Sternberg & E. L. Grigorenko (Eds.), *Intelligence, heredity and environment* (pp. 363–392). Cambridge: Cambridge University Press.

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn effect*. Cambridge: Cambridge University Press.

Flynn, J. (2018). Reflections about intelligence over 40 years. *Intelligence*, *70*, 73–83.

Ganzach, Y. (2017). Cognitive ability and party affiliation: The role of the formative years of political socialization. *Intelligence*, *61*, 56–62.

Glaeser, E., Ponzetto, G., & Shleifer, A. (2007). Why does democracy need education? *Journal of Economic Growth*, *12*, 77–99.

Glewwe, P., & Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (Vol. 2, pp. 945–1017). Amsterdam: North-Holland.

Goodenough, F. L. (1926). *Measurement of intelligence by drawings*. Chicago: World Book.

Gottfredson, L., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science*, *13*, 1–4.

Hallpike, C. R. (1980). *The foundations of primitive thought*. Oxford: Oxford University Press.

Hansen, K. T., Heckman, J. J., & Mullen, K. J. (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics*, *121*, 39–98.

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, *46*, 607–668.

Hanushek, E. A., & Woessmann, L. (2015). The knowledge capital of nations. Education and the economics of growth. Cambridge, MA: MIT Press.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, *73*, 103–130.

Harris, D. B. (1963). *Children's drawings as measures of intellectual maturity*. New York: Harcourt, Brace & World.

Hart, M. (2007). *Understanding human history: An analysis including the effects of geography and differential evolution*. Augusta, GA: Washington Summit.

Heckman, J. J., & Masterov, D. V. (2007). The productivity argument for investing in young children. NBER Working Paper No. 13016. Cambridge, MA: National Bureau of Educational Research.

Henrich, J., Heine, S., & Norenzayan, A. (2010). The WEIRDEST people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83.

Holland, P. W., & Wainer, H. (Eds.). (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hunt, E. (2012). What makes nations intelligent? *Perspectives on Psychological Science*, *7* (3), 284–306.

Hunt, E., & Carlson, J. (2007). Considerations relating to the study of group differences in intelligence. *Perspectives on Psychological Science*, *2*, 194–213.

Jones, G. (2016). *Hive mind: How your nation's IQ matters so much more than your own*. Stanford: Stanford University Press.

Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research*, *78*, 85–123.

Kura, K. (2013). Japanese north-south gradient in IQ predicts differences in stature, skin color, income, and homicide rate. *Intelligence*, *41*, 512–516.

Lakhanpal, M., & Ram, R. (2008). Educational attainment and HIV/AIDS prevalence: A cross-country study. *Economics of Education Review*, *27*, 14–21.

Luria, A. R. (1976). *Cognitive development. Its cultural and social foundations*. Cambridge, MA: Harvard University Press.

Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the development quotients of infants. *Intelligence*, *37*, 16–24.

Lynn, R. (2010). National IQs updated for 41 nations. *Mankind Quarterly*, *50*, 275–296.

Lynn, R., & Meisenberg, G. (2010). The average IQ of sub-Saharan Africans: Comments on Wicherts, Dolan, and van der Maas. *Intelligence*, *28*, 21–29.

Lynn, R., & Vanhanen, T. (2001). National IQ and economic development: A study of eighty-one nations. *Mankind Quarterly*, *41*, 415–435.

Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations*. Westport, CT: Praeger.

Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality*. Augusta GA: Washington Summit Publishers.

Lynn, R., & Vanhanen, T. (2012). *Intelligence. A unifying construct for the social sciences*. London: Ulster Institute for Social Research.

Meisenberg, G. (2012). National IQ and economic outcomes. *Personality and Individual Differences*, *53*, 103–107.

Meisenberg, G. (2015). Verbal ability as a predictor of political preferences in the United States, 1974–2012. *Intelligence*, *50*, 135–143.

Melhuish, E. C., Phan, M. B., Sylva, K., Sam-mons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues*, *64*, 95–114.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015: International results in mathematics*. Chestnut Hill, MA: International Study Center.

Murnane, R. J., Willet, J. B., & Levy, F. (1995). The growing importance of reading comprehension cognitive skills in wage determination. *Review of Economics and Statistics*, *77*, 251–266.

Murray, Ch. (2008). *Real education: Four simple truths for bringing America's schools back to reality*. New York: Crown Forum.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Nisbett, R. E. (2009). *Intelligence and how to get it. Why schools and cultures count*. New York: W. W. Norton.

Oesterdiekhoff, G. W. (2014). Evolution of law and justice from ancient to modern times. *Journal on European History of Law*, *5*, 54–64.

Piffer, D. (2015). A review of intelligence GWAS hits: Their relationship to country IQ and the issue of spatial autocorrelation. *Intelligence*, *53*, 43–50.

Psacharopoulos, G., & Patrinos, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, *12*, 111–134.

Radding, C. (1985). *A world made by men: Cognition and society, 400–1200*. Chapel Hill: University of North Carolina Press.

Rindermann, H. (2018). *Cognitive capitalism: Human capital and the wellbeing of nations*. Cambridge: Cambridge University Press.

Rindermann, H., & Baumeister, A. E. E. (2015). Validating the interpretations of PISA and TIMSS tasks: A rating study. *International Journal of Testing*, *15*, 1–22.

Rindermann, H., Becker, D., & Coyle, T. R. (2016). Survey of expert opinion on intelligence: Causes of international differences in cognitive ability tests. *Frontiers in Psychology*, *7*, 1–9.

Rindermann, H., & Carl, N. (2018). Human rights: Why countries differ. *Comparative Sociology*, *17*, 29–69.

Rindermann, H., & Ceci, S. J. (2009). Educational policy and country outcomes in international cognitive competence studies. *Perspectives on Psychological Science*, *4*, 551–577.

Rindermann, H., & Ceci, S. J. (2018). Parents' education is more important than their wealth in shaping their children's intelligence: Results of 19 samples in seven countries at different developmental levels. *Journal for the Education of the Gifted*, *41*, 298–326.

Rindermann, H., Sailer, M., & Thompson, J. (2009). The impact of smart fractions, cognitive ability of politicians and average competence of peoples on social development. *Talent Development and Excellence*, *1*, 3–25.

Roivainen, W. (2010). European and American WAIS III norms: Cross-national differences in performance subtest scores. *Intelligence*, *38*, 187–192.

Rushton, J. P., Bons, T. A., Vernon, P. A., & Cvorovic, J. (2007). Genetic and environmental contributions to population group differences on the Raven's Progressive Matrices estimated from twins reared together and apart. *Proceedings of the Royal Society*, *274*, 1773–1777.

Rushton, J. P., & Templer, D. I. (2009). National differences in intelligence, crime, income, and skin color. *Intelligence*, *37*, 341–346.

Sandefur, J. (2016). Internationally comparable mathematics scores for fourteen African countries. Working Paper No. 444. Center for Global Development.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, *87*, 355–374.

Stelzl, I., Merz, F., Remer, H., & Ehlers, T. (1995). The effect of schooling on the development of fluid and crystallized intelligence: A quasi-experimental study. *Intelligence*, *21*, 279–296.

Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist*, *59*, 325–338.

Strauss, J., & Thomas, D. (1998). Health, nutrition, and economic development. *Journal of Economic Literature 36*, 766–817.

Thomson, G. H. (1937). Intelligence and civilisation: A Ludwig Mond lecture delivered at the University of Manchester on October 23, 1936. *Journal of the University of Manchester*, *1*, 18–38.

Warne, R. T., & Burningham, C. (2019). Spearman's g found in 31 non-Western nations: Strong evidence that g is a universal phenomenon. *Psychological Bulletin*, *145*, 237–272.

Weede, E., & Kämpf, S. (2002). The impact of intelligence and institutional improvements on economic growth. *Kyklos*, *55*, 361–380.

Whetzel, D. L., & McDaniel, M. A. (2006). Prediction of national wealth. *Intelligence*, *34*, 449–458.

Wicherts, J. M., Borsboom, D., & Dolan, C. V. (2010). Evolution, brain size, and the national IQ of people around 3000 years B.C. *Personality and Individual Differences*, *48*, 104–106.

Wicherts, J. M., & Dolan, C. V. (2010). Measurement invariance in confirmatory factor analysis: An illustration using IQ test performance of minorities. *Educational Measurement: Issues and Practice*, *29*(3), 39–47.

Wicherts, J. M., Dolan, C. V., Carlson, J. S., & van der Maas, H. L. J. (2010a). Raven's Test performance of sub-Saharan Africans: Average performance, psychometric properties, and the Flynn effect. *Learning and Individual Differences*, *20*, *135–151*.

Wicherts, J. M., Dolan, C. V., Carlson, J. S., & van der Maas, H. L. J. (2010b). Another failure to replicate Lynn's estimate of the average IQ of sub-Saharan Africans. *Learning and Individual Differences*, *20*, *155–157*.

Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. J. (2010a). A systematic literature review of the average IQ of sub-Saharan Africans. *Intelligence*, *38*, 1–20.

Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. J. (2010b). The dangers of unsystematic selection methods and the representativeness of 46 samples of African test-takers. *Intelligence*, *38*, 30–37.

Wober, M. (1969). Distinguishing centri-cultural from cross-cultural tests and research. *Perceptual and Motor Skills*, *28*, 488.

Woodworth, K. R., David, J. L., Guha, R., Wang, H., & Lopez-Torkos, A. (2008). San Francisco Bay Area KIPP schools: A study of early implementation and achievement. Final report. Menlo Park, CA: SRI International.

41 Environment and Intelligence

David Bellinger

In 2016, recognizing that human activities have profoundly affected the Earth's climate, geology, and ecosystems, the International Geological Congress declared that the Holocene era had ended, and that we had entered a new age called the Anthropocene. One reason was the increasing extent to which present-day humans, compared to our ancient ancestors, are exposed to potentially hazardous chemicals. Although humans have mined and worked elemental chemicals, such as lead and mercury, for thousands of years, the pace with which such chemicals are dispersed into sectors of the environment with which humans come into intimate contact has increased steadily over the last two millennia. The rate of increase accelerated dramatically over the past 250 years, following the onset of the industrial revolution. Moreover, because of the development in recent decades of methods for producing complex synthetic chemicals such as plastics, flame retardants, and pesticides, humans are now exposed to chemicals that, from an evolutionary perspective, are entirely novel to biological systems. The detoxification strategies that humans gradually developed to mitigate the impacts of natural chemicals are not always effective in protecting against such exposures.

Although environmental chemicals can impair the function of many organ systems, the central nervous system (CNS) appears to be especially vulnerable. Six of the ten chemicals identified by the World Health Organization (WHO) as being of greatest public health concern are known to adversely affect brain-based functions and their development (air pollution, arsenic, dioxin- and dioxin-like compounds, lead, mercury, pesticides) (WHO, 2019), with some data implicating at least two others (cadmium, fluoride). The goal of this chapter is to describe, in brief, important stages of brain development and to provide examples of how exposure to environmental chemicals during a critical stage of this process can alter the course of development in ways that might impact an individual's intelligence. Many of the examples involve lead because the body of data available for lead is much more extensive than for any other chemical. The principles induced on the basis of lead are broadly applicable to other chemicals, however.

One reason for the enhanced vulnerability of the developing brain to exogenous insults is the complexity of the processes that occur over a protracted period, beginning shortly after conception and lasting for at least two decades. Brain development is characterized by an exquisite temporal and spatial choreography of processes that result in an organ that consists of billions of precisely located, highly interconnected, specialized cells. Many factors can interfere with the developmental processes involved in brain development during the prenatal and postnatal years, including pathogens (viruses such as zika, cytomegalovirus, rubella; bacteria such as syphilis and *Neisseria meningitidis*; protozoa such as *Toxoplasma gondii* and schistosoma).

In the sixteenth century, the Swiss physician Paracelsus articulated the most famous dictum in toxicology: "Solely the dose determines that a thing is not a poison." In other words, at a sufficient dose, any chemical becomes toxic to biological systems whereas, at lower doses, it might be harmless. Based on the temporal choreography of brain development, described in the preceding paragraph, it is evident that Paracelsus' dictum must be amended to indicate that not only the dose but the timing of exposure with regard to the stage of development determines a chemical's neurotoxicity. Nevertheless, Paracelsus' contribution provided the foundation for fundamental concepts of toxicology such as the dose-response relationship and threshold for toxicity. The major questions, then, in evaluating the magnitude of the public health hazard posed by a chemical, pertain to whether the doses incurred by members of the general population, at critical times in brain development, exceed the dose known to cause harm.

The major stages of development during prenatal life are primary neurulation (occurring during weeks 3 and 4), development of the forebrain (prosencephalon) (months 2–3), neuronal proliferation (months 3–4), neuronal migration (months 3–5), neuronal organization (later gestation and in the period following birth), and myelination (beginning in mid-pregnancy and continuing postnatally into young adulthood). The sequential nature of the processes involved and the long period over which they occur provide the basis for hypotheses that "critical windows of vulnerability" exist, such that the impact of exposure to a certain chemical will differ depending on when it occurs in the sequence.

The coordination of these complex processes is regulated by myriad signaling pathways that must work as evolution intended if an individual is to end up with a species-typical brain. Although minor variations occur across individuals in the fidelity of these processes and are inconsequential, some variations can result in abnormalities that might have substantial impact on an individual's abilities to carry out brain-based functions.

Which aspects of CNS development a chemical perturbs, and thus when exposure is likely to be most deleterious to a child, depends on its mechanism of action. Many chemicals affect multiple aspects of brain development. Alcohol, methylmercury, and chlorpyrifos, an organophosphate pesticide, disturb neural cell proliferation, while methylmercury, alcohol, and X-irradiation affect cell migration. Differentiation of neuroblasts is affected by alcohol, nicotine, methylmercury, and lead. The creation of glial cells and subsequent myelination is affected by endocrine-disrupting chemicals, alcohol, lead, and postnatal malnutrition. Synaptogenesis is affected by alcohol, polychlorinated biphenyls, triethyltin, and the pesticides parathion and permethrin, while neuroapoptosis (the orderly process of programmed cell death) is affected by lead, alcohol, and methylmercury. Many chemicals affect neurotransmitter systems, including organophosphate pesticides, which are designed to be cholinesterase inhibitors, alcohol, lead, methylmercury, aluminum, and certain pharmaceutical agents such as many antidepressants (e.g., selective serotonin reuptake inhibitors).

The fetus is not always fully protected from environmental chemicals by the placenta or the blood-brain barrier. Many chemicals, such as lead, passively diffuse across the placenta so that the concentration in fetal blood is approximately the same as in maternal blood (Aylward et al., 2014). Moreover, not only is a fetus exposed to whatever lead a woman is exposed to during the course of her pregnancy but they are also exposed to lead to which the woman was exposed in the past. To support ossification of the fetus (development of the skeleton), calcium is mobilized from maternal bone in large quantities during pregnancy. Because most of the lead in an adult's body is stored in bone, and both lead and calcium are divalent cations (i.e., have a valence of +2), lead, as well as calcium, is released from bone. It is estimated that this accounts for up to 70 percent of the lead in fetal blood. In the case of methylmercury, the concentration in fetal blood is approximately 70 percent greater than in maternal blood, consistent with active transport of this chemical across the placenta. Other chemicals, such as cadmium, do not cross the placenta but can accumulate in this tissue and impair its support functions. A study of a nationally representative sample of pregnant women in the United States showed that certain polychlorinated biphenyls, organochlorine pesticides, perfluorinated chemicals, phenols, polybrominated diphenyl ethers (flame retardants), phthalates (plasticizers), polycyclic aromatic hydrocarbons (produced by burning organic materials), and perchlorate were detected in 99–100 percent of pregnant women (Woodruff, Zota, & Schwartz, 2011).

A fully developed blood-brain barrier, which consists of tight junctions between the endothelial cells lining cerebral microvessels, prevents larger water-soluble chemicals from entering the brain (Zheng, Aschner, & Ghersi-Egea, 2003). Fat-soluble chemicals, such as alcohol, may cross, however. This barrier is not fully developed at birth and studies in nonhuman primates using radioactive tracers demonstrate that chemicals pass from the circulating blood into the brain more readily in immature than mature individuals.

Certain behavioral and physiologic factors also place a developing child at greater risk than an adult to the deleterious effects of chemical exposures (Selevan, Kimmel, & Mendola, 2000). First, certain pathways of exposure are unique to children, including transplacental passage and breastfeeding. It is primarily fat-soluble chemicals (e.g., dioxins, PCBs, perfluorinated compounds) that are of concern with regard to ability to pass into breast milk. Young children engage in behaviors that can potentially bring them into more intimate contact than adults with toxic chemicals. Behaviors more common among children, including hand-to-mouth activity, oral exploration of objects, and nonnutritive ingestion (i.e., pica) will result in greater exposure in an environmental that is contaminated by chemicals present in household dust and in soil. On a body weight basis, they consume more food and breathe a greater volume of air than adults and so will experience greater exposures than adults to foodborne and airborne hazards. Children and adults tend to experience different breathing zones. Children spend more time near the floor, where chemical concentrations in the air might be greater as a result of residential pesticide application. Differences between the diets of children and adults can also result in different exposures. Because children's relative consumption of fruit juices is typically greater than that of adults, the presence of pesticide residues on these products can pose a greater risk to them than to adults. Certain micronutrient deficiencies, such as iron and calcium, are more common in children, which can result in greater fractional absorption of chemicals with which these essential metals share binding sites in the gut. For example, children absorb up to 50 percent of ingested lead, whereas adults absorb approximately 10 percent. Finally, some of the detoxification pathways in the liver are not fully developed in the early postnatal years. Specifically, the xenobiotic biotransforming enzymes that convert lipid-soluble compounds into water-soluble metabolites that can be excreted in the urine are less effective in children, with the result that parent compounds that can damage cellular processes remain active in children for a longer period following exposure.

Although acute poisonings as a result of a single, high-dose exposure to a chemical can be serious, and in some cases even fatal, a broader concern is the potential neurological impact of children's chronic, low-dose exposure to chemicals that, in some cases, are ubiquitous in the human environment. The specific concern is that, while such exposures might not produce clinical signs of intoxication, they nevertheless adversely impact brain function, increasing a child's risk of reduced intellectual capacities. The following sections discuss in more detail what is known about the effects of particular chemicals or classes of chemicals on children's intellectual development.

Metals

The heavy metal methylmercury provides a striking example of age-dependent vulnerability. Industrial discharge of mercury salts into Minamata Bay in Japan produced what came to be called Minamata disease. Children born to women who, during their pregnancies, consumed large amounts of seafood contaminated by methylmercury suffered a distinctive constellation of signs after birth that was called Congenital Minamata disease (CMD) (Harada, 1995). This included growth disturbances, retention of primitive reflexes, sensory impairments, intellectual disability (tenfold increase in risk), cerebral palsy (fiftyfold increase in risk), and movement and coordination disorders such as cerebellar ataxia (loss of control of movements), chorea (jerky, involuntary movements), athetosis (involuntary writhing movements), and dysarthria (difficulties with articulation). A special hospital was built in Minamata City solely to provide lifelong care for these children, as they were unable to function independently. It was striking that many of the mothers of children with CMD manifested no symptoms of mercury intoxication or only mild sensory symptoms such as paresthesia (burning or prickly sensations). Neuropathological examination of individuals who were of different ages when their mercury exposure began revealed very different patterns of brain abnormalities (Choi, 1989). In individuals who had already reached adulthood at the onset of exposure, the lesions were highly localized, clustering in the pre- and postcentral gyri, the calcarine fissure of the occipital cortex, and the cerebellum. This is consistent with the clinical signs of adult intoxication, which include movement disorders, tremors, sensory disturbances, and constriction of the visual fields. In individuals exposed throughout gestation, however, lesions were found throughout the brain, with no apparent localization. This is consistent with the global developmental delay characteristic of patients with CMD. One reason the impacts are diffuse rather than focal in fetuses is that exposure to methylmercury arrests mitotic cells in metaphase, impairing the cytoskeletal proteins (microtubule assemblies) that form the mitotic spindle. As a result, cell proliferation and migration are perturbed, producing widespread abnormalities in the developing brain, including heterotopias (islands of cells in the wrong location), reduced cell densities, anomalous cytoarchitecture, disturbance in the laminar pattern of cerebral cortex, incomplete myelination, glial proliferation, and limited gyral differentiation.

Consumption of seafood is the major pathway of exposure of the general population to methylmercury. Inorganic mercury that is dispersed into the environment by both natural (e.g., volcanoes, forest fires) and industrial processes (e.g., combustion of fossil fuels) settles into waterbodies, where it is biotransformed (specifically methylated) by microbes in the sediments. It enters the aquatic food chain and bioconcentrates in tissues as it ascends trophic levels. The concentrations of methylmercury are therefore greatest in long-lived predatory species (e.g., whales, shark, swordfish, albacore tuna). The devastating neurological effects observed in poisoning episodes such as occurred in Minamata and elsewhere stimulated concern that low-level chronic prenatal exposure to methylmercury has less serious but still deleterious effects on children's brain development. Therefore, numerous studies have been conducted in cohorts for whom seafood is an important component of the diet. The population of the Faroe Islands (North Atlantic Ocean) consumes pilot whale as well as other types of seafood. A large prospective study showed that the performance of children on tests of language, attention, and memory was inversely related to blood mercury concentrations of their mothers during pregnancy (Grandjean et al., 1997). In this cohort, each increase of 1 μg/g (part per million) in the maternal hair mercury is associated with, on average, a loss of about one-half of an IQ point (Bellanger et al., 2013). Functional magnetic resonance imaging (fMRI) studies of a subset of the cohort showed dose-related alterations in patterns of activation (White et al., 2011). Studies conducted in areas in which local seafood is heavily contaminated originating due to the use of mercury as an amalgamator in artisanal gold mining have found deficits among the children with the greatest prenatal exposure to methylmercury (Gibb & O'Leary, 2014). Some large studies of populations of fish consuming methylmercury have not found significant associations with children's cognitive development, however (e.g., Davidson et al., 1998). One source of the inconsistency has to do with the vehicle of exposure. Consumption of seafood is a source not only of methylmercury but of a variety of macronutrients (e.g., protein) and micronutrients (e.g., selenium, choline) that promote brain development. Failure to structure statistical analyses in such a way that adjustments are made can obscure the detection of methylmercury toxicity, that is, produce negative confounding (Choi et al., 2008). More sophisticated analyses have shown that, by careful selection of the particular species of fish that are consumed, it is possible to both achieve the benefits of these nutrients and minimize exposure to methylmercury, benefiting the cognitive outcomes of children (e.g., Oken et al., 2005).

Lead can produce devastating effects on the developing fetus by interfering with myriad processes. Depending on the dose, it increases apoptosis and excitotoxicity; reduces cellular energy metabolism by impairing the functioning of mitochondria; reduces heme synthesis so that the oxygen-carrying capacity of red blood cells is reduced; increases oxidative stress and lipid peroxidation thus damaging cell membrane lipids; alters the activity of first and second messenger systems in neurons, receptor densities, and dendritic branching patterns; impairs the development and function of oligodendroglia, resulting in abnormal myelin formation; disturbs neurotrophic processes, including thyroid transport into brain; and alters the regulation of gene transcription. Severe lead poisoning can cause brain hemorrhage, edema, seizures, and coma. Even children whose lead poisoning is not so severe as to cause such an encephalopathy are left with a variety of residual difficulties. In an early case series, Byers and Lord (1943) demonstrated the error of the view, widespread at the time, that children fully recover from subencephalopathic lead poisoning. They observed that, "after recovery from their lead poisoning, these . . . children made an extremely poor record in competition with their fellows," insofar as, "with one definite and a second possible exception, none of the 20 children succeeded in school" (p. 479). They also noted that the children exhibited severe behavioral pathologies, including hyperactivity, reduced impulse control, and aggression (e.g., fire-setting, biting others, stealing supplies, repeatedly dancing on desks, attacking classmates). Other problem behaviors noted included cruelty to animals and a lack of response to punishment.

Based on the substantial amount of evidence that accrued on lead neurotoxicity over the past few decades, the blood lead concentration considered to be the "upper limit of normal" by bodies such as the WHO and the US Centers for Disease Control and Prevention (CDC) steadily dropped. In the 1960s, a value of 60 μg/dL was considered the limit but, at present, because of the apparent absence of a threshold for the appearance of adverse effects on children's cognition and behavior, no level is considered to be "safe." A set of analyses of IQ data collected in seven prospective studies ($N = 1{,}333$) conducted in several countries provided much of the justification for the present consensus (Lanphear et al., 2005). The motivation for pooling the data from these studies was to increase the precision with which the shape of the dose-effect relationship could be ascertained for blood lead concentrations <10 μg/dL. At the time of these analyses, 10 μg/dL was the US CDC "action" level, at which intervention activities were triggered to reduce a child's lead exposure. The relationships between IQ, measured at age five to ten years, and four indices of lead exposure were evaluated: the blood lead concentration measured the closest in time to the IQ assessment (concurrent), the maximum blood lead concentration prior to IQ measurement (peak), the mean blood lead concentration between six months of age and IQ measurement (lifetime average), and the mean level between six and twenty-four months of age (early childhood). A variety of statistical models were compared in terms of goodness-of-fit to the data, adjusting for ten covariates, such as maternal IQ and home quality. A log-linear model involving concurrent blood lead concentration provided the best fit and indicated that an increase in blood lead from 2.4 to 30 μg/dL (the 5th and 95th percentiles of the blood lead distribution in the pooled dataset) was associated with an IQ reduction of 6.9 points. One of the most important findings was that the association was nonlinear, insofar as much of the total reduction, 3.9 points, occurred over the range of 2.4 to 10 μg/dL. The increase from 10 to 20 μg/dL was associated with a further reduction of 1.9 points and the increase from 20 to 30 μg/dL with an additional reduction of 1.1 points. Although the mechanism(s) by which the proportional loss in IQ is greater at lower than at higher blood lead concentrations remains unknown, similar supralinear relationships between blood lead and other cognitive outcomes were subsequently reported (e.g., Kordas et al., 2006; Tellez-Rojo et al., 2006).

Epidemiological studies have generally included assessment of other aspects of children's neuropsychological functioning. IQ is an apical measure, integrating children's performance in diverse verbal and nonverbal domains. If the effects of a neurotoxicant are focal, limited to only certain domains, one would expect that tests that focus more narrowly on the vulnerable domains would be more strongly related to lead biomarkers than is IQ. It is somewhat surprising, therefore, that IQ is more consistently associated with lead biomarkers than performance on tests that focus on more narrowly defined domains (e.g., memory, language). This might provide some insight into the mechanism of lead-associated neurotoxicity, suggesting diffuse rather than focal neuronal and/or white matter injury, thereby affecting global higher cortical functioning. It is also possible that the domains most affected vary across studies because of the joint effect of cohort differences in the timing of lead exposure and differences across domains in the timing of greatest susceptibility. The domains vulnerable to lead exposure also depend, to some extent, on cohort characteristics or context-specific aspects such as genetic susceptibility, the social environment, and mixed chemical exposures.

Many studies have reported significant inverse associations, in the general population of children, between lead exposure and success in school, expressed as lower scores on standardized tests, receipt of special education services, grade retention, and failure to complete qualifications (e.g., Amato et al., 2012; Delgado et al., 2017; Fergusson, Horwood, & Lynskey, 1997; Magzamen et al., 2013; Magzamen et al., 2015; Needleman et al., 1990). In ecologic analyses, Nevin (2009) found an inverse relationship between preschool blood lead concentrations and SAT scores, lagged by seventeen years. Evens and colleagues (2015) evaluated the relationship between blood lead concentration and performance on the Illinois Standard Achievement Test in a cohort of 47,168 Chicago schoolchildren. Adjusting for covariates, they found dose-dependent reductions, extending below 10 μg/dL, in both reading and math scores, with each 5 μg/dL increase in blood lead associated with an increase of 1.3 in the risk of reading and math scores considered to represent "failure." A nonlinear dose-response relationship was found for reading failure, with the slope steeper for blood lead concentrations less than 10 μg/dL compared to greater than 10.

In a similar study, Miranda and colleagues (2009) combined a North Carolina blood lead surveillance database with a database containing scores on a reading test given to schoolchildren in the state at the end of the fourth grade. In the sample of 57,678 children, blood lead concentration ranged from 1 to 16 μg/dL, with a 75th percentile of 6 μg/dL. A significant dose-response effect relationship was found between blood lead and test score, without apparent threshold. Quantile regression analyses indicated that the adverse impact was more pronounced at the lower end of the test score distribution than at the higher end, indicating that the impact of lead is disproportionally greater on children who are already at academic risk than on children who are at low academic risk. This might be attributable to contextual factors that influence a child's resilience or susceptibility to lead exposure, such as nutrition, social environment, and other chemical exposures.

Skerfving and colleagues (2015) evaluated the performance of 3,176 Swedish children on an examination taken at the end of compulsory schooling (age sixteen years) in relation to blood lead concentration measured when the children were in primary school (age seven to twelve years). They found an inverse nonlinear association between blood lead concentration (range 0.6 to 16.2 μg/dL, 90th percentile 6 μg/dL) and school performance, with a steeper slope at concentrations below 5 μg/dL than above 5. They estimated that the adverse impact of an increase in blood lead concentration from 2.5 to 5 μg/dL was similar in magnitude to the impact of having a mother with a university versus a primary school education. Surkan and colleagues (2007) reported that children with a blood lead concentration of 5–10 μg/dL, compared to children with a concentration of 1–2 μg/dL, had significantly lower reading and mathematics scores, even after adjustment for Full-Scale IQ score. This suggests that the achievements of children with greater exposures were not at a level commensurate with their natural ability.

The effects of lead are persistent as children age, with greater childhood exposure associated with reduced success in life. In the large Dunedin Multidisciplinary Health and Development Study, Reuben and colleagues (2017) found that blood lead concentration measured at eleven years of age was inversely related to IQ at age thirty-eight years, even after adjustment for IQ at age eleven years. Furthermore, early lead exposure was inversely related to socioeconomic status at age thirty-eight, suggesting that children who were more highly exposed enjoyed lower upward social mobility and, in fact, failed to match the socioeconomic achievements of their parents.

One aspect of lead neurotoxicity that receives little attention is the possible role of early-life exposure to lead as, itself, an effect modifier of later events and exposures, or even processes associated with normal aging, by increasing their adverse impacts. For example, Weiss, Clarkson, and Simon (2002) speculated that if early-life exposure to a neurotoxicant increases the annual rate of neuronal loss by less than 1 percent, clinical signs of neurodegeneration would appear several years earlier than they would in the absence of such an exposure. Limited animal studies support the hypothesis that early-life exposure is a risk factor for less optimal resilience in adulthood in response to an unrelated neurological insult. Rats exposed to lead early in development but not thereafter were less successful than control rats in recovering function (beam walking, limb placing) after a laser-induced stroke in the somatosensory cortex in adulthood (Schneider & DeKamp, 2007). Early-life exposure to lead might also influence function in adulthood by altering the trajectory of processes associated with aging. Provocative studies in rodents and nonhuman primates suggest that exposure in infancy initiates epigenetic processes, perhaps involving altered patterns of DNA methylation, that result in adult-onset overexpression of proteins involved in the neurodegenerative processes characteristic of Alzheimer's disease (e.g., increased deposition of β-amyloid, increased hyperphosphorylation of tau protein) (Gąssowska et al., 2016).

In the 1970s, the observations of Byers and Lord (1943) regarding the behavioral pathologies were followed up, with case-control and chelation challenge studies suggesting that children diagnosed with hyperactivity had greater lead burdens (e.g., David, Clark, & Voeller, 1972). Subsequent studies indicated that, even in children who were neither clinically lead poisoned nor diagnosed with hyperactivity, a greater lead burden was associated with increased distractibility, a reduced ability to follow directions, disorganization, daydreaming, and a lack of task persistence (e.g., Needleman et al., 1979). A meta-analysis of thirty-three studies conducted between 1972 and 2010, involving more than 10,000 children, found significant effect sizes linking greater exposure and dimensional measures of both inattentive and hyperactive/impulsive symptoms (Goodlad, Marcus, & Fulton, 2013).

It is now established that increased childhood lead exposure also increases the risk that a child will meet diagnostic criteria for attention-deficit/hyperactivity disorder (ADHD). Using data from the National Health and Nutrition Examination Survey (NHANES) 1999–2002, Braun and colleagues (2006) found that the adjusted odds ratio for parent-reported ADHD among six- to sixteen-year-old children with a blood lead concentration in the 5th quintile (>2 μg/dL) was 4.1, compared to children in the 1st quintile (<0.8 μg/dL). A dose-response relationship was observed, with adjusted odds ratios of 1.1, 2.1, and 2.7 for children in the 2nd quintile, 3rd, and 4th quintiles, respectively. Among the limitations of this study are the absence of a clinician-confirmed diagnosis of ADHD and the fact that data on important covariates, such as family history of ADHD, were not available. A study by Froehlich and colleagues (2009), using NHANES 2001–2004 data, addressed the first issue, as the Diagnostic Interview Schedule for Children, a clinician-administered diagnostic interview based on DSM-IV, was used to confirm a diagnosis of ADHD. Children with a blood lead concentration in the upper tertile (>1.3 μg/dL) were 2.3 times more likely to meet diagnostic criteria than were children with a concentration in the lowest tertile.

This association has been replicated in several subsequent case-control studies (Choi at al., 2016; Park et al., 2016; Wang et al., 2008). Nigg and colleagues (2008) clarified possible behavioral mechanisms of the association between lead and ADHD symptoms. Children eight to seventeen years old who met rigorous criteria for the diagnosis of ADHD had a significantly higher blood lead concentration than controls, even though concentration for all participants ranged only from 0.4 to 3.5 μg/dL (mean of 1.03). A significant relationship was found between blood lead concentration and total ADHD symptoms. In this study, IQ was measured and a Stop task was administered, providing assessments of a child's ability to suppress a prepared response (stop signal reaction time) and variability of reaction time on the "go" trials (response variability, readiness, and control). Mediation analyses suggested that lead exposure might increase a child's risk of ADHD by impairing cognitive control abilities and that the association between blood lead and IQ was mediated by the association between blood lead and hyperactivity-impulsivity, not vice versa. Nigg and colleagues argued that the plausibility of the link between lead and ADHD is supported by the evidence that lead disrupts midbrain dopamine circuitry (striatum and frontostriatal networks), the same circuitry that is thought to underlie ADHD. Nigg and colleagues (2010) replicated the associations between blood lead concentration and ADHD symptoms in a larger sample, including adjustment for additional covariates, and at even lower blood lead concentrations (range 0.3–2.2 μg/dL).

Several studies in the last decade have suggested that children with greater early-life lead exposure are at increased risk of social pathologies such as criminal activities (e.g., Boutwell et al., 2017; Wright et al., 2008). It seems most likely that this represents a developmental cascade, with reduced intelligence, reduced school success, behavioral impairments such as reduced impulse control and ADHD, and possible substance abuse, leading some children to make poor decisions (Bellinger, Matthews-Bellinger, & Kordas, 2016).

Neuroimaging studies have explored the associations between lead exposure history and brain structure and function among individuals from the general population. Most of the data are from the Cincinnati Prospective Lead Study in which participants were enrolled prenatally and followed into young adulthood (nineteen to twenty-four years of age). Volumetric imaging revealed significant inverse associations between annual mean blood lead concentration between three and six years of age and gray matter volume, particularly in the frontal regions of the brain, including the anterior cingulate cortex and the ventrolateral prefrontal cortex, areas usually considered to be related to executive functions, mood regulation, and decision-making (Cecil et al., 2008). On diffusion tensor imaging, reduced fractional anisotropy and axial diffusivity were also associated with greater childhood lead exposure (Brubaker et al., 2009). These findings suggest impaired myelination and reduced axonal integrity in regions that regulate executive functions.

Brain function in adulthood is also inversely related to childhood lead exposure. Proton magnetic resonance spectroscopy studies showed that greater childhood blood lead concentration was associated with reductions in several metabolites, including N-acetyl aspartate and creatine and phosphocreatine, in both gray and white matter (Cecil et al., 2011). During a verb generation task, fMRI imaging showed that individuals with greater blood lead concentrations in childhood showed dose-dependent changes in activation pattern in the left frontal cortex and the left middle temporal gyrus (Yuan et al., 2006). In another cohort, in individuals with greater lead exposure, fMRI showed reduced activation in the dorsolateral prefrontal cortex, ventrolateral prefrontal cortex, pre-supplementary motor areas, and inferior parietal cortex on the Wisconsin Card Sorting Test and the n-back task (particularly on trials that imposed a greater memory load) (Seo et al., 2015). These findings suggest that exposure to lead impairs the frontoparietal working memory network.

Pesticides

All pesticides are neurotoxicants because they act by targeting the functioning of the activity of the CNS. Because the CNSs of insects and mammals share many features, these chemicals are not species-selective with regard to targets of toxicity. There are many different classes of pesticides, differing in their modes of action and in their toxicities. Organochlorine pesticides (e.g., DDT) were developed in the first half of the twentieth century. They are fat-soluble pesticides that accumulate in the food chain, remain in the environment for long periods, and bioaccumulate. They act by altering the electrophysiological properties of cell membranes (particularly axons), disturbing sodium and potassium ion exchange. Because of their toxicity and persistence in the environment, their use has largely been banned or restricted in most developed countries in the past several decades, although they are still used in certain regions of the world. Compounds that degrade more rapidly than organochlorines were introduced in the mid-twentieth century (e.g., organophosphates, carbamates). The organophosphates were originally developed as nerve gas agents in Germany during World War II and later put to use as insecticides. They are generally considered to be less toxic than organochlorines and are widely used on food crops and in homes, parks, schools, and golf courses, resulting in human exposure. Organophosphate pesticides inhibit the activity of acetylcholinesterase, the enzyme that catalyzes the breakdown of the neurotransmitter acetylcholine. However, it has recently been learned that certain organophosphate pesticides have adverse impacts on children's neurodevelopment at doses that do not cause acetylcholinesterase inhibition, working by a different mechanism, such as inflammation, C-reactive protein receptor signaling, insulin resistance, and interference with nuclear transcription factor functioning. Another major class of pesticides, the pyrethroids, were developed in the 1970s. Still another class of pesticides, the neonicotinoids, were introduced in the 1980s, although controversy quickly arose because they were implicated in colony collapse disorder. This involves a massive die-off of workers in a honeybee colony, with serious implications for the pollination of food crops.

In the last decades, a substantial number of studies, conducted in both cohorts presumed to have greater pesticide exposures due to the location of their residence and in general population cohorts, have reported that pesticide body burden is inversely related to children's IQ scores. In a group of children living in the agricultural Salinas Valley of California, Bouchard and colleagues (2011) showed that offspring of mothers with a concentration of the dialkyl phosphate metabolites of organophosphate pesticides greater than 50 nmol/L during pregnancy had lower IQ scores at age seven years than the children of mothers with concentrations less than 50 nmol/L, which corresponds to the mean in US women of reproductive age. In an urban New York City cohort, Rauh and colleagues (2011) found an inverse association between the concentration of the organophosphate pesticide chlorpyrifos in umbilical cord blood plasma and child IQ and working memory at age seven years. In this same cohort, morphometric magnetic resonance imaging (MRI) analysis showed dose-related perturbations in the volumes of many regions of the brain

(Rauh et al., 2012). Recently, Eskenazi and colleagues (2018) reported inverse relationships in a rural, agricultural cohort between prenatal exposure to pyrethroid pesticides and children's social-emotional scores at age one year and language/expressive communication scores at age two years.

Air Pollution

The findings of the rapidly developing body of literature on exposure to air pollutants provide confidence in the conclusion that pollutants produce subclinical impacts on children's cognition (Clifford et al., 2016). Approximately 95 percent of the world's population live in areas where outdoor fine particulate matter (particles less than 2.5 microns in diameter) concentrations (dust or soot particles) exceed WHO's Air Quality Guideline of 10 $\mu g/m^3$, with most areas of Africa, the Middle East, and South Asia exceeding 35 $\mu g/m^3$ (Health Effects Institute, 2019). In addition, more than one-third of the world's population is exposed to potentially hazardous indoor air pollution as a result of the combustion of biomass (i.e., wood, dung, peat, crop, wastes) or coal. Because of the closed space and the large amount of time spent indoors, especially during winter months, the indoor concentrations of particulate matter can be as much as twentyfold greater than outdoor concentrations. Air pollution is a complex exposure, the composition of which differs depending on what the different components are in a local area. Components of the mixture include polycyclic aromatic hydrocarbons (which are produced by the incomplete combustion of organic matter), oxides of sulfur, nitrogen, and carbon, ozone, and metals. In the United States, the concentrations of many air pollutants have steadily declined in recent years, although the concentration of ozone and short-term particle pollution have increased (American Lung Association, 2019). Small particles (smaller than 2.5 micrometers in diameter) are generally considered to be the most hazardous because they can be inhaled and deposited deep in the lung, reaching terminal bronchioles and alveoli.

It is only in recent years that the impacts of air pollution on the brain have been investigated and several potential mechanisms have been identified, including oxidative stress/inflammation (i.e., elevation of cytokines and reactive oxygen species), altered levels of dopamine and/or glutamate, and changes in synaptic plasticity/structure (Allen, Oberdorster et al., 2017). Studies of children and young adults growing up in Mexico City have reported the emergence of exposure-related signs of neurodegeneration, including early stages of the development of neurofibrillary tangles (hyperphosphorylated tau protein) and neuritic plaques (beta-amyloid deposits), with one in four individuals showing later stages (Braak stages III-V) neurofibrillary tangles by the fourth decade of life (Calderón-Garcidueñas et al., 2018). They also show other abnormalities, including prefrontal white matter hyperintensities, damage to epithelial and endothelial barriers, and tight junction and neural autoantibodies (Calderón-Garcidueñas et al., 2016). Studies of a cohort in Spain showed that, even in the absence of morphological changes in brain, greater airborne exposure to elemental carbon and nitrogen dioxide was associated with lower functional connectivity in key brain networks (e.g., the default mode network) as well as altered activation pattern on a sensory task (Pujol et al., 2016). A prospective study conducted in New York City found that greater prenatal exposure to polycyclic aromatic hydrocarbon air pollutants was associated with slower processing speed, ADHD symptoms, and externalizing problems at age seven to nine years (Peterson et al., 2015). Morphometric neuroimaging indicated that these effects were mediated by disruptions of white matter in the left hemisphere. Greater postnatal exposure to polycyclic aromatic hydrocarbons was associated with disruptions of white matter in the dorsal prefrontal regions.

Synthetic Organic Chemicals

The research literatures on the impacts of exposure to synthetic organic chemicals are not as well developed as those on the pollutants discussed in the previous sections. These chemicals include polyhalogenated compounds such as polychlorinated biphenyls (PCBs), and polybrominated diphenyl ethers (PBDEs). PCBs were banned in the United States in the 1970s but, because of their resistance to degradation, they persist in the environment. PBDEs are used as flame retardants in a wide variety of products. Sharing many of the properties of PCBs, they accumulate and persist for long periods in the environment and in human fat tissue. A systematic review and meta-analysis of studies on children’s intelligence and prenatal exposure to PBDEs at levels typical of the general population found a consistent inverse relationship (Lam et al., 2017). A tenfold increase in PBDE exposure was associated with a decrement of nearly four IQ points.

Perfluorinated compounds (PFCs) are commonly used in a variety of consumer products (e.g., nonstick cookware, stain-resistant fabrics, fast food packaging). To date, the evidence pertaining to the neurodevelopmental risks associated with such exposures is mixed (Liew, Goudarzi, & Oulhote, 2018).

Endocrine Disruptors

Concerns have been raised about developmental exposures to chemicals that alter the function(s) of the hormonal system, causing adverse effects in an organism or its progeny (Braun, 2017). Such chemicals are called “endocrine-disrupting chemicals” and can mimic the effects of endogenous hormones, antagonize the effects of endogenous hormones, disrupt the synthesis and metabolism of endogenous hormones, disrupt the synthesis of hormone receptors, and alter target cell sensitivity. Hormone levels in early development are critical in organizing brain development and perturbations can have long-lasting effects on hormonal programming. For example, adequate levels of thyroid hormone are critical for various processes of brain development, including cell migration, differentiation, and signaling. Given that intellectual disability is a result of congenital hypothyroidism, and even subclinical reductions in thyroid function during pregnancy are associated with

IQ deficits in children (Levie et al., 2018), it is plausible to hypothesize that prenatal exposures to chemicals that affect thyroid hormone levels produce more modest impacts on children's intelligence. Increased concentrations of chemicals such as phthalates are inversely associated with total serum thyroid hormone levels in pregnant women and newborns and thyroid-stimulating hormone in newborns (Romano et al., 2018). However, evidence that the alterations that occur at levels of phthalate exposure typical in the general population affect intelligence is presently inconsistent (Factor-Litvak et al., 2014; Nakiwala et al., 2018). Several studies have reported that early exposure to endocrine-disrupting chemicals such as phthalates do, however, influence sexually dimorphic behaviors, that is, those that tend to differ between the sexes. For example, prenatal phthalate exposure might reduce masculine play in boys (Swan et al., 2010).

Estimating the Population Impact of Environmental Chemicals and the Burden of Disease

An argument frequently advanced by those skeptical about the importance of environmental chemicals is that their impact on the neurodevelopment of an individual child is modest, failing to reach the level of clinical significance. This argument fails to consider the issue in the context of population health. Effect estimates from epidemiologic data are in essence population average effects and should be interpreted in the context of a population and not at the individual level. Some individuals will be resistant and some will be more sensitive. Moreover, the impact of a factor at the population level depends not only on the effect size but also on the distribution of the factor or, in the case of a dichotomous factor, its incidence or prevalence. In a set of comparative analyses of pediatric diseases and events, such as brain tumors, congenital heart disease, traumatic brain injury, iron deficiency, and lead exposure, Bellinger (2012) estimated the total number of IQ points lost among US children younger than five years of age associated with each disease or event. The estimate for the loss associated with lead exposure was nearly 23 million IQ points, exceeded only by preterm birth. Among the reasons for this is the absence of a threshold for its inverse relationship with IQ and the fact that virtually every child has a blood lead concentration above the detection limit. As a result, and in contrast to most other diseases and events, every child contributes to the total IQ loss in the population that is associated with lead exposure. In fact, the greatest contribution to the total loss care is contributed by the very large proportion of children with blood lead concentrations at the low end of the distribution (because that is where most children fall). A similar calculation of the total IQ losses among the cohort of young US children from the late 1970s produced the figure of approximately 125 million points, suggesting that the public health interventions implemented to reduce population lead exposure have produced a benefit of about 100 million IQ points. Given that approximately 25 million children fall into this age range, the average IQ benefit has been about 4 points, close to the estimate of 4–5 points reached by Kaufman and colleagues (2014) for the gain in adult IQ.

As the research reviewed in this chapter indicates, environmental chemicals cause adversities that extend well beyond a reduction in intelligence, affecting an exposed individual's success in many aspects of future life. Although current efforts to estimate the burden of disease associated with environmental chemicals consider only IQ deficit as the sequelae (GBD 2016 Risk Factors Collaborators, 2017), a full accounting of the burden of disease imposed by environmental chemicals must include these downstream impacts that can seriously impair quality of life (Bellinger, 2018).

The current approach in regulating chemical exposures is to impose restrictions on their uses only after evidence of harm to people who are exposed becomes apparent. In effect, this results in large-scale natural experiments on the population. Because children are the most vulnerable subgroup of the population, this practice is unconscionable. A more protective alternative would be to proactively require industries to provide evidence of a chemical's safety before allowing its introduction into the marketplace. It is our responsibility to future generations.

References

Allen, J. L., Oberdorster, G., Morris-Schaffer, K., Wong, C., Klocke, C., Sobolewski, M., et al. (2017). Developmental neurotoxicity of inhaled ambient ultrafine particle air pollution: Parallels with neuropathological and behavioral features of autism and other neurodevelopmental disorders. *Neurotoxicology*, *59*, 140–154.

Amato, M. S., Magzamen, S., Imm, P., Havlena, J. A., Anderson, H. A., Kanarek, M. S., & Moore. C. F. (2013). Early lead exposure (<3 years old) prospectively predicts fourth grade school suspension in Milwaukee, Wisconsin (USA). *Environmental Research*, *126*, 60–65.

Amato, M. S., Moore, C. F., Magzamen, S., Imm, P., Havlena, J. A., Anderson, H. A., & Kanarek, M. S. (2012). Lead exposure and educational proficiency: Moderate lead exposure and educational proficiency on end-of-grade examinations. *Annals of Epidemiology*, *22*(10), 738–743.

American Lung Association. (2019). Key findings: The State of the Air 2019. www.lung.org/our-initiatives/healthy-air/sota/key-findings/

Aylward, L. L., Hays, S. M., Kirman, C. R., Marchitti, S. A., Kenneke, J. F., English, C., et al. (2014). Relationships of chemical concentrations in maternal and cord blood: A review of available data. *Journal of Toxicology and Environmental Health B Critical Reviews*, *17*(3), 175–203.

Bellanger, M., Pichery, C., Aerts, D., Berglund, M., Castano, A., Cejchanova, M., et al. (2013). Economic benefits of methylmercury control in Europe: Monetary value of neurotoxicity prevention. *Environmental Health*, *12*, 3.

Bellinger, D. C. (2012). A strategy for comparing the contributions of environmental chemicals and other risk factors to neurodevelopment of children. *Environmental Health Perspectives*, *120*(4), 501–507.

Bellinger, D. C. (2018). Applying methods of the Global Burden of Diseases, Injuries, and Risk Factors Study to developmental neurotoxicants: A commentary. *Environmental Health*, *17*, 53.

Bellinger, D. C., Matthews-Bellinger, J. A., & Korkas, K. (2016). A developmental perspective on early-life exposure to neurotoxicants. *Environment International*, *94*, 103–112.

Bouchard, M. F., Chevrier, J., Harley, K. G., Kogut, K., Vedar, M., Calderon, N., et al. (2011). Prenatal exposure to organophosphate pesticides and IQ in 7-year-old children. *Environmental Health Perspectives*, *119*(8), 1189–1195.

Boutwell, B. B., Nelson, E. J., Qian, Z., Vaughn, M. G., Wright, J. P., Beaver, K. M., et al. (2017). Aggregate-level lead exposure, gun violence, homicide, and rape. *PLoS One*, *12*(11), e0187953.

Braun, J. (2017). Early-life exposure to EDCs: Role in childhood obesity and neurodevelopment. *Nature Reviews Endocrinology*, *13*(3), 161–173.

Braun, J. M., Kahn, R. S., Froehlich, T., Auinger, P., & Lanphear, B. P. (2006). Exposures to environmental toxicants and attention deficit hyperactivity disorder in U.S. children. *Environmental Health Perspectives*, *114*(12), 1904–1909.

Brubaker, C. J., Schmithorst, V. J., Haynes, E. N., Dietrich, K. N., Egelhoff, J. C., Lindquist, D. et al. (2009). Altered myelination and axonal integrity in adults with childhood lead exposure: A diffusion tensor imaging study. *Neurotoxicology*, *30*(6), 867–875.

Byers, R. K., & Lord, E. E. (1943). Late effects of lead poisoning on mental development. *American Journal of Diseases of Children*, *66*(5), 471–494.

Calderón-Garcidueñas, L., Gónzalez-Maciel, A., Reynoso-Robles, R., Delgado-Chávez, R., Mukherjee, P. S., Kulesza, R. J., et al. (2018). Hallmarks of Alzheimer disease are evolving relentlessly in Metropolitan Mexico City infants, children and young adults. APOE4 carriers have higher suicide risk and higher odds of reaching NFT stage V at ≤ 40 years of age. *Environmental Research*, *164*, 475–487.

Calderón-Garcidueñas, L., Reynoso-Robles, R., Vargas-Martínez, J., Gómez-Maqueo-Chew, A., Pérez-Guillé, B., Mukherjee, P. S., et al. (2016). Prefrontal white matter pathology in air pollution exposed Mexico City young urbanites and their potential impact on neurovascular unit dysfunction and the development of Alzheimer's disease. *Environmental Research*, *146*, 404–417.

Cecil, K. M., Brubaker, C. J., Adler, C. M., Dietrich, K. N., Altaye, M., Egelhoff, J. C., et al. (2008). Decreased brain volume in adults with childhood lead exposure. *PLoS Medicine*, *5*(5), e112.

Cecil, K. M., Dietrich, K. N., Altaye, M., Egelhoff, J. C., Lindquist, D. M., Brubaker, C. J., & Lanphear, B. P. (2011). Proton magnetic resonance spectroscopy in adults with childhood lead exposure. *Environmental Health Perspectives*, *119*(3), 403–408.

Choi, A. L., Cordier, S., Weihe, P., & Grandjean, P. (2008). Negative confounding in the evaluation of toxicity: The case of methylmercury in fish and seafood. *Critical Reviews in Toxicology*, *38*(10), 877–893.

Choi, B. H. (1989). The effects of methylmercury on the developing brain. *Progress in Neurobiology*, *32*(6), 447–470.

Choi, W. J., Kwon, H. J., Lim, M. H., Lim, J. A., & Ha, M. (2016). Blood lead, parental marital status and the risk of attention-deficit/hyperactivity disorder in elementary school children: A longitudinal study. *Psychiatry Research*, *236*, 42–46.

Clifford, A., Lang, L., Chen, R., Anstey, K. J., & Seaton, A. (2016). Exposure to air pollution and cognitive functioning across the life course: A systematic literature review. *Environmental Research*, *147*, 383–398.

David, O., Clark, J., & Voeller, K. (1972). Lead and hyperactivity. *Lancet*, *2*(7783), 900–903.

Davidson, P. W., Myers, G. J., Cox, C., Axtell, C., Shamlaye, C., Sloane-Reeves, J., et al. (1998). Effects of prenatal and postnatal methylmercury exposure from fish consumption on neurodevelopment: Outcomes at 66 months of age in the Seychelles Child Development Study. *Journal of the American Medical Association*, *280*(8), 701–707.

Delgado, C. F., Ullery, M. A., Jordan, M., Duclos, C., Rajagopalan, S., & Scott, K. (2017). Lead exposure and developmental disabilities in preschool-aged children. *Journal of Public Health Management Practice*, *24*(2), e10–e17.

Eskenazi, B., An, S., Rauch, S. A., Coker, E. S., Maphula, A., Obida, M., et al. (2018). Prenatal exposure to DDT and pyrethroids for malaria control and child neurodevelopment: The VHEMBE cohort, South Africa. *Environmental Health Perspectives*, *126*(4), 047004.

Evens, A., Hryhorczuk, D., Lanphear, B. P., Rankin, K. M., Lewis, D. A., Forst, L., & Rosenberg, D. (2015). The impact of low-level lead toxicity on school performance among children in the Chicago public schools: A population-based retrospective cohort study. *Environmental Health*, *14*, 21.

Factor-Litvak, P., Insel, B., Calafat, A. M., Liu, X., Perera, F., Rauh, V. A., & Whyatt, R. M. (2014). Persistent associations between maternal prenatal exposure to phthalates on child IQ at age 7 years. *PLoS One*, *9*(12), e114003.

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1997). Early dentine lead levels and educational outcomes at 18 years. *Journal of Child Psychology and Psychiatry*, *38*, 471–478.

Froehlich, T. E., Lanphear, B. P., Auinger, P., Hornung, R., Epstein, J. N., Braun, J., & Kahn, R. S. (2009). Association of tobacco and lead exposures with attention-deficit/hyperactivity disorder. *Pediatrics*, *124*(6), e1054–e1063.

Gąssowska, M., Baranowska-Bosiacka, I., Moczydłowska, J., Tarnowski, M., Pilutin, A., Gutowska, I., et al. (2016). Perinatalexposure to lead (Pb) promotes Tau phosphorylation in the rat brain in a GSK-3β and CDK5 dependent manner: Relevance to neurological disorders. *Toxicology*, *347–349*, 17–28.

GBD 2016 Risk Factor Collaborators. (2017). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet*, *390*, 1345–1442.

Gibb, H., & O'Leary, K. G. (2014). Mercury exposure and health impacts among individuals in the artisanal and small-scale goldmining community: A comprehensive review. *Environmental Health Perspectives*, *122*(7), 667–672.

Goodlad, J. K., Marcus, D. K., & Fulton, J. J. (2013). Lead and attention-deficit/hyperactivity disorder (ADHD) symptoms: A meta-analysis. *Clinical Psychology Review*, *33*(3), 417–425.

Grandjean, P., Weihe, P., White, R. F., Debes, F., Araki, S., Yokoyama, K., et al. (1997). Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury. *Neurotoxicology and Teratology*, *19*(6), 417–428.

Harada, M. (1995). Minamata disease: Methylmercury poisoning in Japan caused by environmental pollution. *Critical Reviews in Toxicology*, *25*(1), 1–24.

Health Effects Institute. (2019). State of Global Air 2019 (website). www.stateofglobalair.org/air

Kaufman, A. S., Zhou, X., Reynolds, M. R., Kaufman, N. L., Green, G. P., & Weiss, L. G. (2014). The possible societal impact of the decrease in U.S. blood lead levels on adult IQ. *Environmental Research*, *132*, 413–420.

Kordas, K., Canfield, R. L., López, P., Rosado, J. L., Vargas, G. G., Cebrián, M. E., et al. (2006). Deficits in cognitive function and achievement in Mexican first-graders with low blood lead concentrations. *Environmental Research*, *100*(3), 371–386.

Lam, J., Lanphear, B. P., Bellinger, D., Axelrad, D. A., McPartland, J., Sutton, P., et al. (2017). Developmental PBDE exposure and IQ/ADHD in childhood: A systematic review and meta-analysis. *Environmental Health Perspectives*, *125*(8), 086001.

Lanphear, B. P., Hornung, R., Khoury, J., Yolton, K., Baghurst P., Bellinger, D. C., et al. (2005). Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. *Environmental Health Perspectives*, *113*(7), 894–899.

Levie, D., Korevaar, T. I. M., Bath, S. C., Dalmau-Bueno, A., Murcia, M., Espada, M., et al. (2018). Thyroid function in early pregnancy, child IQ, and autistic traits: A meta-analysis of individual-participant data. *Journal of Clinical Endocrinology and Metabolism*, *103*(8), 2967–2979.

Liew, Z., Goudarzi, H., & Oulhote, Y. (2018). Developmental exposures to perfluoroalkyl substances (PFASs): An update of associated health outcomes. *Current Environmental Health Reports*, *5*(1), 1–19.

Magzamen, S., Amato, M. S., Imm, P., Havlena, J. A., Coons, M. J., Anderson, H. A., et al. (2015). Quantile regression in environmental health: Early life lead exposure and end-of-grade exams. *Environmental Research*, *137*, 108–119.

Magzamen, S., Imm, P., Amato, M. S., Havlena, J. A., Anderson, H. A., Moore, C. F., & Kanarek, M. S. (2013). Moderate lead exposure and elementary school end-of-grade examination performance. *Annals of Epidemiology*, *23*(11), 700–707.

Miranda, M. L., Kim, D., Reiter, J., Galeano, M. A. O., & Maxson, P. (2009). Environmental contributors to the achievement gap. *Neurotoxicology*, *30*(6), 1019–1024.

Nakiwala, D., Peyre, H., Heude, B., Bernard, J. Y., Béranger, R., Slama, R., & Philippat, C. (2018). In-utero exposure to phenols and phthalates and the intelligence quotient of child study group: Boys at 5 years. *Environmental Health*, *17*(1), 17.

Needleman, H. L., Gunnoe, C., Leviton, A., Reed, R., Peresie, H., Maher, C., & Barrett, P. (1979). Deficits in psychologic and classroom performance of children with elevated dentine lead levels. *New England Journal of Medicine*, *300*(13), 689–695.

Needleman, H. L., Schell, A., Bellinger, D., Leviton, A., & Allred, E. N. (1990). The long-term effects of exposure to low doses of lead in childhood: An 11-year follow-up report. *New England Journal of Medicine*, *322*(2), 83–88.

Nevin, R. (2009). Trends in preschool lead exposure, mental retardation, and scholastic achievement: Association or causation? *Environmental Research*, *109*(3), 301–310.

Nigg, J. T., Knottnerus, G. M., Martel, M. M., Nikolas, M., Cavanagh, K., Karmaus, W., & Rappley, M. D. (2008). Low blood lead levels associated with clinically diagnosed attention-deficit/hyperactivity disorder and mediated by weak cognitive control. *Biological Psychiatry*, *63*, 325–331.

Nigg, J. T., Nikolas, M., Knottnerus, G. M., Cavanagh, K., & Friderici, K. (2010). Confirmation and extension of association of blood lead with attention-deficit/hyperactivity disorder (ADHD) and ADHD symptom domains at population-typical exposure levels. *Journal of Child Psychology and Psychiatry*, *51*, 58–65.

Oken, E., Wright, R. O., Kleinman, K. P., Bellinger, D., Amarasiriwardena, C. J., Hu, H., et al. (2005). Maternal fish consumption, hair mercury, and infant cognition in a U.S. Cohort. *Environmental Health Perspectives*, *113*(10), 1376–1380.

Park, J. H., Seo, J. H., Hong, Y. S., Kim, Y. M., Kang, J. W., Yoo, J. H., et al. (2016). Blood lead concentrations and attention deficit hyperactivity disorder in Korean children: A hospital-based case control study. *BMC Pediatrics*, *16*(1), 156.

Peterson, B. S., Rauh, V. A., Bansal, R., Hao, X., Toth, Z., Nati, G., et al. (2015). Effects of prenatal exposure to air pollutants (polycyclic aromatic hydrocarbons) on the development of brain white matter, cognition, and behavior in later childhood. *JAMA Psychiatry*, *72*(6), 531–540.

Pujol, J., Martínez-Vilavella, G., Macià, D., Fenoll, R., Alvarez-Pedrerol, M., Rivas, I., et al. (2016). Traffic pollution exposure is associated with altered brain connectivity in school children. *Neuroimage*, *129*, 175–184.

Rauh, V. A., Perera, F. P., Horton, M. K., Whyatt, R. M., Bansal, R., Hao, X., et al. (2012). Brain anomalies in children exposed prenatally to a common organophosphate pesticide. *Proceedings of the National Academy of Sciences*, *109*(20), 7871–7876.

Rauh, V., Arunajadai, S., Horton, M., Perera, F., Hoepner, L., Barr, D. B., & Whyatt, R. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. *Environmental Health Perspectives*, *119*(8), 1196–1201.

Reuben, A., Caspi, A., Belsky, D. W., Broadbent, J., Harrington, H., Sugden, K., et al. (2017). Association of childhood blood lead levels with cognitive function and socioeconomic status at age 38 years and with IQ change and socioeconomic mobility between childhood and adulthood. *Journal of the American Medical Association*, *317*(12), 1244–1251.

Romano, M. E., Eliot, M. N., Zoeller, R. T., Hoofnagle, A. N., Calafat, A. M., Karagas, M. R., et al. (2018). Maternal urinary phthalate metabolites during pregnancy and thyroid hormone concentrations in maternal and cord sera: The HOME Study. *International Journal of Hygiene and Environmental Health*, *221*(4), 623–631.

Schneider, J. S., & DeKamp, E. (2007). Postnatal lead poisoning impairs behavioral recovery following brain damage. *Neurotoxicology*, *28*(6), 1153–1157.

Selevan, S. G., Kimmel, C. A., & Mendola, P. (2000). Identifying critical windows of exposure for children's health. *Environmental Health Perspectives*, *108*(Suppl 3), 451–455.

Seo, J., Lee, B. K., Jin, S. U., Jang, K. E., Park, J. W., Kim, Y. T., et al. (2015). Altered executive function in the lead-exposed brain: A functional magnetic resonance imaging study. *Neurotoxicology*, *50*, 1–9.

Skerfving, S., Lofmark, L., Lundh, T., Mikoczy, Z., & Stromberg, U. (2015). Late effects of low blood lead concentrations in children on school performance and cognitive functions. *Neurotoxicology*, *49*, 114–120.

Surkan, P. J., Zhang, A., Trachtenberg, F., Daniel, D. B., McKinlay, S., & Bellinger, D. C. (2007). Neuropsychological function in children with blood lead levels <10 microg/dL. *Neurotoxicology*, *28*(6), 1170–1177.

Swan, S. H., Liu, F., Hines, M., Kruse, R. L., Wang, C., Redmon, J. B., et al. (2010). Prenatal phthalate exposure and reduced masculine play in boys. *International Journal of Andrology*, *33*(2), 259–269.

Téllez-Rojo, M. M., Bellinger, D. C., Arroyo-Quiroz, C., Lamadrid-Figueroa, H., Mercado-García, A., Schnaas-Arrieta, L., et al. (2006). Longitudinal associations between blood lead concentrations lower than 10 microg/dL and neurobehavioral development in environmentally exposed children in Mexico City. *Pediatrics*, *118*(2), e323–e330.

Wang, H. L., Chen, X. T., Yang, B., Ma, F. L., Wang, S., Tang, M. L., et al. (2008). Case-control study of blood lead levels and attention deficit hyperactivity disorder in Chinese children. *Environmental Health Perspectives*, *116*(10), 1401–1406.

Weiss, B., Clarkson, T. W., & Simon, W. (2002). Silent latency periods in methylmercury poisoning and in neurodegenerative disease. *Environmental Health Perspectives*, *110*(Suppl 5), 851–854.

White, R. F, Palumbo, C. L., Yurgelun-Todd, D. A., Heaton, K. J., Weihe, P., Debes, F., & Grandjean. P. (2011). Functional MRI approach to developmental methylmercury and polychlorinated biphenyl neurotoxicity. *Neurotoxicology*, *32*(6), 975–980.

WHO (World Health Organization). (2009). Ten chemicals of major public health concern. www.who.int/ipcs/assessment/public_health/chemicals_phc/en/

Woodruff, T. J., Zota, A. R., & Schwartz, J. M. (2011). Environmental chemicals in pregnant women in the United States. *Environmental Health Perspectives*, *119*(6), 878–885.

Wright, J. P., Dietrich, K. N., Ris, M. D., Hornung, R. W., Wessel, S. D., Lanphear, B. P., Ho, M., & Rae, M. N. (2008). Association of prenatal and childhood blood lead concentrations with criminal arrests in early adulthood. *PLoS Medicine*, *5*(5), e101.

Yuan, W., Holland, S. K., Cecil, K. M., Dietrich, K. N., Wessel, S. D., Altaye, M., et al. (2006). The impact of early childhood lead exposure on brain organization: a functional magnetic resonance imaging study of language function. *Pediatrics*, *118*(3), 971–977.

Zheng, W., Aschner, M., & Ghersi-Egea, J. F. (2003). Brain barrier systems: A new frontier in metal neurotoxicological research. *Toxicology and Applied Pharmacology*, *192*(1), 1–11.

PART VIII

Intelligence and Allied Constructs

42 Intelligence and Personality

Colin G. DeYoung

Intelligence and personality have often been viewed as distinct psychological domains that intersect only to a very limited degree. However, research over the last four decades suggests the possibility that, both conceptually and empirically, intelligence could be integrated with larger models of personality. Such an integration may allow a more unified conception of the structure and sources of individual differences. Since the previous edition of this handbook, the prospect for integrating intelligence with personality has been strengthened by new research clarifying the relation of intelligence to broad taxonomies of personality. The first purpose of this chapter is to explore the conceptual relation of intelligence to personality. The second is to review empirical research on the relation of intelligence to a wide range of personality traits.

Following a presentation of working definitions for intelligence and personality, the chapter reviews arguments for and against three of the most common distinctions that are drawn between intelligence and personality. These three dichotomies provide an overview of the major conceptual issues at stake. Given the amount of thought that has been devoted to the conceptual relation of intelligence to personality, this chapter cannot hope to be comprehensive. Additional perspectives can be found in three excellent edited collections (Collis & Messick, 2001; Saklofske & Zeidner, 1995; Sternberg & Ruzgis, 1994). Additionally, the chapter discusses whether intelligence can be located within the Big Five model (John, Naumann, & Soto, 2008). Finally, the Big Five personality dimensions serve to organize a review of empirical associations of intelligence with various personality traits, with a separate section at the end for associations with sociopolitical orientation.

Definition of Intelligence

In this chapter, "intelligence," without a modifier, is used to refer to general intelligence, often known as the *g*-factor, the ability to perform well on a wide variety of challenging cognitive tasks (Spearman, 1904). Intelligence contributes to solving problems (including problems of comprehension) through thinking and reasoning, and it is well measured by IQ and similar tests (Gottfredson, 1997a; Neisser et al., 1996). General intelligence occupies the apex of a hierarchy of more specific cognitive abilities that are all related to each other. At levels of the hierarchy below *g*, divisions can be made among narrower forms of intelligence, which are empirically distinct despite being correlated with each other (Carroll, 1993; Johnson & Bouchard, 2005a, 2005b). In this chapter, it will be necessary to refer to a distinction between the two major types of ability at the level of the intelligence hierarchy immediately below *g*. These have often been referred to as *crystallized* and *fluid* intelligence, but I will refer to them as *verbal* and *nonverbal* intelligence, for several reasons.

Research indicates that the difference between these two factors is not best captured by the terms "fluid" and "crystallized," which were originally developed based on the theory that some abilities (those called "fluid") were genetically determined and uninfluenced by experience or education, whereas other abilities (those that "crystallized" over time) relied on knowledge or skill acquired from experience (Horn & Cattell, 1966). Factor analyses of the most extensive test batteries available show that individual differences in ability do not covary according to how much they depend on acquired knowledge but rather according to whether they require solving problems using stimuli that are verbal or nonverbal (Johnson & Bouchard, 2005a, 2005b; Johnson, Nijenhuis, & Bouchard, 2007; Major, Johnson, & Deary, 2012). Johnson and Bouchard (2005a) labeled the nonverbal factor "perceptual," but nonverbal memory and reasoning tasks were also encompassed by this factor, and "nonverbal" seems a more adequately inclusive label. (Their model also identifies a third factor at the same level of the hierarchy, representing the ability to rotate three-dimensional images mentally, but this is a much smaller factor than the other two.)

Both verbal and nonverbal intelligence are determined by a combination of innate ability and acquired knowledge and skills. Verbal intelligence cannot be entirely crystallized (dependent on experience) because it is just as heritable (that is, its variation among people is genetically influenced) as nonverbal intelligence, even after controlling for *g* (Johnson & Bouchard, 2007; Johnson et al., 2007). Even vocabulary tests (often held up as prototypically "crystallized") usually require people to reason fluidly about the meaning and relations of words and concepts in order to articulate appropriate definitions on the fly, given that most people do not memorize dictionaries. Complementarily, nonverbal intelligence is obviously not entirely fluid (independent of experience) because it is influenced by environmental factors in studies of heritability (Johnson & Bouchard, 2007; Johnson et al., 2007), and it has increased in the populations of industrialized nations on a timescale too short to be due to genetic change (a phenomenon known as the Flynn effect, after its discoverer). In fact, the Flynn effect for nonverbal ("fluid") intelligence is considerably greater than that for verbal intelligence (Pietschnig & Voracek, 2015).

Most tests traditionally considered to measure crystallized intelligence are verbal, whereas most tests traditionally considered to measure fluid intelligence are nonverbal. Thus, most past findings regarding fluid and crystallized intelligence and personality can be translated effectively into a verbal-nonverbal framework for the present review.

Definition of Personality

Personality is a broader concept than intelligence, as can be seen in the following definition by McAdams and Pals (2006):

> Personality is an individual's unique variation on the general evolutionary design for human nature, expressed as a developing pattern of dispositional traits, characteristic adaptations, and integrative life stories, complexly and differentially situated in culture. (p. 212)

This definition highlights three distinct levels at which personality can be described: characteristic adaptations, life stories, and traits. Characteristic adaptations and life stories both describe the individual's adaptation to their particular sociocultural context (e.g., as a lawyer), with characteristic adaptations reflecting different strategies and goals that one has adopted and life stories reflecting one's narrative descriptions of one's history and identity (DeYoung, 2015a). Traits describe relatively stable patterns of behavior, motivation, emotion, and cognition (Pytlik Zillig, Hemenover, & Dienstbier, 2002; Wilt & Revelle, 2015) that are not bound to a particular sociocultural context but could be observed in any such context (e.g., argumentativeness). This is not to say that all traits will have the same average scores, or identical manifestations, in all cultures, nor that all traits can be observed in any situation, but rather that any trait can be observed in a range of situations in any culture (DeYoung, 2015a). Traits will be the primary level of focus in this chapter. For this reason, vocational interests will not be discussed, despite their relevance to intelligence and related personality traits (Ackerman & Heggestad, 1997), as they are more like characteristic adaptations than traits, in their cultural specificity.

A central project in personality psychology has been the development of a comprehensive taxonomy of traits. Research based both on trait descriptors drawn from the natural language (as represented in dictionaries) and on large collections of existing personality questionnaires has provided evidence for a five factor solution, leading to a taxonomy known as the Five Factor Model or Big Five (Goldberg, 1990; John et al., 2008; Markon, Krueger, & Watson, 2005; Waller, DeYoung, & Bouchard, 2016). This model includes the broad trait domains of Extraversion, Neuroticism, Agreeableness, Conscientiousness, and Openness/Intellect, each of which will be defined in its own section below. The Big Five are substantially genetically influenced (Rieman, Angleitner, & Strelau, 1997), and the genetic factor structure of the Big Five appears to be invariant across European, North American, and East Asian samples, suggesting the biological universality of this model (Yamagata et al., 2006).

Personality traits are hierarchically organized, with more specific traits (e.g., talkativeness, sociability, enthusiasm) varying together, such that one can deduce the presence of broader traits (e.g., Extraversion, for the three traits just mentioned) that account for their covariance. Higher-order traits may exist above the Big Five (DeYoung, 2006; Digman, 1997), but they appear to be minimally related to intelligence (DeYoung et al., 2008). For the present purpose, therefore, they are of less interest than levels of trait structure below the Big Five. Each Big Five domain comprises a large number of lower-level traits, called *facets*, with no consensus as to how many facets exist for each domain. Additionally, research suggests the existence of a level of personality structure between the Big Five and their facets. In two samples, two genetic factors were necessary to account for the shared genetic variance among the facets within each of the Big Five (Jang et al., 2002). If the Big Five were the next level above the facets, only one genetic factor should have been necessary for each domain.

In factor analysis of phenotypic data, using fifteen facets for each domain, two factors similar to the genetic factors were found for each of the Big Five (DeYoung, Quilty, & Peterson, 2007). These factors were then characterized empirically by their correlations with more than 2,000 items from the International Personality Item Pool (Goldberg, 1999). Of particular relevance for intelligence, the two factors in the Openness/Intellect domain clearly differentiated between Openness to Experience and Intellect, with Openness reflecting aesthetically oriented traits related to engagement in sensation and perception (e.g., “Believe in the importance of art”; “See beauty in things that others might not notice”) and Intellect reflecting intellectual interest or engagement (e.g., “Avoid philosophical discussions”–reversed) and perceived intelligence (e.g., “Am quick to understand things”).

Importantly, traits are probabilistic entities. Each of the Big Five encompasses many subtraits, and a high score on a Big Five trait indicates an increased likelihood of high scores on its various subtraits but is not deterministic. This entails that people scoring high in Intellect will, on average, score higher in Openness than people scoring low in Intellect. However, the correlation between Openness and Intellect is far from perfect, which means that some people will score high in Intellect but only moderate or low in Openness and vice versa. Distinguishing these two aspects of the broader Openness/Intellect domain from the Big Five turns out to be crucial for understanding the empirical relation of intelligence to personality, as explained following a discussion of their conceptual relation.

The Conceptual Relation of Intelligence to Personality

Given a broad definition of personality, like the one presented in the section “Definition of Personality,” the possibility of describing intelligence as a personality trait seems clear. Indeed, some early theorists considered personality to include intelligence (Cattell, 1950; Guilford, 1959). However, most theorists have not considered intelligence to be part of personality, instead asserting either that intelligence (as defined in the section “Definition of Intelligence”) is unrelated to personality (e.g., Eysenck, 1994) or that intelligence and personality are related but nonetheless categorically distinct (e.g., Chamorro-Premuzic & Furnham, 2005). The large body of empirical evidence reviewed in the latter half of this chapter rules out the possibility that intelligence is unrelated to personality. A number of personality traits show consistent and meaningful relations to intelligence. Thus, the important contrast is between the view that intelligence is a personality trait and the more common view that intelligence is fundamentally different from personality traits.

Three dichotomies seem to be largely responsible for the view that intelligence and personality may be related but should be considered categorically distinct. (Because many researchers have advanced similar dichotomies, with slight variations, what follows represents a distillation of many viewpoints.) First, a distinction is often made between cognitive and noncognitive traits, with intelligence considered to be cognitive and personality considered to be noncognitive. Second, intelligence and personality differ in their typical methods of measurement: Intelligence is usually assessed using ability tests, whereas personality is usually assessed by questionnaire. Third, the difference in typical measurement corresponds to a conceptual distinction in which intelligence is often considered to reflect “maximal performance” (i.e., performance when individuals are trying their hardest), whereas personality is considered to reflect “typical behavior” (Cronbach, 1949). The following section reviews arguments for and against the validity of these dichotomies.

The cognitive/noncognitive dichotomy is widely used, but the evidence against it is strong enough that even some of the people who popularized it have acknowledged that it is a “misnomer” (Duckworth, 2009, p. 279). The distinction between cognitive and noncognitive fails because almost all traits have cognitive attributes (even when “cognitive” is used to designate something like “conscious thought” rather than referring to any form of information processing), though these are more prominent in some traits than in others. In a study of common Big Five questionnaires, items describing cognitive traits were found in all five domains, with Openness/Intellect containing the most such items and Extraversion and Neuroticism containing the fewest (Pytlik Zillig et al., 2002). Examples of cognitive attributes are easily provided, even for traits that might be considered relatively less cognitive: Neuroticism is associated with rumination, compulsive thinking about possible threats (Nolan, Roberts, & Gotlib, 1998); Agreeableness is associated with “social-cognitive theory of mind,” understanding and reasoning about the mental states of others (Allen et al., 2017; Nettle & Liddle, 2008). Personality includes stable patterns of cognition, in addition to behavior, motivation, and emotion. Duckworth (2009; Duckworth & Yeager, 2105) suggests that psychologists may continue to employ this problematic dichotomy because “cognitive” is a convenient shorthand for “cognitive ability.” “Noncognitive,” therefore, is used as shorthand to indicate all variables other than cognitive ability or intelligence, even though many of those other variables have cognitive attributes. Thus, the existence of the cognitive/noncognitive dichotomy appears to reflect imprecise use of language rather than a strong theoretical assertion that intelligence is categorically distinct from personality.

The second dichotomy involves methods of measurement. Historically, research on intelligence has been separated from research on personality by the fact that personality has typically been assessed by questionnaire, whereas intelligence has typically been assessed by ability tests. These two research traditions thus represent two *paradigms*, in Kuhn’s (1970) original sense, separated from each other by differing sets of conventional scientific practices. Nonetheless, most psychologists would not assert that different methods of measurement, in and of themselves, justify a categorical distinction between the constructs that have been measured. (Whether the differences in measurement are necessary because of an underlying conceptual distinction is a separate question and the focus of the third dichotomy, discussed later in the current section.) Psychometricians warn against confusing constructs with measures (e.g., Jensen, 1998; Loevinger, 1957). Personality traits are not identical to scores on personality questionnaires, just as intelligence is not identical to an IQ score. In both cases, the measures merely provide estimates of what researchers typically want to investigate – namely, latent traits, actual patterns of human functioning that persist over time – and these cannot be measured without error. Multiple methods can be used to measure a single latent trait; each method may incorporate different sources of error or bias, and one method may be better than another for the purposes intended, but nonetheless each can be said to measure the same latent trait. For example, given our working definition of intelligence as a general mental ability, one should expect it to be best measured by ability tests, but one could also measure it, albeit less accurately, using questionnaires that require self-, peer, or observer ratings of subjects' mental ability (this approach is discussed in more detail in the section "Openness/Intellect"). Differences in typical methods of measurement, therefore, would not usually be seen as sufficient to rule out the possibility that intelligence is part of personality.

What makes the issue of measurement more complicated, however, is the possibility that the different types of measures typically used for intelligence and personality correspond to a valid dichotomy between maximal performance and typical behavior. If intelligence really involves only maximal performance, and if personality really involves only typical behavior, then one would be forced to conclude that intelligence and personality are categorically distinct. Our working definition of intelligence can be read to imply that maximal performance is what matters. However, some theorists have questioned the sharpness of the distinction between maximal performance and typical behavior (e.g., Ackerman, 1996). This distinction is blurred by the fact that ability can affect typical behavior, as illustrated by the fact that IQ scores are good predictors of outcomes that depend on typical behavior – including job success, academic performance, and health (Deary, 2012). If being intelligent did not typically entail that one often used one's intelligence, IQ would be unlikely to predict real-world outcomes. Because the complexity of the world always outstrips our simplified mental models (Peterson & Flanders, 2002), intelligence will often be expressed in typical behavior (Gottfredson, 1997b). Even idle thoughts seem likely to be different for those high as opposed to low in intelligence. Any ability for which there is frequent demand or possibility for application will influence typical behavior, and tests of that ability will provide indices of both maximal performance and typical behavior. This is not to say that maximal performance is identical to typical behavior – underachievers who fail to make the best use of their abilities are a clear counterexample – but a case can be made that intelligence, as a trait, entails typical behavior as well as maximal performance, even while acknowledging that ability and typical behavior are not the same thing.

The idea that personality involves only typical behavior has also been contested. The personality research framework provided by the lexical hypothesis has generally not excluded abilities. Traits that describe ability have been included in all selections of personality descriptors from natural languages (though more in some than others; John et al., 2008), and these have not fallen exclusively within the Openness/Intellect domain in factor analysis. For example, empathy is a component of Agreeableness that involves the ability to detect the mental states of others, and many components of Conscientiousness, such as self-discipline and patience, can be considered abilities. Large differences in outcome may be evident when people are trying their hardest to be patient, rather than not attempting to restrain themselves, and some people may be more successful in the attempt than others. Abilities thus appear to be relatively common within the Big Five.

One complement to the observation that numerous personality traits involve abilities is the idea that ability tests could be used to measure traits other than intelligence (Ackerman, 2009; Cattell & Birkett, 1980; Cattell & Warburton, 1967; Wallace, 1966; Willerman, Turner, & Peterson, 1976). For example, tests of the ability to detect and understand others' mental and emotional states might be reasonable measures of at least some facets of Agreeableness (Allen et al., 2017; Nettle & Liddle, 2008), and tests of the ability to remain calm under stress might be good measures of Neuroticism. Personality includes many abilities that could potentially be measured by tests of maximal performance, and better progress may be made in this area if such tests are designed to reflect theories regarding the key underlying processes involved in different personality traits (e.g., DeYoung, 2015a). In creating such tests, one must remember that, because of the differences in method, correlations between questionnaires and tests measuring the same trait are unlikely to be very high, even if the tests are valid.

Having reviewed arguments for and against the three dichotomies commonly used to separate intelligence from personality, one can conclude that viewing intelligence as a personality trait is a viable, if relatively uncommon, conceptual strategy. Many personality traits appear to involve both cognitive processes and abilities, two categories that have sometimes been considered exclusive to intelligence. One might argue that maximal performance (relative to typical behavior) is more important in intelligence than in other traits, but this suggests a difference of degree between intelligence and other traits, rather than a qualitative or categorical difference. The question of whether intelligence should be considered a personality trait thus remains open.

Intelligence in the Big Five

The previous section raised the question of whether intelligence can be considered part of personality. Given the potential viability of an affirmative answer, another important question is whether intelligence can be integrated with models of personality, like the Big Five, that are derived from questionnaire measures and attempt to provide comprehensive taxonomies of traits. Any trait model that would claim comprehensiveness should presumably include intelligence. Based on lexical and questionnaire studies, verbal descriptions of intelligence fall within the Intellect aspect of the Openness/Intellect domain in the Big Five.

The compound label "Openness/Intellect" reflects a history of debate about how best to characterize the content of this domain, with some researchers preferring "Openness to Experience" (e.g., Costa & McCrae, 1992a) and others "Intellect" (e.g., Goldberg, 1990). This debate was largely resolved conceptually by the observation that "Openness" and "Intellect" describe two central aspects of the larger domain (DeYoung, 2015b; DeYoung et al., 2007; Johnson, 1994; Saucier, 1992; Woo, Chernyshenko, Longley et al., 2014). Lexical studies made it clear that both aspects are represented in natural language and appear within a single Big Five factor (e.g., Goldberg, 1990; Saucier, 1992). Many words describe Intellect – *intellectual, intelligent, philosophical, erudite, clever* – and many words describe Openness – *artistic, perceptive, poetic, fantasy-prone*. Additionally, many words could characterize people high in Intellect or Openness or both – *imaginative, original, innovative*. In fact, Saucier (1992, 1994) proposed that "Imagination" might be a better single label for the domain as a whole, given the existence of both intellectual and aesthetic forms of imagination. This broad sense of "imagination" seems appropriate for a trait domain that has, as its central characteristic, the disposition to detect, explore, appreciate, and utilize both abstract and sensory information (DeYoung, 2015a, 2015b). Importantly, general measures of Openness/Intellect (such as the Revised NEO Personality Inventory; NEO PI-R; Costa & McCrae, 1992b; the Trait Descriptive Adjectives; Goldberg, 1992; or the Big Five Inventory; John et al., 2008; Soto & John, 2017) contain content reflecting both Openness and Intellect and they predict other variables very similarly, no matter which label their authors prefer (e.g., DeYoung, Peterson, & Higgins, 2005).

In studies of the Big Five in languages other than English, less agreement about the nature of the factor corresponding to Openness/Intellect has emerged, relative to the other four factors. In a Dutch study, for example, this factor was most strongly characterized by descriptors of unconventionality (Hofstee et al., 1997). (Content related to unconventionality also appears in the English Openness/Intellect factor but less predominantly.) However, these differences between languages appear to be related primarily to criteria for variable selection. In Dutch and Italian lexical studies, for example, descriptors related to abilities were intentionally undersampled, leading to the exclusion of many terms that might reflect intellectual ability (John et al., 2008). Additionally, in a six-factor lexical solution that has been proposed as a slight modification of the Big Five, the content of Openness/Intellect was more consistent across all languages (Ashton et al., 2004). Thus, the relative lack of consensus about the content of Openness/Intellect appears to have been due to methodological issues. The current state of lexical research suggests that Openness/Intellect encompasses a range of trait descriptors related to intellectual and aesthetic curiosity, creativity, imagination, and ability – including descriptors of intelligence.

As measured by questionnaires, therefore, intelligence can be located within the Big Five. Despite this semantic fit, objections have been raised because intelligence tests do not behave quite like questionnaire ratings of descriptors of intelligence. If multiple cognitive ability tests are factor analyzed with personality questionnaires, they tend to form a sixth factor, rather than grouping with questionnaire variables reflecting Openness/Intellect (McCrae & Costa, 1997). However, this result may be due to one or two artifacts, the first of which is the presence of two distinct sources of method variance in these factor analyses. In addition to substantive trait variance, all of the ability tests share method variance that they do not share with any questionnaire variables and vice versa. This shared variance inflates the intercorrelations within each type of measure, relative to their correlations with the other type, and inclines the two types of measure to form separate factors, regardless of what they share substantively.

A second possible artifact resembles what Cattell (1978) called a "bloated specific factor," which could result from the inclusion of many intelligence tests in factor analysis of broad personality questionnaires. A bloated specific factor is one that appears only because measures of a single lower-level trait are overrepresented in the pool of variables to be factor analyzed. Their large number will tend to cause them to form a separate factor, even when the other factors recovered are at a higher level of the trait hierarchy and one of them should subsume the lower-level trait in question. As an analogy, consider what would happen if one included many scales measuring anxiety in a factor analysis with the thirty facets of the Big Five measured by the NEO PI-R. When this is done, one sometimes finds a sixth factor for anxiety, in addition to the usual general Neuroticism factor (Oltmanns & Widiger, 2016). This anxiety factor should be considered a bloated specific factor because the location of anxiety as a lower-level trait within Neuroticism is well established (John et al., 2008; Markon et al., 2005).

The existence of distinct method variance for intelligence tests and questionnaires, plus the possibility of bloated specific factors, makes interpretation ambiguous for results of joint factor analyses of tests and questionnaires. The factor-analytic results summarized by McCrae and Costa (1997) could be taken to indicate that intelligence falls outside of the Big Five (which would imply that descriptors of intelligence do not measure intelligence as much as they measure some other construct), or they can be challenged by the argument that an adequate factor analysis would need to take method variance into account. Unfortunately, there's a catch to the latter argument: Modeling method variance in intelligence is complicated by the fact that the relevant method factor would consist of the variance shared among cognitive ability tests – but this is exactly the definition of *g*. Thus, for intelligence, method variance is thoroughly confounded with substantive variance. One alternative way to test the location of intelligence within the Big Five in factor analysis is to use a single IQ score or other index of *g*, rather than multiple ability tests, because then there are no other tests with which it can share method and hence no separate method factor. When this is done, intelligence loads primarily and substantially (>0.30) on the Openness/Intellect factor, supporting the integration of intelligence as a lower-level trait within the Big Five personality hierarchy (DeYoung, Grazioplene, & Peterson, 2012).

The idea that intelligence could be a lower-level trait in the personality hierarchy might strike some as odd, given the obvious importance of intelligence in human functioning and the number of cognitive abilities that make up the hierarchy below *g*. Nonetheless, the location of descriptors of intelligence within the Big Five seems clear. As noted in the section "Definition of Personality," the existence of Openness and Intellect as two correlated but separable aspects of Openness/Intellect was supported by factor analysis of fifteen facet scales in this domain, and empirical characterization of the Intellect factor by correlations with thousands of personality items indicated that it includes at least two facets, intellectual or cognitive engagement and perceived intelligence or cognitive capacity (DeYoung et al., 2007; Smillie et al., 2016, appendix). In the Big Five personality hierarchy, therefore, intelligence appears to be at a relatively low level: one facet out of at least two within Intellect, which is itself one of two aspects of the broader Openness/Intellect domain (Figure 42.1; DeYoung, 2015b). This structural finding highlights the great complexity of the personality hierarchy, in terms of how many different patterns of emotion, motivation, cognition, and behavior it encompasses. Intelligence is not unique in being an extremely important and multifaceted construct that is, nonetheless, relatively narrow when compared with traits like the Big Five that represent very broad regularities in personality. Anxiety, for example, appears to be one facet of the Withdrawal aspect of Neuroticism (DeYoung et al., 2007, 2016) and thus exists at the same level of the personality hierarchy as intelligence. The relative breadth of a trait places no limitation on its importance to human beings and seems to place little limitation on the extent to which it may be further subdivided.

Having located intelligence within the personality hierarchy conceptually, we can turn in more depth to the question of how it relates empirically to the Big Five and their lower-order traits. Its putative position within Intellect suggests that it should be most strongly related to questionnaire measures of Intellect and to general measures of the Openness/Intellect domain but less strongly to specific measures of Openness and to other Big Five domains. Having asserted that ability tests are better measures of intelligence than questionnaires are, this chapter will continue to focus on these tests and, when "intelligence" is discussed in relation to empirical work, it has been measured by ability tests, unless otherwise noted.

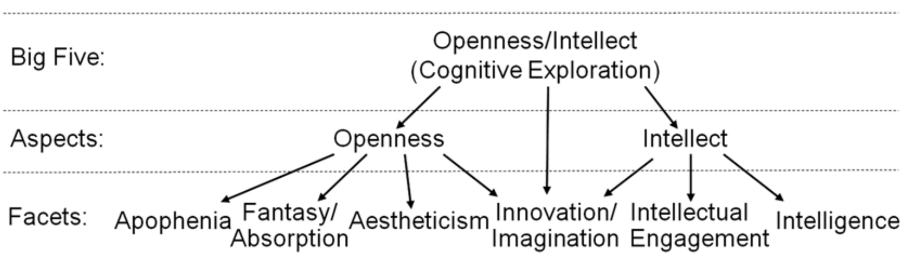

Figure 42.1 *Hierarchical structure of the Openness/Intellect trait domain (from DeYoung, 2015b).*
Levels of the hierarchy are labeled at left. Facets are arranged such that those closest together are most strongly related and those farthest apart are least related (DeYoung et al., 2012). Facet labels represent categories of facets and are not indivisible entities; no consensus exists as to the exact number and identity of facets. Apophenia is the tendency to detect patterns or causal connections where none exist.

Openness/Intellect

Several thorough reviews of associations between intelligence and personality have been published (Ackerman, 2009; Chamorro-Premuzic & Furnham, 2005; Eysenck, 1994; Zeidner & Matthews, 2000), but until recently only one had been meta-analytic (Ackerman & Heggestad, 1997), and this meta-analysis included only three studies reporting the correlation of Openness/Intellect with *g*. In the previous version of this chapter (DeYoung, 2011), I informally meta-analyzed nine additional studies that had been published since 1997 and found a very similar correlation to Ackerman and Heggestad ($r = 0.3$). A more recent dissertation provides an impressively comprehensive meta-analysis of the relation of intelligence to personality, analyzing effects from more than 900 studies, and serves as an important source for my discussion in this chapter (Stanek, 2014). Unfortunately, the dissertation is currently under embargo, and the results are being updated for publication (Stanek, personal communication, June 2018). Therefore, I will not be citing exact numbers from this meta-analysis but will provide approximations and assessments of how well it supports conclusions from already published data.

One additional complication is that this meta-analysis includes many studies using non-Big Five questionnaires, whenever possible categorizing scales from those measures within the Big Five or their aspects and facets and including them in meta-analytic estimates of associations with constructs from the Big Five hierarchy. This has the potential to introduce noise and attenuate correlations if any scale categorizations are inaccurate or merely approximate. Nonetheless, it provides the most extensive analysis to date of intelligence-personality associations and confirms that, of the Big Five, Openness/Intellect shows by far the strongest association with intelligence, with a correlation around 0.25. Although this correlation is moderate in magnitude (Hemphill, 2003), it is consistent with the possibility of including intelligence as a facet of Openness/Intellect, given the lack of shared method. Note that the average correlation between facets of Openness/Intellect in the NEO PI-R, which do share method, is only 0.28 (Costa & McCrae, 1992b).

Both Stanek's (2014) meta-analysis and research using a purpose-built measure of the Intellect and Openness aspects (the Big Five Aspect Scales; BFAS; DeYoung et al., 2007) confirm that Intellect is more strongly related to intelligence than Openness is. Whereas the correlation of intelligence with Intellect is about 0.35, that with Openness is only in the range of about 0.15 to 0.20 (DeYoung et al., 2012, 2014; Kaufman et al., 2016; Stanek, 2014). Further, when Intellect and Openness are used as simultaneous predictors, only Intellect is uniquely associated with intelligence (DeYoung et al., 2012, 2014; Woo, Chernyshenko, Stark, & Conz, 2014). This pattern is consistent with the idea that intelligence can be seen specifically as a facet of the Intellect aspect of Openness/Intellect within the Big Five.

It is also consistent with research based on scales that were not explicitly designed to measure Intellect and Openness but that measure facets of these two traits (DeYoung et al., 2012; Mussel, 2013). Scales measuring components of Intellect can be categorized as measuring either intellectual engagement or perceived intelligence. Commonly used scales measuring intellectual engagement include Typical Intellectual Engagement (TIE; Goff & Ackerman, 1992), Need for Cognition (NFC; Cacioppo et al., 1996), and the Ideas facet of the NEO PI-R (Costa & McCrae, 1992b). The Ideas facet is much more strongly correlated with TIE ($r = 0.77$; Ackerman & Goff, 1994) and NFC ($r = 0.78$; Cacioppo et al., 1996) than with any of the other NEO PI-R facets (Costa & McCrae, 1992b). Like Ideas, TIE and NFC have been found to be substantially associated with intelligence (Ackerman & Heggestad, 1997; Cacioppo et al., 1996; Espejo, Day, & Scott, 2005; Frederick, 2005; Gow et al., 2005; Hill et al., 2013).

Whereas Ideas is the only NEO PI-R facet that is a good marker of Intellect (DeYoung et al., 2007; DeYoung et al., 2012), four NEO PI-R facets are good markers of Openness; listed from largest to smallest loading, they are Aesthetics, Fantasy, Feelings, and Actions.[1] (The sixth Openness/Intellect facet, Values, does not mark either Openness or Intellect strongly and is discussed below in the section "Sociopolitical Orientation.") In studies that consider the NEO PI-R facets individually, Ideas typically predicts intelligence more strongly than do the four Openness facets (DeYoung et al., 2005, 2009, 2012; Furnham et al., 2007; Holland et al., 1995; McCrae, 1993; Moutafi, Furnham, & Crump, 2003, 2006). Further, a behavioral genetic study found that a genetic factor influencing intelligence tests was marked strongly by Ideas but not by the facets that reflect Openness (Wainwright et al., 2008).

Measures of perceived intelligence are less standardized than measures of intellectual engagement, with some involving Likert-ratings of descriptors of intelligence and others involving more direct estimations of intelligence with reference to a normal distribution or percentiles. Nonetheless, a number of studies have examined their association with performance on intelligence tests, and a meta-analysis of forty-one such studies found a correlation of 0.33 (Freund & Kasten, 2012). Again, this effect size is consistent with the location of intelligence within the personality hierarchy depicted in Figure 42.1, but it also clearly indicates that self-reported intelligence should not be used as a proxy for tested intelligence (Freund & Kasten, 2012; Paulhus, Lysy, & Yik, 1998). Other-ratings of intelligence fare somewhat better, though they have been less well studied. Teacher-ratings of intelligence strongly predict student IQ, with reported correlations ranging from about 0.45 all the way up to 0.80 (Alvidrez & Weinstein, 1999; Brickenkamp, 1975, cited in Ostendorf & Angleitner, 1994; Pedulla, Airasian, & Madaus, 1980). Additional research is necessary on how well intelligence can be rated by others who are not teachers, such as friends or family members.

The relative lack of accuracy for self-ratings of intelligence suggests the utility of studying discrepancies between self-rated and tested intelligence (Ackerman, Beier,

[1] That the NEO PI-R contains only one Intellect facet and four Openness facets is an idiosyncrasy of that instrument and does not constitute evidence that Intellect is not central to the larger Openness/Intellect domain. The facets of the NEO PI-R were derived rationally, rather than empirically, and its authors have argued against Intellect as a valid interpretation of content in this domain (Costa & McCrae, 1992a; McCrae & Costa, 1997). As noted above, however, considerable evidence in both lexical and questionnaire research indicates that Intellect is just as central to the larger domain as Openness.

& Bown, 2002; Paulhus & John, 1998). Self-reported intelligence may reflect a combination of actual intelligence and inaccurate self-perception that could be due to over- or underconfidence. Indeed, self-esteem predicts the tendency to rate one's intelligence more highly than is warranted by one's tested intelligence (Gabriel, Critelli, & Ee, 1994). Gender is another predictor of self-rated intelligence, with men rating themselves higher than women do, even though no gender difference exists in general intelligence (Johnson & Bouchard, 2007; Syzmanowicz & Furnham, 2011). Men also score higher on Intellect and measures of intellectual engagement than women do (whereas women score higher in Openness), suggesting that men's tendency to be overconfident in their intelligence might also encourage them to be more intellectually engaged (Costa et al., 2001; Weisberg, DeYoung, & Hirsh, 2011). In addition to the male tendency to exaggerate intelligence, there is also a female tendency to underestimate (Kaufman, 2012; Steinmayr & Spinath, 2009; Syzmanowicz & Furnham, 2011).

So far in this section we have considered associations with general intelligence only. Studies that have examined verbal and nonverbal intelligence separately consistently show that Openness/Intellect is more strongly correlated with verbal than nonverbal intelligence (Ackerman & Heggestad, 1997; Ashton et al., 2000; Austin, Deary, & Gibson, 1997; Baker & Bichsel, 2006; Bates & Shieles, 2003; Beauducel et al., 2007; DeYoung et al., 2005, 2014; Holland et al., 1995; Stanek, 2014). This differential association has led many researchers to theorize that Openness/Intellect causes increased "crystallized" intelligence through increased motivation to learn and through investment in educational pursuits (Chamorro-Premuzic & Furnham, 2005; von Stumm & Ackerman, 2013).

The most thoroughly elaborated theory of this type is the Openness-Fluid-Crystallized-Intelligence (OFCI) model, which hypothesizes a number of developmental influences of Openness/Intellect on intelligence and vice versa (Ziegler et al., 2012). (Note, however, that OFCI does not distinguish between the Openness and Intellect aspects and refers to the broader Big Five dimension of Openness/Intellect as "Openness.") The OFCI's *environmental success hypothesis* posits that higher intelligence leads to higher Openness/Intellect because success in the intellectual domain leads to greater interest and engagement in that domain. The OFCI's *environmental enrichment hypothesis* posits a causal effect in the other direction, in which heightened curiosity associated with Openness/Intellect leads to greater exposure to complex environments, which encourages the development of (fluid) reasoning ability and, in turn, the greater acquisition of (crystallized) knowledge.

Although the developers of OFCI recognize that what has traditionally been called "fluid" intelligence can be influenced by environmental factors (as in their environmental enrichment hypothesis), they nonetheless conflate verbal tests with "crystallized" intelligence, and many of the other developmental investment theories do too (Chamorro-Premuzic & Furnham, 2005; von Stumm & Ackerman, 2013). The problem with interpreting the stronger correlation of Openness/Intellect with verbal than nonverbal intelligence as evidence of a developmental investment process is that, as discussed in the section "Definition of Intelligence," verbal intelligence cannot be equated to crystallized intelligence. Because both verbal and nonverbal intelligence are influenced by a mix of genetic and environmental forces, their differential associations with Openness/Intellect are uninformative regarding the causal relations between Openness/Intellect and intelligence.

To begin to elucidate such causal relations requires longitudinal data, preferably in a genetically informative sample that can help to rule out likely genetic confounds. One very lengthy longitudinal (but not genetically informative) study found no support for the idea that Openness/Intellect is related to change in intelligence over time, using IQ at ages eleven and seventy-nine years (Gow et al., 2005). Although Openness/Intellect, assessed at seventy-nine, was correlated with IQ at both ages (r = 0.32 at age eleven and 0.22 at age seventy-nine), it ceased to predict IQ at age seventy-nine after controlling for IQ at age eleven. Consistent with the argument that intelligence is a facet of Openness/Intellect, Gow and colleagues concluded that the variance shared between Openness/Intellect and intelligence probably just reflects the same stable trait of intelligence across the life span. A smaller longitudinal study, with more limited assessment of intelligence, did find some association of Openness/Intellect (rated by parents at age seventeen) with change in intelligence between the ages seventeen and twenty-three (Ziegler et al., 2012). Clearly, additional research is needed and genetically informative samples, such as in twin studies, would be a useful next step.

The differential association of Openness/Intellect with verbal and nonverbal intelligence can be clarified by separating the Openness and Intellect aspects. In the previous version of this chapter (DeYoung, 2011), I noted a pattern in which facets from the NEO PI-R that are markers of Openness appeared to be more weakly associated with nonverbal intelligence than did the Ideas facet (a marker of Intellect), whereas they had similar strength of association with verbal intelligence. Subsequently, we confirmed this pattern with the BFAS, finding in two samples that Intellect was almost equally strongly associated with both verbal and nonverbal intelligence, whereas Openness was associated only with verbal intelligence (DeYoung et al., 2014). Another study similarly found that Intellect but not Openness predicts nonverbal intelligence (though it did not assess verbal intelligence) (Nusbaum & Silvia, 2011). These findings were also supported by meta-analysis (Stanek, 2014) and explain why total Openness/Intellect is associated more strongly with verbal than with nonverbal intelligence, as well as casting further doubt on developmental theories that rely on this differential association. It also supports locating intelligence within Intellect taxonomically, given its relation to both the verbal and the nonverbal subfactors.

The question remains, however, as to why the Openness aspect is related to verbal but not nonverbal intelligence and remains related to verbal intelligence even after controlling for Intellect (DeYoung et al., 2014). One possible answer to this question was provided by a study demonstrating (1) that Intellect and Openness showed a double dissociation, whereby Intellect predicted working memory and Openness predicted implicit learning, and (2) that implicit learning was specifically associated with verbal ability but not with g (i.e., with its unique variance, as opposed to its variance shared with g) (Kaufman et al., 2010). Implicit learning refers to the ability to detect and learn patterns in sensory information automatically without conscious awareness. Given that Openness is associated with the tendency to perceive and enjoy patterns in sensory information, implicit learning is a sensible candidate as one of its functional substrates (DeYoung, 2015b). This implicit-learning ability may also facilitate language learning, much of which involves detecting statistical regularities in speech.

The link between intelligence and Intellect is reinforced by studies of working memory and brain function. Intelligence is very strongly associated with working memory, the ability to maintain and manipulate information in short-term memory, despite distraction (Kovacs & Conway, 2016). Further, the brain systems in the prefrontal cortex (PFC) and parietal cortex that support both working memory and intelligence overlap substantially, supporting the theory that working memory is one of the primary cognitive substrates of intelligence (Deary, Johnson, & Penke, 2010). Not surprisingly, therefore, Intellect appears to be associated with working memory capacity and also with its neural substrates (DeYoung et al., 2009; Kaufman et al., 2010). A neuroimaging study of brain activity during a difficult working memory task found that Intellect predicted neural activity associated with better working memory performance, in both the left frontal pole (most anterior region) of the PFC and a region of the medial PFC involved in monitoring performance and detecting the likelihood of error (DeYoung et al., 2009). The left frontal pole has been strongly implicated in *g* (e.g., Gläscher et al., 2010), and, indeed, when controlling for intelligence, the association between Intellect and neural activity in this area was attenuated, suggesting that this association reflects the fact that questionnaire measures of Intellect partially capture actual intelligence. The association between Intellect and neural activity in medial PFC, however, remained significant after controlling for intelligence, suggesting that this association might reflect the tendency toward cognitive engagement and effort that is also captured by Intellect. People who are more motivated to do well in cognitive tasks may be more likely to expend energy on monitoring their ongoing performance to detect and avoid errors.

A recent study inspired by this finding tested experimentally the hypothesis that Intellect is partly associated with performance on cognitive tasks because those higher in Intellect exert more effort (Smillie et al., 2016). Using a dual-task paradigm, this study showed that those high in Intellect were more susceptible to decrements in cognitive performance when required to engage in an additional secondary task, indicating that they were allocating more of their available cognitive resources to the primary task than were those low in Intellect. These results suggest that, although Intellect is associated with intelligence, this may not be exclusively due to differences in ability; Intellect reflects motivation as well as ability in the intellectual domain.

Another trait that falls within Openness/Intellect in lexical studies is creativity (Saucier, 1992), and both Openness/Intellect and intelligence are consistently positively associated with creativity, whether the latter is measured by trait-descriptive questionnaires, by real-world achievement, or by measures of creative ability in the laboratory, such as divergent thinking (Feist, 1998; Kaufman et al., 2016; Silvia, 2008). Another chapter provides in-depth review of the association of intelligence with creativity (see Chapter 45, by Plucker, Karwowski, & J. C. Kaufman, this volume). Creativity has often been considered a personality trait, which provides yet another reason to endorse the possibility of considering other cognitive abilities, including intelligence, as personality traits. In Figure 42.1, creativity could be listed with innovation and imagination in the central facet category of the Openness/ Intellect domain, with relations to both Openness, primarily for artistic creativity, and Intellect, primarily for scientific creativity (Kaufman et al., 2016).

One personality trait positively associated with both creativity and Openness is often (though not always) weakly negatively related to intelligence; this is apophenia, the tendency to detect patterns or causal connections where none in fact exist (DeYoung et al., 2012; Miller & Tal, 2007). The word "apophenia" was coined to describe the central symptoms of psychosis – hallucinations and delusions (Brugger, 2001) – but milder apophenia is a common phenomenon, including things like mistakenly thinking that one has heard one's name in a crowd, seeing faces in inanimate objects, and holding superstitious beliefs, such as astrology. A more common label for apophenia is "positive schizotypy," referring to characteristics associated with schizotypal personality disorder and risk for schizophrenia. People high in Openness are more likely to experience apophenia presumably because they detect more patterns in general and some of those patterns are Type I errors. Intelligence, however, should facilitate screening out false positives from real patterns, thus encouraging lower levels of apophenia. Despite their weak or negative correlation, intelligence and apophenia both load positively on the general Openness/Intellect factor and can potentially be considered peripheral facets of that Big Five dimension, with apophenia as a facet of Openness and intelligence as a facet of Intellect (Figure 42.1; DeYoung et al., 2012, 2016).

Extraversion

Extraversion comprises a set of lower-level traits related to approach behavior and positive affect, including assertiveness, talkativeness, drive, sociability, activity level, and positive emotionality. Extraversion appears to represent the manifestation in personality of sensitivity to rewards, both anticipated and received (DeYoung, 2015a; Wacker & Smillie, 2015). Meta-analyses of many studies shows that Extraversion is negligibly related to intelligence, with a correlation of 0.05 or less (Stanek, 2014; Wolf & Ackerman, 2005). Further, any weak positive association of intelligence with Extraversion might be artifactual, simply reflecting Extraversion's positive correlation with Openness/Intellect (DeYoung, 2006; Digman, 1997) rather than a real association with intelligence specifically.

Another possibility is that any weak associations of Extraversion with intelligence could reflect individual differences in low-level cognitive processes that are themselves only weakly related to intelligence. For example, Extraversion has been found to predict better short-term memory (Zeidner & Matthews, 2000), although it does not typically predict working memory, in which information in short-term memory must be manipulated or maintained despite distraction (DeYoung et al., 2005, 2009). Extraversion may be related to some aspects of intelligence test-taking, rather than to actual intelligence. Faster speed of test-taking and a lack of persistence during tests have been associated with Extraversion but results are equivocal (Chamorro-Premuzic, & Furnham, 2005; Doerfler & Hornke, 2010). In general, the cognitive correlates of Extraversion seem to be moderated by contextual factors, such as sensory stimulation and incentives (Eysenck, 1994; Pickering, 2004; Zeidner & Matthews, 2000). Perhaps because it primarily reflects basic positive emotional and motivational tendencies, Extraversion appears to be related to the stylistic ways in which people solve problems that require intelligence, while predicting their ability to solve them correctly only slightly, if at all.

Neuroticism

Neuroticism encompasses a variety of traits reflecting the tendency to experience negative emotion, including anxiety, depression, irritability, panic, and insecurity. It appears to reflect the primary manifestation in personality of sensitivity to threat and punishment (DeYoung, 2015a; Gray & McNaughton, 2000). Neuroticism exhibits a small but reliable negative correlation with intelligence, in the range of −0.10 to −0.15 (Ackerman & Heggestad, 1997; Stanek, 2014). This correlation is likely to be due to the facts that negative emotion typically interferes with higher cognition, in part by interrupting the functions of the PFC (Fales et al., 2008; Keightley et al., 2003), and that neurotic individuals are more likely to experience anxiety under the pressures of testing situations (Ackerman & Heggestad, 1997). Measures specifically designed to assess test anxiety are negatively correlated with intelligence, $r = -0.33$ (Ackerman & Heggestad, 1997). The most likely reason that this correlation is considerably stronger than the correlation of intelligence with Neuroticism is that traits are probabilistic, such that not everyone high in Neuroticism will experience a lot of test anxiety. Individuals who are high in Neuroticism and generally anxious may nonetheless be nonanxious while taking tests because of their particular histories and characteristic adaptations. (Similarly, individuals scoring low in Neuroticism, who are not generally anxious, may nonetheless be anxious about taking tests for reasons related to their personal histories.) Neuroticism is not inevitably associated with test anxiety, but the substantial correlation between the two ($r \approx 0.5$; Ackerman & Heggestad, 1997) means that high levels of Neuroticism increase the probability of anxiety during tests, which presumably leads to the small negative correlation between Neuroticism and intelligence. Hence the association of Neuroticism with intelligence is probably mediated by test anxiety (Moutafi, Furnham, & Tsaousis, 2006).

Longitudinal studies suggest a link between Neuroticism and change in IQ that may indicate a substantive association, rather than just a confounding by test anxiety. One such study, which assessed a large cohort at eleven years old and again at seventy-nine, found a small negative correlation ($r = -0.18$) of Neuroticism with change in IQ over that sixty-eight-year span (Gow et al., 2005), suggesting either that Neuroticism influences the development of intelligence or that it is linked to age-related declines in intelligence. The latter possibility is further supported by a longitudinal study of more than 600 adults over seventy-one years old, which found that Neuroticism predicted a steeper rate of cognitive decline over seven years (Chapman et al., 2012).

Agreeableness

Agreeableness reflects traits involved in altruism and cooperation, contrasting empathy, politeness, kindness, and humility with callousness, rudeness, aggression, and dishonesty. Meta-analysis has consistently indicated that Agreeableness is not associated with intelligence (Ackerman & Heggestad, 1997; DeYoung, 2011; Stanek, 2014). However, like Openness/Intellect, the two major subfactors or aspects of Agreeableness show differential association with intelligence. The aspects of Agreeableness are Compassion, reflecting the tendency to experience and express empathy, sympathy, and concern for others, and Politeness, reflecting the tendency to avoid being rude or belligerent and to refrain from manipulating or taking advantage of other people. (The term "compassion" is sometimes used more specifically to describe the desire to help others, explicitly differentiating this from "empathy," defined as sharing others' emotions, but Compassion in the Big Five hierarchy encompasses both of these things.) Whereas Compassion reflects emotional concern for others, Politeness seems to be less based in emotional connection and more in following social rules and inhibiting belligerent or socially disruptive impulses.

Research that separates Compassion and Politeness shows that, although Politeness is unrelated to intelligence, Compassion is positively related to intelligence, and more strongly to verbal than nonverbal ability, just like Openness, with correlations around 0.2 (DeYoung et al., 2014; Stanek, 2014). One possible explanation for this pattern is psychometric: Compassion is correlated with Openness, which may lead to an artifactual correlation with intelligence. Indeed, one study found that Compassion did not remain significantly associated with intelligence after controlling for Openness (DeYoung et al., 2014). However, another possible explanation is that Openness and Compassion share some of their underlying mechanisms. Openness involves the capacity for imagination, in the sense of simulating experience, such as an imagined future or a fictional world. Compassion also involves the capacity for imagination because, to understand what others are experiencing (known as "mentalizing ability"), one must imagine the world from their perspective.

Openness, Compassion, mentalizing, and imagination have all been linked to the brain's so-called default network, an extensive brain system involved in self-directed thought (in contrast to attention directed toward external stimuli) and in the simulation of experience in episodic memory, prospection, and mentalizing (Allen et al., 2017; Andrews-Hanna, Smallwood, & Spreng, 2014; Beaty et al., 2016). Further, the subnetwork of the default network that is most strongly linked to mentalizing ability is also linked to language processing, potentially helping to explain why Compassion and Openness are specifically linked to verbal intelligence (Andrews-Hanna et al., 2014).

Although Politeness appears not to be related to intelligence, aggression, which is a facet of Politeness (reversed), often is found to be negatively related to intelligence (Ackerman & Heggestad, 1997; DeYoung et al., 2008; Frisell, Pawitan, & Långström, 2012; Huesmann, Eron, & Yarmel, 1987; Séguin et al., 1999). However, results are inconsistent and meta-analysis suggests little association (Stanek, 2014), which may be due to the existence of different measures, types, and severities of aggression. Some questionnaire measures of aggression include being rude or pushy as instances of aggression, rendering the construct similar to Politeness, which is not correlated with intelligence. The examples of negative correlation cited above tend to focus on physical aggression, and perhaps it is physical aggression specifically that is linked to low intelligence.

Consistent with an association with physical aggression, intelligence is negatively associated with the broader trait of externalizing behavior, which includes antisocial behavior, impulsivity, and drug abuse, in addition to aggression (DeYoung et al., 2008; Krueger et al., 2007; Raine et al., 2005; Séguin et al., 1999). Among the Big Five, Agreeableness and Conscientiousness show the strongest (negative) correlations with externalizing behavior (Miller & Lynam, 2001). Behavioral genetic research suggests that the association between externalizing behavior and intelligence is genetically based (Koenen et al., 2006). Many questions remain regarding the association of aggression and antisocial behavior with intelligence, which will hopefully be clarified by future research that distinguishes between different types of aggression.

When components of Agreeableness such as detecting the emotional states of others or facilitating harmonious social relations are measured by ability tests rather than questionnaires, they are correlated with intelligence (Mayer, Salovey, & Caruso, 2004; Mayer, Roberts, & Barsade, 2008; Roberts, Schulze, & MacCann, 2008). This finding has emerged primarily from work on emotional intelligence, which has been defined as "the ability to engage in sophisticated information processing about one's own and others' emotions and the ability to use this information as a guide to thinking and behavior" (Mayer, Salovey, & Caruso, 2008, p. 503). Many questionnaires have been developed to assess emotional intelligence, but they reflect a diverse and rather incoherent collection of different conceptualizations of the construct (Mayer, Salovey, & Caruso, 2008; Roberts et al., 2008). Of more interest are ability tests that have been developed to assess emotional intelligence, most prominently the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT), which comprises a battery of subtests that involve tasks like identifying emotions in facial expressions or judging how best to manage others' emotions in social situations. Despite psychometric limitations (Barchard, 2003; Brody, 2004), the MSCEIT can be considered an encouraging example of the assessment of personality using ability tests rather than questionnaires. Scores on the MSCEIT are consistently associated with intelligence, with a correlation of about 0.3 (Mayer et al., 2004; Roberts et al., 2008). Like Openness and Compassion, the MSCEIT appears to be more strongly associated with verbal intelligence than with nonverbal intelligence (Mayer et al., 2004; Roberts et al., 2008).

Despite the fact that the MSCEIT is at least moderately related to intelligence, its primary association with the Big Five is with Agreeableness, rather than with

Openness/Intellect. Across a number of studies, scores on the MSCEIT have been found to be correlated with Agreeableness in the range of 0.20 to 0.30 (Mayer et al., 2008; Roberts et al., 2008). They are also correlated with Openness/Intellect, but more weakly, in the range of 0.10 to 0.20. Correlations with Extraversion, Neuroticism, and Conscientiousness are lower still (Mayer et al., 2004, 2008; Roberts et al., 2008). Thus, emotional intelligence has roughly the same magnitude of relation to Agreeableness that intelligence has to Openness/Intellect and self-reported intelligence. The ability to recognize and manage emotions effectively in social situations can potentially be considered a component of Agreeableness and one that is positively associated with general intelligence. One study found that most of the variance in the MSCEIT could be accounted for by *g*, Agreeableness, and gender (Schulte, Ree, & Carretta, 2004), suggesting that emotional intelligence might reasonably be considered a compound of Agreeableness and general intelligence.

Conscientiousness

Conscientiousness describes the tendency to be organized, self-disciplined, responsible, and hardworking, as opposed to lazy, messy, impulsive, and distractible. It appears to reflect the ability and tendency to prioritize nonimmediate or abstract goals, leading to the exertion of effort to pursue goals or follow rules (DeYoung, 2015a). Among the Big Five, it is the best predictor of both academic and occupational success, as well as health and longevity (Roberts et al., 2014). In fact, the only psychological trait that predicts these outcomes more strongly is intelligence. Interestingly, however, intelligence and Conscientiousness are nearly unrelated, and it may even be that Conscientiousness is weakly negatively related to intelligence, although the evidence is somewhat inconsistent (Ackerman & Heggestad, 1997; DeYoung, 2011; Stanek, 2014).

One potential explanation for a weak negative association of intelligence with Conscientiousness is provided by a theory of *compensation* (Chamorro-Premuzic & Furnham, 2005; Moutafi, Furnham, & Paltiel, 2004). People who are unintelligent may be more orderly in order to avoid complexity that they find difficult to manage because of their low intelligence. Similarly, they may tend to work extra hard, so as to accomplish tasks that could be performed more quickly or easily by someone more intelligent. Given that intelligence and Conscientiousness both predict academic and occupational success substantially but independently, this theory is plausible.

Evidence suggests, however, that it may be only one aspect of Conscientiousness that is negatively associated with intelligence and therefore a candidate as a compensatory mechanism. The two aspects of Conscientiousness are Orderliness and Industriousness (DeYoung et al., 2007), and meta-analysis suggests that Orderliness is weakly negatively correlated with intelligence, whereas Industriousness is positively related to intelligence (Stanek, 2014). A positive association of Industriousness with intelligence is consistent with the finding, noted in the previous section, that externalizing behavior is negatively correlated with both intelligence and Conscientiousness. When comparing the aspects of Conscientiousness, externalizing problems are more strongly associated with Industriousness than with Orderliness (DeYoung et al., 2016). Impulsivity is a core feature of externalizing problems related to Conscientiousness, and it too has been found to correlate negatively with intelligence (Kuntsi et al., 2004; Lynam et al., 1993; Vigil-Colet & Morales-Vives, 2005). (Note that some forms of impulsivity are more strongly associated with Neuroticism or Extraversion than with Conscientiousness, and different forms of impulsivity may be differentially associated with intelligence; DeYoung & Rueter, 2016; Whiteside & Lynam, 2001).

Conceptually, Conscientiousness is clearly linked to the tendency to forgo immediate rewards, in favor of longer-term goals. Normatively, people discount rewards that are delayed (Frederick, Loewenstein, & O'Donoghue, 2002), but the strength of this *delay discounting* shows considerable variability and has the characteristics of a stable personality trait (Kirby, 2009). Delay discounting is typically measured through a series of choices between smaller, more immediate rewards and larger, delayed rewards, with similar outcomes obtained whether these choices are hypothetical or actually result in reward (Shamosh & Gray, 2008). A meta-analysis of twenty-four studies indicated a correlation of −0.23 between delay discounting and intelligence (Shamosh & Gray, 2008). In one study, this association was partially mediated by working memory capacity and by neural activity in the same frontopolar region of the PFC discussed in relation to Intellect (Shamosh et al., 2008). Delay discounting is positively correlated with questionnaire measures of impulsivity (Hinson, Jameson, & Whitney, 2003; Ostaszewski, 1996; Richards et al., 1999; Swann et al., 2002) but only weakly correlated with Conscientiousness, with a correlation around −0.1 (Mahalingam et al., 2014).

Finally, in both childhood and adulthood, ratings of intelligence and Intellect in questionnaires are associated positively with Conscientiousness, and especially with Industriousness (Costa & McCrae, 1992a; DeYoung et al., 2007). In adults, this association does not prevent Intellect descriptors from loading primarily on a broader Openness/Intellect factor. In preschool-age children, however, this association appears to be strong enough that traits reflecting Intellect may group with Conscientiousness in factor analysis, rather than with traits that reflect Openness, such as perceptual sensitivity and enjoying low intensity sensations, which form their own separate factor (De Pauw, Mervielde, & Van Leeuwen, 2009; Shiner & DeYoung, 2013).

A link between Intellect and Conscientiousness may reflect their related biological substrates in the PFC (Shamosh et al., 2008). The lateral PFC is responsible for maintaining focus on nonimmediate goals and inhibiting impulsive responses (Bunge & Zelazo, 2006; Rueter et al., 2018), functions associated with Conscientiousness, but it is also responsible for manipulating information in working memory, functions associated with Intellect and intelligence (DeYoung et al., 2005, 2009). These two classes of PFC function, one more stabilizing and the other more flexible and exploratory, may be in tension, though both have been described as "executive function." As the PFC is developing rapidly in young children, differences in overall state of development might cause Intellect and Conscientiousness to covary (Shiner & DeYoung, 2013). After the PFC is more fully developed, however, the functional similarity of Intellect and Openness, as forms of exploratory cognition, may link Intellect more strongly with Openness than with Conscientiousness. Further, Conscientiousness and intelligence appear to be related to two distinct neural networks that both have nodes in lateral PFC: a goal priority network and a cognitive control network, respectively (Rueter et al., 2018). At biological, behavioral, and psychometric levels of analysis, the relation of intelligence to Conscientiousness and related traits is a pressing topic for investigation in personality psychology.

Sociopolitical Orientation

Although culturally specific social and political attitudes are clearly characteristic adaptations rather than traits, a general tendency toward conservatism versus liberalism (broadly defined) is a trait that might be found in any culture and that has been studied along with related traits such as right-wing authoritarianism and traditionalism (Bouchard et al., 2003; Koenig & Bouchard, 2006). Sociopolitical orientation receives a separate section here because it cannot easily be categorized within any one of the Big Five. Conservatism, authoritarianism, and traditionalism are associated negatively with Openness/Intellect but also positively with Conscientiousness and particularly Orderliness (Carney et al., 2008; Hirsh et al., 2010; Goldberg & Rosolack, 1994). Additionally, conservatism is associated negatively with the Compassion aspect of Agreeableness but positively with the Politeness aspect (Hirsh et al., 2010; Osborne, Wootton, & Sibley, 2012). Sociopolitical orientation thus appears to reflect a complex blend of multiple basic traits and this blend is consistent with the characterization of the core of conservatism as dislike of change and uncertainty, plus anti-egalitarianism, and the core of liberalism as openness to change, plus egalitarianism (Hirsh et al., 2010; Jost, 2017). (Note, however, that openness to change and egalitarianism are distinct dimensions that are nearly uncorrelated among people who are not politically engaged; Malka, Lelkes, & Soto, 2017.)

In keeping with their negative association with Openness/Intellect, conservatism and authoritarianism are negatively associated with intelligence. A meta-analysis estimates this correlation at around −0.15 for conservatism and −0.30 for authoritarianism (Onraet et al., 2015). However, the correlation with conservatism varies as a function of sample characteristics, such as age and the quality of assessment, and a number of studies find correlations of conservatism with intelligence in the range of −0.20 to −0.35 (e.g., Bouchard et al., 2003; Deary, Batty, & Gale, 2008; Koenig & Bouchard, 2006; Ludeke, Rasmussen, & DeYoung, 2017). Longitudinal studies have even found that childhood or adolescent IQ negatively predicts conservatism in adulthood (Block & Block, 2006; Hodson & Busseri, 2012). Like Openness and Compassion (but in the opposite direction), conservatism is more strongly related to verbal than nonverbal intelligence (Ludeke et al., 2017; Onraet et al., 2015).

In the NEO PI-R, the Values facet of Openness/Intellect assesses liberal versus conservative sociopolitical attitudes and an alternative measure of this facet has been labeled "Liberalism" (Goldberg, 1999). The Values facet seems to behave most like the Ideas facet in its association with intelligence and working memory, typically showing stronger correlations than the four Openness facets (Chamorro-Premuzic et al., 2005; DeYoung et al., 2005, 2009). However, Values does not clearly mark either the Intellect or the Openness aspect of Openness/Intellect, potentially because liberalism represents a compound of Openness/Intellect with other traits (DeYoung et al., 2007; Hirsh et al., 2010).

Liberalism is characterized by appreciation of diverse points of view and embrace of change, which may be facilitated by intelligence and working memory in part because change and consideration of diverse perspectives produce higher levels of complexity in experience. Such complexity may be difficult to manage for those of lesser intelligence (note the similarity of this argument to the one described in the "Conscientiousness" section regarding the compensatory negative association between Orderliness and intelligence; Orderliness is a strategy for reducing complexity). Further, liberalism is characterized by concern for the welfare of others, as reflected in its association with Compassion, and Compassion is also positively correlated with intelligence. Thus, most of the personality traits correlated with liberalism are correlated in the same direction with intelligence, which may reflect the fact that sociopolitical orientation is best considered to be a blend or compound of several basic traits, rather than a basic trait itself.

Nonlinear and Interactive Associations of Personality and Intelligence

Thus far, all associations of intelligence with other traits considered in this chapter have been linear and nonmultiplicative. A few studies, however, have examined more complex effects. Analyses of two large samples ($N > 1000$) and one larger still ($N > 70{,}000$) suggested an absence of nonlinear associations between intelligence and personality (Austin et al., 2002; Reeve, Meyer, & Bonaccio, 2006), but analysis of the even larger Project TALENT sample ($N > 360{,}000$) found a number of nonlinear associations, which could be important when considering the extremes of the intelligence distribution (Major, Johnson, & Deary, 2014). Participants both high and low in intelligence scored lower on Orderliness than those intermediate in intelligence. Additionally, two scales that showed nonlinear effects in opposite directions reflected the two major aspects of Extraversion: Sociability (corresponding to Enthusiasm) was lower at both extremes of intelligence, whereas Leadership (corresponding to Assertiveness) was higher at both extremes. This suggests the existence of substantive associations of Extraversion with intelligence, despite the absence of a linear relation, which might be suppressed if the two aspects are not considered separately.

In addition to nonlinear effects, there may also be important multiplicative or interactive effects of intelligence, in which the effect of intelligence differs depending on the level of other traits or vice versa. For example, intelligence may influence the effects of Neuroticism, as suggested by studies of interactions between Neuroticism and intelligence in predicting various outcomes. One such study found that leadership performance was predicted by this interaction (Perkins & Corr, 2006): For individuals high in Neuroticism, intelligence was positively associated with performance, whereas for those low in Neuroticism, intelligence was unrelated to performance. Another study found a similar effect for the interaction of Neuroticism and intelligence, among military conscripts, in predicting performance, physical health, and adjustment to military life (Leikas et al., 2009). Those high in Neuroticism showed poor performance, health, and adjustment only if they were low in intelligence. Intelligence, therefore, may act as a buffer for neurotic individuals, allowing them to cope with stressors despite heightened sensitivity to negative affect.

Intelligence has also been found to interact with Openness/Intellect in several studies. In a study of 180 psychology students, Openness/Intellect predicted vocabulary only at low levels of intelligence, suggesting that those who are highly intelligent do not need to make any particular effort to learn new words, whereas those who are relatively unintelligent will learn new words only if they are high in Openness/Intellect and hence curious and motivated to explore new information (Ziegler et al., 2012). (Because the vocabulary test in this study used a multiple-choice format designed to prohibit deduction of the correct answer, it was more reasonable as a measure of learned "crystallized" information than vocabulary tests that require spoken definitions of words.) A similar finding emerged in a sample of 836 Chinese secondary students, in which nonverbal intelligence interacted with Openness/Intellect to predict academic performance in Math, Chinese, and English (Zhang & Ziegler, 2015). Again, there was a positive correlation of performance with Openness/Intellect only for those low in intelligence, suggesting that the motivation associated with Openness/Intellect can compensate for low intelligence in challenging cognitive tasks, like schoolwork. Two studies in Germany, however, did not replicate this effect and instead found interactions of intelligence with Conscientiousness in predicting academic performance (Bergold & Steinmayr, 2018). In these samples, intelligence predicted performance more strongly in those high rather than low in Conscientiousness, suggesting that, among students low in Conscientiousness, highly intelligent students may fail to achieve up to their full potential or that, among relatively unintelligent students, being more conscientious may not lead to much improvement in grades. Given the disparities in results across different studies, additional research on the interaction of intelligence with other personality traits is clearly warranted.

Conclusions and Future Directions

Intelligence can be viewed either as a construct that is categorically distinct from personality or as one construct within the larger domain of personality. Neither viewpoint is supported by incontrovertible evidence, but I believe that psychology would benefit from the conceptual integration of intelligence and personality. The mandate of personality psychology is to understand the whole person as a coherent entity (DeYoung, 2015a; McAdams & Pals, 2006), and this goal can be furthered by consideration of intelligence as a personality trait. In discussing the relation of intelligence to Openness/Intellect, Saucier (1994, p. 294) wrote, “Intelligence is prone to suck in, or perturb the orbit of, any construct that comes near it.” This assertion evokes an image of personality traits as small planets orbiting a massive sun of intelligence. Framed grandiosely, one purpose of this chapter is to propose a Copernican revolution, whereby intelligence is now simply one trait among many, orbiting the central concept of personality. As mentioned in the section “The Conceptual Relation of Intelligence to Personality,” this proposal is not entirely novel but similar proposals in the past have not been much heeded. Given recent developments in understanding the difference between Openness and Intellect and their differential association with intelligence, the time may have come when this revolution is sufficiently empirically supported to gain traction.

The major conceptual barrier to integrating intelligence and personality is the old distinction between maximal performance and typical behavior. I suggest that this dichotomy, although intuitively appealing, may ultimately fail to distinguish personality from intelligence, both because individual differences in intelligence entail individual differences in typical behavior and because many personality traits encompass abilities other than intelligence. Broad personality traits reflect pervasive regularities in human functioning, and such regularities are likely to reflect types of challenge that are common in everyday life (DeYoung, 2015a; Nettle, 2006). Any such challenge provides an opportunity, or even a demand, for the application of relevant ability, ensuring that ability will be intimately tied to typical behavior. From this perspective, underlying most traits is both a motivational component – how likely the relevant mechanism is to be engaged – and an ability component – how likely the mechanism is to succeed when engaged (DeYoung, 2015a).

A full integration of intelligence with personality requires locating intelligence within hierarchical trait taxonomies, like the Big Five model. In the Big Five, descriptors of intelligence are located within the Intellect aspect of the broader domain of Openness/Intellect. As reviewed in this chapter, this location is strongly consistent with the patterns of correlation of intelligence tests with trait questionnaires. Having located intelligence within Intellect one can address what is perhaps a more interesting question: Are there personality traits *other* than Intellect that are associated with intelligence and, if so, why? Utilizing the Big Five framework, this chapter reviewed what is known about these associations and highlighted a number of empirical questions that should be addressed in future research.

Particularly interesting are the associations of intelligence with Agreeableness and Conscientiousness. As typically measured in Big Five questionnaires, both show little or no association. However, some of their lower-level aspects and facets, as well as conceptually related constructs such as delay discounting and emotional intelligence, do show significant associations with intelligence. Agreeableness reflects the mechanisms by which we are able to cooperate with others and Conscientiousness reflects the mechanisms by which we are able to follow rules and work toward distant goals. Understanding exactly how intelligence relates to these sophisticated psychological functions is of paramount importance for understanding personality as a coherent system. The relatively strong association of intelligence with Compassion (the strongest of any trait from the Big Five hierarchy outside the Openness/Intellect domain) is particularly interesting and warrants further study. It may be relevant for understanding the well-established negative correlation between intelligence and prejudice (Hodson & Busseri, 2012; Onraet et al., 2015).

Given that individual differences in the intelligence hierarchy below *g* appear to cluster according to whether they involve verbal or nonverbal operations, rather than according to whether they are crystallized or fluid (Johnson & Bouchard, 2005a, 2005b) and that Intellect is related to nonverbal intelligence almost as strongly as to verbal intelligence, new theories regarding the causal and developmental links between Openness/Intellect and intelligence probably need to be developed. Clearly, genetic versus experience-dependent aspects of intelligence are still of interest, but investigating them will be more challenging now that one cannot simply assume that any verbal tests assess crystallized intelligence while nonverbal tests assess fluid intelligence. One promising approach to experience-dependent abilities is to investigate domain-specific knowledge, while controlling for verbal and nonverbal intelligence (e.g., Ackerman, 2000). To test developmental theories adequately will require longitudinal, genetically informative designs.

Our understanding of personality generally and intelligence specifically will be enriched by considering how the psychological functions and biological systems that underlie intelligence are related to and interact with those that underlie other personality traits. A biological layer can be added to all of the questions raised in this chapter. In each case, we know relatively little about how the biological systems that underlie intelligence (Deary, Penke, & Johnson, 2010; Santarnecchi, Emmendorfer, Pascual-Leone, 2017) interact with the biological systems that underlie other personality traits (Allen & DeYoung, 2017). Pinpointing specific genetic and neurobiological mechanisms involved in the association of intelligence with other traits is an important project that has barely begun.

This project may be usefully guided by a cybernetic perspective on personality, in which traits are presumed to reflect variations in parameters of mechanisms that contribute to human goal pursuit (DeYoung, 2015a). In this framework, a key function associated with Openness/Intellect is to generate interpretations of the world through cognitive exploration, with Openness more oriented toward comprehending correlational patterns of association (such as those manifest in sensory experience) and Intellect more oriented toward comprehending causal and logical structure (DeYoung, 2015b; Kaufman et al., 2010). With this in mind, one can understand intelligence – a “capability for comprehending our surroundings” (Gottfredson, 1997a, p. 13) – as an important mechanism for interpretation of the causal and logical structure of experience, one that is complemented by intellectual engagement and by the aesthetic interests and abilities encompassed by Openness.

Acknowledgment

Because this chapter is an updated version of my chapter in the previous edition of this handbook, I would like, once again, to thank the people who read drafts of the chapter in the previous edition and provided helpful feedback: Tom Bouchard, Wendy Johnson, Niels Waller, Auke Tellegen, Aldo Rustichini, Raymond Mar, and Jacob Hirsh. Any remaining errors or infelicities are my own and may even have been introduced as I revised the chapter for this new edition.

References

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*, 229–259.

Ackerman, P. L. (2000). Domain-specific knowledge as the "dark matter" of adult intelligence: Gf/Gc, personality, and interest correlates. *Journal of Gerontology: Psychological Sciences*, *55B*, 69–84.

Ackerman, P. L. (2009). Personality and intelligence. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (pp. 162–174). New York: Cambridge University Press.

Ackerman, P. L., Beier, M. E., & Bowen, K. R. (2002). What we really know about our abilities and our knowledge. *Personality and Individual Differences*, *34*, 587–605.

Ackerman, P. L. & Goff, M. (1994). Typical intellectual engagement and personality: Reply to Rocklin (1994). *Journal of Educational Psychology*, *86*, 150–153.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.

Allen, T. A., & DeYoung, C. G. (2017). Personality neuroscience and the Five Factor Model. In T. A. Widiger (Ed.). *Oxford handbook of the Five Factor Model* (pp. 319–349). New York: Oxford University Press.

Allen, T. A., Rueter, A. R., Abram, S. V., Brown, J. S., & DeYoung, C. G. (2017). Personality and neural correlates of mentalizing ability. *European Journal of Personality*, *31*, 599–613.

Alvidrez, J., & Weinstein, R. S. (1999). Early teacher perceptions and later student academic achievement. *Journal of Educational Psychology*, *91*, 731–746.

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, *1316*(1), 29–52.

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Blas, L. D., Boies, K., & De Raad, B. (2004). A six-factor structure of personality descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, *86*, 356–366.

Ashton, M. C., Lee. K., Vernon, P. A., & Jang, K. L. (2000). Fluid intelligence, crystallized intelligence, and the Openness/Intellect factor. *Journal of Research in Personality*, *34*, 197–207.

Austin, E. J., Deary, I. J., & Gibson, G. J. (1997). Relationship between ability and personality: Three hypotheses tested. *Intelligence*, 25, 49–70.

Austin, A. J., Deary, I. J., Whiteman, M. C., Fowkes, F. G. R., Padersen, N. L., Rabbitt, P., Bent, N., & McInnes, L. (2002). Relationships between ability and personality: Does intelligence contribute positively to personal and social adjustment? *Personality and Individual Differences*, *32*, 1391–1411.

Baker, T. J., & Bichsel, J. (2006). Personality predictors of intelligence: Differences between young and cognitively healthy older adults. *Personality and Individual Differences*, *41*, 861–871.

Barchard, K. A. (2003). Does emotional intelligence assist in the prediction of academic success? *Educational and Psychological Measurement*, *63*, 840–858.

Bates, T. C., & Shieles, A. (2003). Crystallized intelligence as a product of speed and drive for experience: The relationship of inspection time and openness to g and Gc. *Intelligence*, *31*, 275−287.

Beaty, R. E., Kaufman, S. B., Benedek, M., Jung, R. E., Kenett, Y. N., Jauk, E., . . . Silvia, P. J. (2016). Personality and complex brain networks: The role of openness to experience in default network efficiency. *Human Brain Mapping*, *37*(2), 773–779.

Beauducel, A., Liepmann, D., Felfe, J., & Nettelnstroth, W. (2007). The impact of different measurement models for fluid and crystallized intelligence on the correlation with personality traits. *European Journal of Psychological Assessment*, *23*, 71–78.

Bergold, S., & Steinmayr, R. (2018). Personality and intelligence interact in the prediction of academic achievement. *Journal of Intelligence*, *6*, 27.

Block, J., & Block, J. H. (2006). Nursery school personality and political orientation two decades later, *Journal of Research in Personality*, *40*, 734–749.

Bouchard, T., Segal, N., Tellegen, A., McGue, M., Keyes, M., & Krueger, R. (2003). Evidence for the construct validity and heritability of the Wilson–Patterson conservatism scale: A reared-apart twins study of social attitudes. *Personality and Individual Differences*, *34*, 959–969.

Brickenkamp, R. (1975). *Handbuch psychologischer und padagogischer Tests* [Handbook of Psychological and Educational Tests]. Göttingen: Hogrefe.

Brody, N. (2004). What cognitive intelligence is and what emotional intelligence is not. *Psychological Inquiry*, *15*, 234–238.

Brugger, P. (2001). From haunted brain to haunted science. A cognitive neuroscience view of paranormal and pseudoscientific thought. In J. Houran & R. Lange (Eds.), *Hauntings and poltergeists: Multidisciplinary perspectives* (pp. 195–215). Jefferson, NC: McFarland & Company.

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, *15*, 118–121.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals differing in need for cognition. *Psychological Bulletin*, 119, 197–253.

Carney, D., Jost, J., Gosling, S., & Potter, J. (2008). The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, *29*(6), 807–840.

Carroll, J. B. (1993). *Human cognitive abilities*. New York: Cambridge University Press.

Cattell, R. B. (1950). *Personality*. New York: McGraw-Hill.

Cattell, R. B. (1978). *Scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.

Cattell, R. B., & Birkett, H. (1980). The known personality factors found aligned between first order T-data and second order Q-data factors, with new evidence on the inhibitory control, independence and regression traits. *Personality and Individual Differences*, *1*, 229–238.

Cattell, R. B., & Warburton, F. W. (1967). *Objective personality and motivation tests: A theoretical introduction and practical compendium*. Champaign: University of Illinois Press.

Chamorro-Premuzic, T., & Furnham, A. (2005). *Personality and intellectual competence*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chamorro-Premuzic, T., Moutafi, J., & Furnham, A. (2005). The relationship between personality traits, subjectively-assessed and fluid intelligence. *Personality and Individual Differences*, *38*, 1517–1528.

Chapman, B., Duberstein, P., Tindle, H. A., Sink, K. M., Robbins, J., Tancredi, D. J., & Franks, P. (2012). Personality predicts cognitive function over 7 years in older persons. *The American Journal of Geriatric Psychiatry*, *20*, 612–621.

Collis, J. M., & Messick, S. (Eds.) (2001). *Intelligence and personality: Bridging the gap in theory and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Costa, P. T., & McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences*, *13*, 653–665.

Costa, P. T., & McCrae, R. R. (1992b). *NEO PI-R Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T. Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*, 322–331.

Cronbach, L. J. (1949). *Essentials of psychological testing*. New York: Harper and Row.

De Pauw, S. S. W., Mervielde, I., & Van Leeuwen, K. G. (2009). How are traits related to problem behavior in preschoolers? Similarities and contrasts between temperament and personality. *Journal of Abnormal Child Psychology*, *37*, 309–325.

Deary, I. J. (2012). Intelligence. *Annual Review of Psychology*, *63*, 453–82

Deary, I. J., Batty, G. D., & Gale, C. R. (2008). Bright children become enlightened adults. *Psychological Science*, *19*, 1–6.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11*(3), 201–211.

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, *91*, 1138–1151.

DeYoung, C. G. (2011). Intelligence and personality. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 711–737). New York: Cambridge University Press.

DeYoung, C. G. (2015a). Cybernetic Big Five Theory. *Journal of Research in Personality*, *56*, 33–58.

DeYoung, C. G. (2015b). Openness/Intellect: A dimension of personality reflecting cognitive exploration. In M. L. Cooper & R. J. Larsen (Eds.), *The APA handbook of personality and social psychology: Personality processes and individual differences*, Vol. 4 (pp. 369–399). Washington, DC: American Psychological Association.

DeYoung, C. G., Carey, B. E., Krueger, B. F., & Ross, S. R. (2016). Ten aspects of the Big Five in the Personality Inventory for DSM-5. *Personality Disorders: Theory, Research, and Treatment*, *7*, 113–123.

DeYoung, C. G., Grazioplene, R. G., & Peterson, J. B. (2012). From madness to genius: The Openness/Intellect trait domain as a paradoxical simplex. *Journal of Research in Personality*, *46*, 63–78.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of Openness/Intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality*, *73*, 825–858.

DeYoung, C. G., Peterson, J. B., Séguin, J. R., Pihl, R. O., & Tremblay, R. E. (2008). Externalizing behavior and the higher-order factors of the Big Five. *Journal of Abnormal Psychology*, *117*, 947–953.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*, 880–896.

DeYoung, C. G., Quilty, L. C., Peterson, J. B., & Gray, J. R. (2014). Openness to Experience, Intellect, and cognitive ability. *Journal of Personality Assessment*, *96*, 46–52.

DeYoung, C. G., & Rueter, A. R. (2016). Impulsivity as a personality trait. In K. D. Vohs & R. F. Baumeister (Eds.). *Handbook of self-regulation: Research, theory, and applications* (3rd ed., pp. 345–363). New York: Guilford Press.

DeYoung, C. G., Shamosh, N. A., Green, A. E., Braver, T. S., & Gray, J. R. (2009). Intellect as distinct from Openness: Differences revealed by fMRI of working memory. *Journal of Personality and Social Psychology*, *97*, 883–892.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *73*, 1246–1256.

Doerfler, T., & Hornke, L. F. (2010). Working style and extraversion: New insights into the nature of item response latencies in computer-based intelligence testing. *Journal of Research in Personality*, *44*(1), 159–162.

Duckworth, A. L. (2009). (Over and) beyond high-stakes testing. *American Psychologist*, *64*, 279–280.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44* (4), 237–251.

Espejo, J., Day, E. A., & Scott, G. (2005). Performance evaluations, need for cognition, and the acquisition of a complex skill: An attribute–treatment interaction. *Personality and Individual Differences*, *38*(8), 1867–1877.

Eysenck, H. J. (1994). Personality and intelligence: Psychometric and experimental approaches. In R. J. Sternberg & P. Ruzgis (Eds.), *Personality and intelligence* (pp. 3–31). New York: Cambridge University Press.

Fales, C. L., Barch, D. M., Burgess, G. C., Schaefer, A., Mennin, D. S., Braver, T. S., & Gray, J. R. (2008). Anxiety and cognitive efficiency: Differential modulation of transient and sustained neural activity during a working memory task. *Cognitive, Affective, and Behavioral Neuroscience*, *8*, 239–253.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality & Social Psychology Review*, *2*, 290–309.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25–42.

Frederick, S., Loewenstein, G., & O’Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*, 351–401.

Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin*, *138*, 296–331.

Frisell, T., Pawitan, Y., & Långström, N. (2012). Is the association between general cognitive ability and violent crime caused by family-level confounders? *PLoS ONE*, *7*, e41783. https://doi.org/10.1371/journal.pone.0041783

Furnham, A., Dissou, G., Sloan, P., Chamorro-Premuzic, T. (2007). Personality and intelligence in business people: A study of two personality and two intelligence measures. *Journal of Business and Psychology*, *22*, 99–109.

Gabriel, M. T., Critelli, J. W., & Ee, J. S. (1994). Narcissistic illusions in self-evaluations of intelligence and attractiveness. *Journal of Personality*, *62*, 143–155.

Gläscher, J., Rudrauf, D., Colom, R., Paul, L. K., Tranel, D., Damasio, H., & Adolphs, R. (2010). Distributed neural system for general intelligence revealed by lesion mapping. *Proceedings of the National Academy of Sciences*, *107*(10), 4705–4709.

Goff, M., & Ackerman, P. L. (1992). Personality–intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology*, *84*, 537–552.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216–1229.

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, *4*, 26–42.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg: Tilburg University Press.

Goldberg, L. R., & Rosolack, T. K. (1994). The big five factor structure as an integrative framework: An empirical comparison with Eysenck's P-E-N model. In C. F. Halverson, Jr., G. A., Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 7–35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gottfredson, L. S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*, 13–23.

Gottfredson, L. S. (1997b). Why g matters: The complexity of everyday life. *Intelligence*, *24*, 79–132.

Gow, A. J., Whiteman, M. C., Pattie, A., & Deary, I. J. (2005). The personality–intelligence interface: Insights from an ageing cohort. *Personality and Individual Differences*, *39*, 751–761.

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system* (2nd ed.). New York: Oxford University Press.

Guilford, J. P. (1959). *Personality*. New York: McGraw-Hill.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*, 78–80.

Hill, B. D., Foster, J. D., Elliott, E. M., Shelton, J. T., McCain, J., & Gouvier, W. D. (2013). Need for cognition is related to higher general intelligence, fluid intelligence, and crystallized intelligence, but not working memory. *Journal of Research in Personality*, *47*(1), 22–25.

Hinson, J. M., Jameson, T. L., & Whitney, P. (2003). Impulsive decision making and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 298−306.

Hirsh, J. B., DeYoung, C. G., Xu, X., & Peterson, J. B. (2010). Compassionate liberals and polite conservatives: Associations of Agreeableness with political ideology and values. *Personality and Social Psychology Bulletin*, 36, 655–664.

Hodson, G., & Busseri, M. A. (2012). Bright minds and dark attitudes: Lower cognitive ability predicts greater prejudice through right-wing ideology and low intergroup contact. *Psychological Science*, *23*, 187–195.

Hofstee, W. K. B., Kiers, H. A., De Raad, B., & Goldberg, L. R. (1997). A comparison of Big Five structures of personality traits in Dutch, English, and German. *European Journal of Personality*, *11*, 15–31.

Holland, D. C., Dollinger, S. J., Holland, C. J., & MacDonald, D. A. (1995). The relationship between psychometric intelligence and the five-factor model of personality in a rehabilitation sample. *Journal of Clinical Psychology*, *51*, 79–88.

Horn, J. L., & Cattell. R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*, 253–270.

Huesmann, L. R., Eron, L. D., & Yarmel, P. W. (1987). Intellectual functioning and aggression. *Journal of Personality and Social Psychology*, *52*, 232–240.

Jang, K. L., Hu, S., Livesley, W. J., Angleitner, A., Riemann, R., & Vernon, P. A. (2002). Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. *Personality and Individual Differences*, 33, 83–101.

Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History: measurement, and conceptual issue. In O. P. John, R. W. Robins, & L. A. Pervin (Eds). *Handbook of personality: Theory and research* (pp. 114–158). New York: Guilford Press.

Johnson, J. A. (1994). Clarification of factor five with the help of the AB5C model. *European Journal of Personality*, *8*, 311–334.

Johnson, W., & Bouchard, T. J., Jr. (2005a). The structure of human intelligence: It's verbal, perceptual, and image rotation (VPR), not fluid-crystallized. *Intelligence*, *33*, 393–416.

Johnson, W., & Bouchard, T. J., Jr. (2005b). Constructive replication of the visual–perceptual-image rotation model in Thurstone's (1941) battery of 60 tests of mental ability, *Intelligence*, *33*, 417–430.

Johnson, W., & Bouchard, T. J., Jr. (2007). Sex differences in mental abilities: g masks the dimensions on which they lie. *Intelligence*, *35*, 23–39.

Johnson, W., Bouchard, T. J., Jr., McGue, M., Segal, N. L., Tellegen, A., Keyes, M., & Gottesman, I. I. (2007). Genetic and environmental influences on the Verbal-Perceptual-Image Rotation (VPR) model of the structure of mental abilities in the Minnesota study of twins reared apart. *Intelligence*, *35*, 542–562.

Johnson, W., Nijenhuis, J., & Bouchard, T. J. (2007). Replication of the hierarchical visual–perceptual-image rotation model in de Wolff and Buiten's (1963) battery of 46 tests of mental ability. *Intelligence*, *35*, 69−81.

Jost, J. T. (2017). Ideological asymmetries and the essence of political psychology. *Political Psychology*, *38*, 167–208.

Kaufman, J. C. (2012). Self estimates of general, crystallized, and fluid intelligences in an ethnically diverse population. *Learning and Individual Differences*, *22*, 118–122.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. J. (2010). Implicit learning as an ability. *Cognition*, *116*, 321–340.

Kaufman, S. B., Quilty, L. C., Grazioplene, R. G., Hirsh, J. B., Gray, J. R., Peterson, J. B., & DeYoung, C. G. (2016). Openness to Experience and Intellect differentially predict creative achievement in the arts and sciences. *Journal of Personality*, *84*, 248–258.

Keightley, M. L., Seminowicz, D. A., Bagby, R. M., Costa, P. T., Fossati, P., & Mayberg, H. S. (2003). Personality influences limbic-cortical interactions during sad mood. *NeuroImage*, *20*, 2031–2039.

Kirby, K. N. (2009). One-year temporal stability of delay-discount rates. *Psychonomic Bulletin & Review*, *16*, 457–462.

Koenen, K. C., Caspi, A., Moffitt, T. E., Rijsdijk, F., & Taylor, A. (2006). Genetic influences on the overlap between low IQ and antisocial behavior in young children. *Journal of Abnormal Psychology*, *115*, 787–797.

Koenig, L. B., & Bouchard, T. J., Jr. (2006). Genetic and environmental influences on the Traditional Moral Values Triad – Authoritarianism, Conservatism and Religiousness – as assessed by quantitative behavior genetic methods. In P. McNamara (Ed.), *Where god and science meet: How brain and evolutionary studies alter our understanding of religion*, Vol. 1, *Evolution, genes, and the religious brain*. Westport, CT: Praeger.

Kovacs, K., & Conway, A. R. A. (2016). Process Overlap Theory: A unified account of the general factor of intelligence, *Psychological Inquiry*, *27*, 151–177.

Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: An integrative quantitative model of the adult externalizing spectrum. *Journal of Abnormal Psychology*, *116*, 645–666.

Kuhn, T. (1970). *The structure of scientific revolution* (2nd ed.). Chicago: University of Chicago Press.

Kuntsi, J., Eley, T.C., Taylor, A., Hughes, C., Asherson, P., Caspi, A., et al. (2004). Co-occurrence of ADHD and low IQ has genetic origins. *American Journal of Medical Genetics*, *124*, 41–47.

Leikas, S., Mäkinen, S., Lönnqvist, J. E., & Verkasalo, M. (2009). Cognitive ability × emotional stability interactions on adjustment. *European Journal of Personality*, *23*(4), 329–342.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, *3*, 635–694.

Ludeke, S. G., Rasmussen, S. H. R., & DeYoung, C. G. (2017). Verbal ability drives the link between intelligence and ideology in two American community samples. *Intelligence*, *61*, 1–6.

Lynam, D. R., Moffitt, T. E., & Stouthamer-Loeber, M. (1993). Explaining the relation between IQ and delinquency: Class, race, test motivation, school failure, or self-control? *Journal of Abnormal Psychology*, *102*, 187–196.

Mahalingam, V., Stillwell, D., Kosinski, M., Rust, J., & Kogan, A. (2014). Who can wait for the future? A personality perspective. *Social Psychological and Personality Science*, *5*(5), 573–583.

Major, J. T., Johnson, W., & Deary, I. J. (2012). Comparing models of intelligence in Project TALENT: The VPR model fits better than the CHC and extended Gf–Gc models. *Intelligence*, *40*, 543–559.

Major, J. T., Johnson, W., & Deary, I. J. (2014). Linear and nonlinear associations between general intelligence and personality in Project TALENT. *Journal of personality and social psychology*, *106*(4), 638–654.

Malka, A., Lelkes, Y., & Soto, C. J. (2017). Are cultural and economic conservatism positively correlated? A large-scale cross-national test. *British Journal of Political Science*, 1–25. https://doi.org/10.1017/S0007123417000072

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, *88*, 139–157.

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, *59*, 507–536.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, *60*, 197–215.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2008). Emotional intelligence: New ability or eclectic traits? *American Psychologist*, *63*, 503–517.

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, *61*, 204–217.

McCrae, R. R. (1993). Openness to Experience as a basic dimension of personality. *Imagination, Cognition, and Personality*, *13*, 39–55.

McCrae, R. R., & Costa, P. T., Jr. (1997). Conceptions and correlates of Openness to Experience. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 825–847). Boston: Academic Press.

Miller, G. F., & Tal, I. R. (2007). Schizotypy versus openness and intelligence as predictors of creativity. *Schizophrenia Research*, *93*, 317–324.

Miller, J. D., & Lynam, D. R. (2001). Structural models of personality and their relation to antisocial behavior: A meta-analytic review. *Criminology*, *39*, 765–798.

Moutafi, J., Furnham, A., & Crump, J. (2003). Demographic and personality predictors of intelligence: A study using the NEO Personality Inventory and the Myers-Briggs Type Indicator. *European Journal of Personality*, *17*, 79–94.

Moutafi, J., Furnham, A., & Crump, J. (2006). What facets of openness and conscientiousness predict fluid intelligence score? *Learning and Individual Differences*, *16*, 31–42.

Moutafi, J., Furnham, A., & Paltiel, L. (2004). Why is Conscientiousness negatively correlated with intelligence? *Personality and Individual Differences*, *37*, 1013–1022.

Moutafi, J., Furnham, A., & Tsaousis, I. (2006). Is the relationship between intelligence and trait neuroticism mediated by test anxiety? *Personality and Individual Differences*, *40*, 587–597.

Mussel, P. (2013). Intellect: A theoretical framework for personality traits related to intellectual achievements. *Journal of Personality and Social Psychology*, 104(5), 885–906.

Neisser, U., Boodoo, G., Bouchard T. J., Jr, Boykin, A. W., Brody, N., Ceci, S. J., . . . & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American psychologist*, *51* (2), 77.

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, *61*, 622–631.

Nettle, D., & Liddle, B. (2008). Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. *European Journal of Personality*, *22*, 323–335.

Nolan, S. A., Roberts, J. E., & Gotlib, I. H. (1998). Neuroticism and ruminative response style as predictors of change in depressive symptomatology. *Cognitive Therapy and Research*, *22*, 445–455.

Nusbaum, E. C., & Silvia, P. J. (2011). Are openness and intellect distinct aspects of openness to experience? A test of the O/I model. *Personality and Individual Differences*, *51*, 571–574.

Oltmanns, J. R., & Widiger, T. A. (2016). Self-pathology, the five-factor model, and bloated specific factors: A cautionary tale. *Journal of Abnormal Psychology*, *125*, 423–434.

Onraet, E., Van Hiel, A., Dhont, K., Hodson, G., Schittekatte, M., & De Pauw, S. (2015). The association of cognitive ability with right-wing ideological attitudes and prejudice: A meta-analytic review. *European Journal of Personality*, *29*(6), 599–621.

Osborne, D., Wootton, L. W., & Sibley, C. G. (2012). Are liberals agreeable or not? Politeness and Compassion differentially predict political conservatism via distinct ideologies. *Social Psychology*, *44*, 354–360.

Ostaszewski, P. (1996). The relation between temperament and rate of temporal discounting. *European Journal of Personality*, *10*, 161–172.

Ostendorf, F., & Angleitner, A. (1994). Reflections on different labels for Factor V. *European Journal of Personality*, *8*, 341–349.

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, *66*, 1025–1060.

Paulhus, D. L., Lysy, D. C., & Yik, M. S. M. (1998). Self-report measures of intelligence: Are they useful as proxy IQ tests? *Journal of Personality*, *66*, 525–554.

Pedulla, J. J., Airasian, P. W., Madaus, G. F. (1980). Do teacher ratings and standardized test results of students yield the same information? *American Educational Research Journal*, *17*, 303–307.

Perkins, A. M., & Corr, P. J. (2006). Cognitive ability as a buffer to neuroticism: Churchill's secret weapon? *Personality and Individual Differences*, *40*, 39–51.

Peterson, J. B., & Flanders, J. L. (2002). Complexity management theory: Motivation for ideological rigidity and social conflict. *Cortex*, *38*, 429–458.

Pickering, A. D. (2004). The neuropsychology of impulsive antisocial sensation seeking personality traits: From dopamine to hippocampal function? In R. M. Stelmack (Ed.), *On the psychobiology of personality: Essays in honor of Marvin Zuckerman* (pp. 453–477). New York: Elsevier.

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science*, *10*(3), 282–306.

Pytlik Zillig, L. M., Hemenover, S. H., & Dienstbier, R. A. (2002). What do we assess when we assess a Big 5 trait? A content analysis of the affective, behavioral and cognitive processes represented in the Big 5 personality inventories. *Personality and Social Psychology Bulletin*, *28*, 847–858.

Raine, A., Moffitt, T. E., Caspi, A., Loeber, R., Stouthamer-Loeber, M., & Lynam, D. (2005). Neurocognitive impairments in boys on the life-course persistent antisocial path. *Journal of Abnormal Psychology*, *114*(1), 38–49.

Reeve, C. L., Meyer, R. D., & Bonaccio, S. (2006). Intelligence–personality associations reconsidered: The importance of distinguishing between general and narrow dimensions of intelligence. *Intelligence*, *34*, 387–402.

Richards, J. B., Zhang, L., Mitchell, S., & de Wit, H. (1999). Delay and probability discounting in a model of impulsive behavior: effect of alcohol. *Journal of the Experimental Analysis of Behavior*, *71*, 121–143.

Riemann, R., Angleitner, A., & Strelau, J. (1997). Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *Journal of Personality*, *65*, 449–476.

Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (2014). What is conscientiousness and how can it be assessed? *Developmental Psychology*, *50*, 1315–1330.

Roberts, R. D., Schulze, R., MacCann, C. (2008). The measurement of emotional intelligence: A decade of progress? In G. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The Sage handbook of personality theory and assessment*, Vol. 2. Los Angeles: Sage.

Rueter, A. R., Abram, S. V., MacDonald, A. W., Rustichini, A., & DeYoung, C. G. (2018). The goal priority network as a neural substrate of Conscientiousness. *Human Brain Mapping*, *39*(9), 3574–3585.

Saklofske, D. H., & Zeidner, M. (Eds.) (1995). *International handbook of personality and intelligence*. New York: Plenum Press.

Santarnecchi, E., Emmendorfer, A., & Pascual-Leone, A. (2017). Dissecting the parieto-frontal correlates of fluid intelligence: A comprehensive ALE meta-analysis study. *Intelligence*, *63*, 9–28.

Saucier, G. (1992). Openness versus intellect: Much ado about nothing? *European Journal of Personality*, *6*, 381–386.
Saucier, G. (1994). Trapnell versus the lexical factor: More ado about nothing? *European Journal of Personality*, *8*, 291–298.
Schulte, M. J., Ree, M. J., & Carretta, T. (2004). Emotional intelligence: Not much more than *g* and personality. *Personality and Individual Differences*, *37*, 1059–1068.
Séguin, J. R., Boulerice, B., Harden, P., Tremblay, R. E., & Pihl, R. O. (1999). Executive functions and physical aggression after controlling for attention deficit hyperactivity disorder, general memory, and IQ. *Journal of Child Psychology and Psychiatry*, *40*, 1197–1208.
Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R. A. et al. (2008). Individual differences in delay discounting: Relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science*, *19*, 904–911.
Shamosh, N.A., & Gray, J.R. (2008). Delay discounting and intelligence: A meta-analysis. *Intelligence*, *38*, 289–305.
Shiner, R. L., & DeYoung, C. G. (2013). The structure of temperament and personality traits: A developmental perspective. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology* (pp. 113–141). New York: Oxford University Press.
Silvia, P. J. (2008). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual differences*, *44*(4), 1012–1021.
Smillie, L. D., Varsavsky, V., Avery, R. E., & Perry, R. (2016). Trait intellect predicts cognitive engagement: Evidence from a resource allocation perspective. *European Journal of Personality*, *30*(3), 215–226.
Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*, 117–143.
Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.
Stanek, K. C. (2014). Meta-analyses of personality and cognitive ability. Doctoral dissertation, University of Minnesota.
Steinmayr, R., & Spinath, B. (2009). What explains boys' stronger confidence in their intelligence? *Sex Roles*, *61*, 736–749.
Sternberg, R. J., & Ruzgis, P. (Eds.) (1994). *Personality and intelligence*. New York: Cambridge University Press.
Swann, A. C., Bjork, J. M., Moeller, F. G., & Dougherty, D. M. (2002). Two models of impulsivity: relationship to personality traits and psychopathology. *Biological Psychiatry*, *51*, 988–994.
Syzmanowicz, A., & Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta analyses. *Learning and Individual Differences*, *21*, 493–504.
Vigil-Colet, A., & Morales-Vives, F. (2005). How impulsivity is related to intelligence and academic achievement. *The Spanish Journal of Psychology*, *8*, 199–204.
Von Stumm, S., & Ackerman, P. L. (2013). Investment and intellect: A review and meta-analysis. *Psychological Bulletin*, *139*, 841–869.
Wacker, J., & Smillie, L. D. (2015). Trait extraversion and dopamine function. *Social and Personality Psychology Compass*, *9*(6), 225–238.

Wainwright, M. A., Wright, M. J., Luciano, M., Geffen, G. M., & Martin, N. G. (2008). Genetic covariation among facets of openness to experience and general cognitive ability. *Twin Research and Human Genetics*, *11*, 275–286.

Wallace, J. (1966). An abilities conception of personality: Some implications for personality measurement. *American Psychologist*, *21*, 132–138.

Waller, N. G., DeYoung, C. G., & Bouchard, T. J. (2016). The recaptured scale technique: A method for testing the structural robustness of personality scales. *Multivariate Behavioral Research*, *51*, 433–445.

Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality across the ten aspects of the Big Five. *Frontiers in Personality Science and Individual Differences*, *2*, 178. https://doi.org/10.3389/fpsyg.2011.00178

Whiteside, S. P., & Lynam, R. W. (2001). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, *30*, 669–689.

Willerman, L., Turner, R. G., & Peterson, M. (1976). A comparison of the predictive validity of typical and maximal personality measures. *Journal of Research in Personality*, *10*, 482–492.

Wilt, J., & Revelle, W. (2015). Affect, behaviour, cognition and desire in the Big Five: An analysis of item content and structure. *European Journal of Personality*, *29*(4), 478–497.

Wolf, M. B., & Ackerman, P. L. (2005). Extraversion and intelligence: A meta-analytic investigation. *Personality and Individual Differences*, *39*, 531–542.

Woo, S. E., Chernyshenko, O. S., Longley, A., Zhang, Z. X., Chiu, C. Y., & Stark, S. E. (2014). Openness to experience: Its lower level structure, measurement, and cross-cultural equivalence. *Journal of Personality Assessment*, *96*(1), 29–45.

Woo, S. E., Chernyshenko, O. S., Stark, S. E., & Conz, G. (2014). Validity of six openness facets in predicting work behaviors: A meta-analysis. *Journal of Personality Assessment*, *96*(1), 76–86.

Yamagata, S., Suzuki, A., Ando, J., Ono, Y., Kijima, N., Yoshimura, K. et al. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology*, *90*, 987–998.

Zeidner, M., & Matthews, G. (2000). Intelligence and personality. In R. Sternberg (Ed.), *Handbook of intelligence* (pp. 581–610). New York: Cambridge University Press.

Zhang, J., & Ziegler, M. (2015). Interaction effects between openness and fluid intelligence predicting scholastic performance. *Journal of Intelligence*, *3*, 91–110.

Ziegler, M., Danay, E., Heene, M., Asendorpf, J., & Bühner, M. (2012). Openness, fluid intelligence, and crystallized intelligence: Toward an integrative model. *Journal of Research in Personality*, *46*, 173–183.

43 Intelligence and Achievement

Richard E. Mayer

This chapter examines the reciprocal relation between intelligence and achievement, particularly within academic domains such as verbal ability and mathematical ability. In particular, the chapter examines the specific knowledge needed for successful performance on tests of verbal ability that focus on decoding or reading comprehension and tests of mathematical ability that focus on solving arithmetic computation problems or arithmetic word problems.

Three Episodes in the History of Intelligence and Achievement

In the waning years of the nineteenth century, the world's first educational psychologist, E. L. Thorndike, undertook his first major experimental study of how learning works (Mayer, 2003a). Working in the attic of his advisor's house in Cambridge, Massachusetts, in a typical study, he put a hungry cat into a crate with a bowl of food just outside. If the cat pulled on a loop of string hanging overhead, a trap door would open and the cat could get out and eat the food. According to Thorndike, the cat began with a family of responses each linked to the situation in varying strengths based on past experience. Furthermore, Thorndike proposed that the cat learned by trial and error – unsuccessful responses were weakened each time they failed and successful responses were strengthened each time they worked. Thorndike called this learning principle *the law of effect* and it went one to become one of the fundamental pillars of learning theory and educational practice. Eventually, Thorndike reported his research in a book that he chose to call *Animal Intelligence* (Thorndike, 1911). Why did he claim to be studying intelligence? Thorndike sought to study the ability to learn, which he saw as "the most important of all original abilities" (p. 278). As you can see, from the very start, psychologists saw intellectual ability as the ability to learn and noted that it was based on prior learning experiences.

Next, let's shift the scene to Paris in the early 1900s where officials of the Paris school system were looking for ways to predict school success so they could identify students who might need special help before they get too far behind. They called on Alfred Binet, who is credited with inventing the world's first intelligence test (Wolf, 1973). Rather than viewing intelligence as a single monolithic ability, he posited that intelligence – or the ability to learn – was reflected in many smaller components. His test measured the many pieces of knowledge that children at various ages had acquired – what can be called achievement – such as the names of the colors of the rainbow or the counting of numbers from one to ten. Children who could answer factual questions customarily known by older children were considered above-average in intelligence because they had learned more from the same experiences as their peers. Similarly, children who could not answer factual questions customarily answered by their peers were considered below-average because they learned less based on the same experiences. His test was effective in predicting school success and became the basis for many subsequent intelligence tests. As you can see, Binet was the first to popularize the idea that intelligence – viewed as one's ability to learn – is reflected in achievement – viewed as what one has learned.

Finally, for our third historical scenario let's consider the saga of college entrance examinations produced by the Educational Testing Service (ETS) – America's largest testing organization, founded in 1947 in Princeton, New Jersey (Zwick, 2002). The SAT-1 is a well-known college entrance exam intended to predict college success by measuring verbal and mathematical abilities. Subsequently, the name of the section measuring verbal ability was changed from "verbal comprehension" to "critical reading" and a writing test was added (Hunt, 2011). Originally, the test was called the "Scholastic Aptitude Test," which was later changed to the "Scholastic Assessment Test," and eventually to simply the "SAT." What does the ambiguity over naming tell us about the relation between intelligence and achievement? It appears that the test was originally intended to measure aptitude – the ability to learn – but seems to have wound up measuring achievement – what students had learned. For example, mathematical test items include solving arithmetic word problems and the verbal test items include reading comprehension items. The newer SAT-II (formerly called "Achievement Tests") was designed to focus on the content of specific school subjects, reflecting the growing focus on past achievement as an indication of future learning ability. As you can see, the line between ability and achievement becomes blurred when tests originally intended to measure ability (e.g., the ability to learn) actually measure achievement (e.g., solving word problems and comprehending text). Thus, the SAT saga provides our third example of how intellectual ability – such as the ability to learn in school – appears to be intimately tied to achievement – such as what has already been learned in school.

One More Historical Clue: The Search for Attribute X Treatment Interactions

Are certain instructional methods better for one kind of learner and other methods better for a different kind of learner? If so, you would have evidence for an *attribute x treatment interaction* (or ATI). The modern search for ATIs dates back to Cronbach and Snow's (1977) heroic efforts, documented in their classic book *Aptitudes and Instructional Methods*, and continues today on many fronts (Massa & Mayer, 2006; Pashler et al., 2008; Sternberg & Zhang, 2001). The overwhelming consensus is that well-documented cases of ATIs are somewhat rare, so that the idea that teaching methods should be synced with the learner's cognitive style is an unproven myth of education (Holmes, 2016).

Does that mean that individual differences should not be taken into account when designing instruction? One important exception is that ATIs have been found when the individual differences dimension is the learner's prior knowledge. For example, Kalyuga (2014) has summarized evidence for the *expertise reversal effect* – the finding that instructional methods that are effective for low knowledge learners are not effective and may even be harmful for high knowledge learners and vice versa. In general, low knowledge learners perform best with well-structured instructional methods whereas high knowledge learners perform best with less-structured instructional methods (Mayer, 2011). This work suggests that, if you are interested in designing instruction for a learner, perhaps the single most important individual differences dimension for you to consider is the learner's prior knowledge (Kalyuga, 2014; Mayer, 2011). The expertise reversal effect has important implications for the relation between achievement and intelligence – showing that your past learning influences your ability to learn under different instructional methods. In short, the history of research on learning is studded with clues concerning the reciprocal relation between intelligence and achievement, which is the theme of this chapter.

What Is the Relation Between Intelligence and Achievement?

Taking an educational perspective, let's define academic intelligence as the ability to learn (e.g., performing a cognitive task) and let's define academic achievement as what is learned (e.g., specific knowledge). As shown in the top row of Table 43.1, academic intelligence can be measured by a person's performance on a cognitive ability test in which someone must accomplish an academic task such as comprehending printed text (i.e., verbal ability) or solving a story problem (i.e., mathematical ability). As shown in the bottom row of Table 43.1, academic achievement can be measured by a person's performance on a knowledge test aimed at assessing specific knowledge components (including facts, concepts, procedures, strategies, and beliefs).

More recently, researchers have reached consensus that the goal of education for life and work should be that students develop *transferable knowledge and skills* – that is, academic achievement that can be used in new situations to support new learning and

Table 43.1 *An educational approach to intelligence and achievement.*

Name	Definition	Example
intelligence	the ability to learn	Performance on an intelligence test intended to measure someone's ability to acquire knowledge from experience.
achievement	what is learned	Performance on an achievement test intended to measure someone's knowledge gained from experience.

creative problem-solving (Pelligrino & Hilton, 2012). In short, instead of simply being able to remember the information on a retention test, transferable knowledge is reflected in the ability to use the information in new situations on a transfer test. Employers report they are looking for people who are creative, innovative, adaptive, and flexible or, put another way, they want people who can solve problems, think critically, and engage in adaptive learning (Pelligrino & Hilton, 2012). These twenty-first-century attributes – reflecting aspects of intelligence – depend on an education system that fosters transferable knowledge and skills.

The unifying theme of this chapter is that there is a reciprocal relation between intelligence and achievement. First, intelligence (which is the ability to learn) helps you to acquire knowledge (which is the outcome of learning). In short, intelligence enables learning. Second, the knowledge that you have (i.e., achievement) improves your ability to learn (i.e., intelligence). In short, achievement enables intelligence. This reciprocal relation is illustrated in Figure 43.1.

How does the reciprocal relation between intelligence and achievement work? Consider the cognitive model of learning shown in Figure 43.2. Based on your experiences in the outside world, sounds and images enter your cognitive system through your ears and eyes, respectively, and are briefly held in your sensory memory. If you pay attention to this fleeting incoming material in sensory memory (indicated by the *selecting* arrow), some of the incoming material enters working memory where you mentally organize it (indicated by the *organizing* arrow) and integrate it with existing knowledge activated from long-term memory (indicated by the *integrating* arrow). Long-term memory is your large-capacity, permanent storehouse of knowledge; and working memory is your limited-capacity, temporary store for processing a small amount of material. Achievement is represented as knowledge

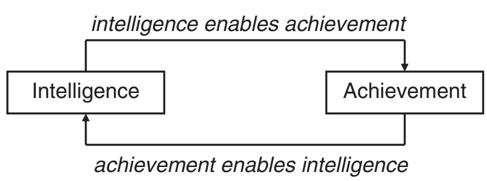

Figure 43.1 *The reciprocal relation between intelligence and achievement.*

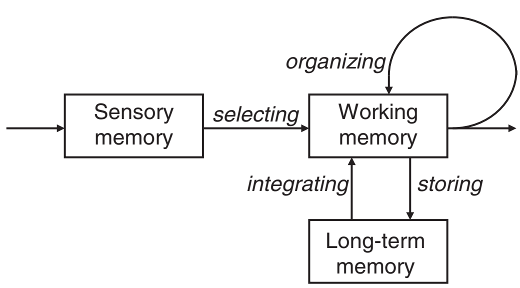

Figure 43.2 *Four cognitive processes in learning.*

in long-term memory and intelligence is represented as the appropriate use of cognitive processes during learning to acquire new knowledge in working memory (such as selecting, organizing, and integrating). These learning processes can be initiated and guided by prior knowledge activated from long-term memory.

What Causes Task Performance?

An important goal of education is to equip learners with what they need to know for accomplishing challenging tasks. Figure 43.3 shows a model of the factors involved in task performance – that is, performance on an academic task such as comprehending a passage or solving a mathematics problem (Mayer, 2003b). As you can see, task performance is indicated by the box on the right side of the figure. What are the determinants of the learner's task performance? The rightmost arrow in Figure 43.3 shows that the learner's knowledge – including facts, concepts, procedures, strategies, and beliefs – determines task performance. Where does the learner's knowledge come from? As shown in the left side of Figure 43.3, knowledge is the result of the combination of intelligence and experience, that is, knowledge depends on the learner having appropriate learning experiences (such as provided by appropriate instruction) and the ability to benefit from those learning experiences.

More recently, the conception of knowledge has been expanded to reflect the idea that learners bring more than their knowledge of reading, writing, and arithmetic to any new cognitive task. To better reflect the importance of metacognition as perhaps the epitome of intelligence, the conceptualization of strategies has been expanded to include metacognitive strategies aimed at monitoring and controlling cognitive processing during learning (Azevedo & Aleven, 2013; Mayer, 2011). To better reflect the role of the learner's motivation to learn, the conceptualization of beliefs has been receiving more attention, including how academic experience can shape and be shaped by a student's interests, values, goals, self-efficacy, attributions, and mindset (Wentzel & Miele, 2016).

The model presented in Figure 43.3 is based on research on the development of expert performance on cognitive tasks (Ericsson & Charness, 2006; Sternberg & Grigorenko, 2003). Let's consider three examples of relevant research findings.

First, when people begin to learn how to perform a cognitive task, their task performance is most strongly correlated with their general ability; but, as they progress from novice to expert, their task performance becomes increasingly more strongly correlated with their specialized knowledge (Ackerman & Beier, 2003; Krampe & Baltes, 2003). In short, as a learner gains expertise on a cognitive task, it appears that specialized knowledge comes to compensate for general ability. However, it is important to note that general

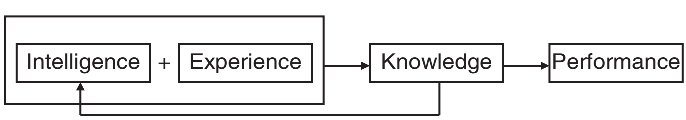

Figure 43.3 *What causes task performance?*

ability is not completely out of the loop because it may have enabled the creation of specialized knowledge, which in turn can be used to help learners to be even more effective in using their general ability for new learning.

Second, consider the Flynn effect. The Flynn effect refers to the finding that IQ scores were rising throughout the twentieth century at a rate of about three points per decade in each of twenty industrialized countries for which data are available (Flynn, 1998, 2009; Martinez, 2000). More recently, Barro and Lee (2015) have shown that educational attainment has been increasing around the world for more than 100 years, along with corresponding increases in economic and personal development. Martinez (2000) interprets the Flynn effect as showing that improvements in access to education serve not only to increase knowledge – what is learned – but also to improve intelligence – the ability to learn. Similarly, Ceci, Barnett, and Kanaya (2003) interpret the Flynn effect as evidence that intelligence and experience interact (as indicated in the left side of Figure 43.3) to produce improvements in the learner's knowledge. Ceci and colleagues propose a multiplier mechanism in which general ability may predispose learners to seek certain experiences, which result in specialized knowledge that enables them to use their general ability to learn even more effectively in that domain, resulting in more specialized knowledge that in turn increases the effectiveness of learning in the domain, and so on. The multiplier mechanism is consistent with viewing "ability + experience" in the left side of Figure 43.3 as an interactive process rather than one in which ability or experience dominates (Mayer, 2003b).

As a third example, consider the finding that deliberate practice can greatly enhance task performance (Ericsson, 2003, 2006). Deliberate practice occurs when a learner continually devotes considerable time and effort to practicing tasks that are challenging – that is, somewhat beyond the learner's current level of performance – and receiving useful feedback until reaching mastery. For example, Ericsson (2003) describes case studies in which people who engaged in concentrated practice in remembering digit lists showed impressive improvements in their digit span – from about 7 digits without deliberate practice to 20 digits after 50 hours of practice to 80 digits after 400 hours of practice. Based on numerous examples of how specialized practice can improve cognitive performance, Ericsson (2003) concludes that expert performance depends on acquiring specialized knowledge, as indicated in the right side of Figure 43.3. Importantly, the learner's willingness to engage in a large amount of deliberate practice may be dependent on the learner's ability and interest (Mayer, 2003b).

What Is Academic Ability?

Academic ability is a kind of intelligence most relevant to academic domains, such as the verbal domain and the mathematical domain. In particular, verbal ability refers to a person's ability to learn and perform verbal tasks, whereas mathematical ability refers to a person's ability to learn and perform mathematical tasks. Table 43.2 lists examples of several kinds of tasks related to verbal ability and mathematical ability. As shown in the top of Table 43.2, two important components of verbal ability are decoding and reading comprehension, whereas two important

Table 43.2 *Performance tasks and supporting knowledge for components of verbal and mathematical ability.*

Name	Performance Task	Supporting Knowledge
	verbal ability	
decoding	pronounce printed words or pseudowords	phonemes
reading comprehension	answer questions after reading a prose passage	prose schemas
	mathematical ability	
arithmetic	solve arithmetic computation problems	number sense
problem solving	solve arithmetic word problems	problem schemas

components of mathematical ability are arithmetic computation and problem-solving. Although this chapter focuses on verbal and mathematical ability, it should be noted that researchers have broadened the conception of academic ability beyond the cognitive domain to include the role of personal skills such as conscientiousness and social skills such as practical social intelligence (Pellegrino & Hilton, 2012; Sternberg et al., 2006). As examples of the relation between intelligence and achievement, the following two sections of this chapter explore the knowledge underlying verbal and mathematical aspects of academic ability.

What Is the Relation Between Intelligence and Achievement in the Verbal Domain?

Verbal ability is widely recognized as an important component of intelligence (Carroll, 1993, Hunt, 2011). Verbal ability refers to learning and performing on tasks that involve words. Within verbal ability two important factors are reading decoding, which is being able to pronounce printed words, and reading comprehension, which is being able to understand the meaning of a printed passage. In this section, let's examine the relation between intelligence and achievement for each of these two important types of verbal tasks.

First, consider the task of reading decoding—when you are given a printed word, you say it out loud. For example, given the printed word CAT, you have to blend the sounds /c/ and /a/ and /t/ into the spoken word, /cat/. Helping students develop decoding skill is perhaps the central mission of language arts instruction in the primary grades and is an essential skill for lifelong learning. As shown in the first row of Table 43.2, a common test of decoding is a word recognition test, which consists of asking students to pronounce a set of printed words (such as CAT), or a word attack test, which consists of asking students to pronounce a set of pseudowords (such as BLUD). Strong performance on such tests is an indication that the test-taker has high reading decoding skill, which is a key aspect of verbal ability.

What knowledge is needed to perform well on a word recognition or word attack test? Research on early reading shows that a particular kind of knowledge called *phonological awareness* is strongly related to decoding performance (Bradley & Bryant, 1983; Ehri et al., 2001; Goswami & Bryant, 1990). Phonological awareness refers to someone's knowledge of the sound units of their language – including knowing how to produce each of the sounds and knowing how to recognize each sound. In English, there are approximately forty-two sound units. For example, one test of phonological awareness involves substitution of the first phoneme, such as when the tester says, "Ball," but instead of /b/ begins the word with /p/. Students who enter primary school with high levels of phonological awareness tend to learn to read more easily and students who lack phonological awareness tend to have difficulty in learning to read (Bradley & Bryant, 1985; Juel, Griffin, & Gough, 1986; Wagner & Torgesen, 1987). Similarly, students who receive training in phonological awareness tend to show later improvements in reading (Bradley & Bryant, 1983; Ehri et al., 2001; Fuchs et al., 2001). Overall, research on phonological awareness is an example of the relation between knowledge (i.e., knowing the forty-two phonemes of English) and verbal ability (i.e., decoding performance).

Second, consider the task of reading comprehension – when you are given a printed passage, you are able to read for understanding so you can remember important information and answer questions about the content of the passage. As shown in the second row of Table 43.2, a common reading comprehension test involves being able to answer integrative questions, such as summarizing the passage or answering a question about the passage content in which you have to make an inference. Performance on reading comprehension tests can be considered a measure of verbal ability (Carroll, 1993; Hunt, 2011).

What knowledge is needed for success on a reading comprehension task? Research on reading comprehension shows that people perform better if they have domain knowledge, including schemas, that allow them to focus on important material (Bartlett, 1932; Lipson, 1983; Marr & Gormley, 1982; Pearson, Hansen, & Gordon, 1979). Importantly, teaching students about the schemas – or structures – for a given kind of prose material serves to improve their reading comprehension performance (Cook & Mayer, 1988; Ponce, Lopez, & Mayer, 2013; Taylor & Beach, 1984). Overall, research shows that domain-specific schemas for prose structure are prerequisites for reading comprehension performance. Similar findings have been reported for the central role of vocabulary in reading comprehension, in which a student's vocabulary (i.e., a type of knowledge) is highly related to reading comprehension (i.e., an aspect of verbal ability; Ash & Baumann, 2017; Perfetti & Stafura, 2014).

What Is the Relation Between Intelligence and Achievement in the Mathematical Domain?

Mathematical ability is widely recognized as an important component of intelligence (Carroll, 1993; Hunt, 2011). Mathematical ability refers to learning and performing on tasks that involve numbers. Within mathematical ability, two important tasks are arithmetic computation – being able to solve computational problems involving addition, subtraction, multiplication, and/or division – and problem-solving – being able to solve arithmetic word problems. These are summarized in the bottom of Table 43.2. In this section, we examine the relation between intelligence and achievement for each of these two important types of mathematical tasks.

First, consider the task of solving arithmetic problems – for example, given a printed problem such as 5 – 2 = ___, you compute a numerical answer. Solving computation problems is a fundamental component in mathematical ability and is part of tests intended to measure mathematical ability (Carroll, 1993).

What do you need to know in order to perform well on numerical computation problems? Research on arithmetic learning shows that an important prerequisite for computational performance is a form of conceptual knowledge that can be called *number sense* – the ability to represent numbers along a mental number line (Case & Okamoto, 1996; Griffin, Case, & Siegler, 1994). For example, number sense is indicated when a student determines which of two numbers is smaller or correctly moves a token along a path in a board game for a certain number of steps. Students who enter the primary grades without number sense tend to have more difficulty in learning arithmetic and students who are given direct instruction in how to use a mental number line tend to learn arithmetic more easily (Case & Okamoto, 1996; Griffin, Case, & Siegler, 1994; Jordan et al., 2012; Moreno & Mayer, 1999). Overall, there is convincing evidence of a strong relation between computational ability and knowledge of the mental number line (i.e., number sense).

Second, consider word problems in which you are given a verbal statement of a quantitative situation and must find an answer, such as:

> A car traveling at a speed of 30 miles per hour left a certain place at 10:00 a.m. At 11:30 a.m., another car departed from the same place traveling at 40 miles per hour and traveled the same route. At what time will the second car overtake the first car?

Performance on solving word problems such as this one is an indication of mathematical ability (Mayer, 2008; Reed, 1999).

What knowledge is needed for success on this test of mathematical ability? Research on mathematical problem-solving shows that students perform better when they possess appropriate problem schemas – mental categories for each kind of situation described in the problem (Hinsley, Hayes & Simon, 1977; Riley, Greeno, & Heller, 1982; Schumacher & Fuchs, 2012). For example, the car problem fits within the category of a time-rate-distance problem involving overtaking (Mayer, 1981). Problem-solvers are better able to mentally represent word problems when they can organize them based on a preexisting problem schema. This work is another example of how a form of academic ability (i.e., solving word problems) is highly related to the student's domain-specific knowledge (i.e., schemas for problem types). Determining the relation between cognitive ability and knowledge as it develops in specific domains is an important challenge for cognitive theory and educational practice.

Discussion

The theme of this chapter is that there is a reciprocal relation between intelligence and achievement, as exemplified within the academic domains of verbal ability and mathematical ability. In examining this theme, it is useful to consider the classic distinction between *fluid intelligence* – cognitive ability that is independent of specific knowledge – and *crystallized intelligence* – cognitive ability that depends on specific knowledge (Carroll, 1993; Hunt, 2011; Sternberg, 1990). In this chapter, my focus has been on crystallized intelligence, because of its importance for education. Crystallized intelligence is important for education because it can be changed through appropriate opportunities for learning. In short, the theme of this chapter is that specific kinds of knowledge that are the result of learning (i.e., achievement) can promote the ability to succeed in new learning (i.e., intelligence) and the ability to learn (intelligence) can help to enhance a learner's storehouse of relevant kinds of knowledge (i.e., achievement).

This analysis places knowledge at the center of the story. Table 43.3 summarizes five important kinds of knowledge and provides examples of each (Anderson et al., 2001; Mayer, 2011) – facts, concepts, procedures, strategies, and beliefs. An important goal of educational research is to pinpoint specific knowledge that enhances new learning, as suggested in the right column of Table 43.3. As you can see, the examples focus mainly on specific kinds of concepts that are useful for performing verbal tasks – namely, categorical knowledge of phonemes and schemas for prose structures – and specific kinds of concepts that are useful for performing mathematical tasks – namely, the concept of a mental number line and schemas for arithmetic word problems. This chapter has provided a glimpse into successful past research on the kinds of knowledge that enhance new learning and encourages a continuation and expansion of this fruitful line of research for the future. A promising direction for future research is to include an expanded view of the knowledge that the learner brings to learning venues – such as expanding *strategies* in

Table 43.3 *Five kinds of knowledge in academic tasks.*

Name	Definition	Example
facts	characteristics of elements	knowing the definitions of words; knowing that cars drive on roads
concepts	categories, principles, models, schemas	phonemes, prose schema, mental number line, problem schema
procedures	step-by-step processes	sound production algorithm, addition algorithm
strategies	general methods	comprehension monitoring strategy, self-evaluation strategy
beliefs	thoughts about one's learning	thinking that success depends on effort

Table 43.3 to include metacognitive strategies for monitoring and controlling cognitive processes during learning and expanding *beliefs* in Table 43.3 to include those that affect the learner's motivation to learn.

References

Ackerman, P., & Beier, M. E. (2003). Trait complexes, cognitive investment, and domain knowledge. In R. J. Sternberg & E. L. Grigorenko (Eds.). *The psychology of abilities, competencies, and expertise* (pp. 1–30). New York: Cambridge University Press.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Ash, G. E., & Baumann. J. F. (2017). Vocabulary and reading comprehension: The nexus of meaning. In S. E. Israel (Ed.), *Handbook of research on reading comprehension* (2nd ed, pp. 377–405). New York: Guilford Press.

Azevedo, R., & Aleven, V. (Eds.). (2013). *International handbook of metacognition and learning technologies*. New York: Springer.

Barro, R. J., & Lee, J-W. (2015). *Education matters: Global schooling gains from the 19th to the 21st century*. New York: Oxford University Press.

Bartlett, F. C. (1932). *Remembering*. London: Cambridge University Press.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read: A causal connection. *Nature*, *301*, 419–421.

Bradley, L., & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. Ann Arbor: University of Michigan Press.

Carroll, J. B. (1993). *Human cognitive abilities*. New York: Cambridge University Press.

Case, R., & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development*, *61*(1 & 2), i–295.

Ceci, S. J., Barnett, S. M., & Kanaya, T. (2003). Developing childhood proclivities into adult competencies: The overlooked multiplier effect. In R. J. Sternberg & E. L. Grigorenko (Eds.). *The psychology of abilities, competencies, and expertise* (pp. 70–93). New York: Cambridge University Press.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods*. New York: Wiley.

Cook, L. K., & Mayer, R. E. (1988). Teaching readers about the structure of scientific text. *Journal of Educational Psychology*, *80*, 448–456.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, *71*, 393–447.

Ericsson, K. A. (2003). The search for general abilities and basic capacities: Theoretical implications from the modifiability and complexity of mechanisms mediating expert performance. In R. J. Sternberg & E. L. Grigorenko (Eds.). *The psychology of abilities, competencies, and expertise* (pp. 93–125). New York: Cambridge University Press.

Ericsson, K. A., & Charness, N. (Eds.). (2006). *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.

Flynn, J. R. (1998). IQ gains over time: Toward finding the causes. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 25–66). Washington, DC: American Psychological Association.

Flynn, J. R. (2009). *What is intelligence? Beyond the Flynn effect*. New York: Cambridge University Press.

Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, A., Yen, L., Yang, N. J., Braun, M., & O'Connor, R. E. (2001). Is reading in reading readiness programs? A randomized field trial with teachers as program implementers. *Journal of Educational Psychology*, *93*, 251–267.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Griffin, S.A., Case, R., & Siegler, R.S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice*. Cambridge, MA: MIT Press.

Hinsley, D., Hayes, J. R., & Simon, H. A. (1977). From words to equations. In P. Carpenter & M. Just (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Holmes, J. D. (2016). *Great myths of education and learning*. Chichester: Wiley Blackwell.

Hunt, E. (2011). *Human intelligence*. New York: Cambridge University Press.

Jordan, N. C., Glutting, J., Dyson, N., Hassinger-Das, B., & Irwin, C. (2012). Building kindergartners' number sense: A randomized controlled study. *Journal of Educational Psychology*, *104*, 647–660.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*, *78*, 243–255.

Kalyuga, S. (2014). The expertise reversal principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed, pp. 576–597). New York: Cambridge University Press.

Krampe, R. T., & Baltes, P. B. (2003). Intelligence as adaptive resource development and resource allocation: A new look through the lenses of SOC and Expertise. In R. J. Sternberg & E. L. Grigorenko (Eds.). *The psychology of abilities, competencies, and expertise* (pp. 31–70). New York: Cambridge University Press.

Lipson, M. Y. (1983). The influence of religious affiliation on children's memory for text information. *Reading Research Quarterly*, *18*, 448–457.

Marr, M. B., & Gormley, K. (1982). Children's recall of familiar and unfamiliar text. *Reading Research Quarterly*, *18*, 89–104.

Martinez, M. E. (2000). *Education as the cultivation of intelligence*. Mahwah, NJ: Lawrence Erlbaum Associates.

Massa, L. J., & Mayer, R. E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer-visualizer cognitive style? *Learning and Individual Differences*, *16*, 321–336.

Mayer, R. E. (1981). Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instructional Science*, *10*, 135–175.

Mayer, R. E. (2003a). E. L. Thorndike's enduring contributions to educational psychology. In B. J. Zimmerman & D. H. Schunk (Eds.), *Educational psychology: A century of contributions* (pp. 113–154). Washington, DC: American Psychology Association.

Mayer, R. E. (2003b). What causes individual differences in cognitive performance? In R. J. Sternberg & E. L. Grigorenko (Eds.). *The psychology of abilities, competencies, and expertise* (pp. 263–274). New York: Cambridge University Press.

Mayer, R. E. (2008). *Learning and instruction* (2nd ed). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Mayer, R. E. (2011). *Applying the science of learning*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Moreno, R., & Mayer, R. E. (1999). Multimedia supported metaphors for meaning making in mathematics. *Journal of Educational Psychology*, *92*, 724–733.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*(3), 105–119.

Pearson, P. D., Hanson, J., & Gordon, C. (1979). The effect of background knowledge on young children's comprehension of explicit and implicit information. *Journal of Reading Behavior*, *11*, 201–209.

Pelligrino, J. W., & Hilton, M. L. (Eds.). (2012). *Education for life and work: Developing transferable knowledge and skills for the 21st century.* Washington, DC: National Academies Press.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*, 22–37.

Ponce, H. R., Lopez, M. J., & Mayer, R. E. (2013). A computer-based spatial learning strategy approach that improves reading comprehension and writing. *Educational Technology Research and Development*, *61*, 819–840.

Reed, S. K. (1999). *Word problems*. Mahwah, NJ: Lawrence Erlbaum Associates.

Riley, M., Greeno, J. G., & Heller, J. (1982). The development of children's problem solving ability in arithmetic. In H. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153–196). New York: Academic Press.

Schumacher, R. F., & Fuchs, L. S. (2011). Does understanding relational terminology mediate effects of intervention on compare word problems? *Journal of Experimental Child Psychology*, *111*, 2012, 607–628.

Sternberg, R. J. (1990). *Metaphors of mind*. New York: Cambridge University Press.

Sternberg, R. J., & Grigorenko, E. L. (Eds). (2003). *The psychology of abilities, competencies, and expertise*. New York: Cambridge University Press.

Sternberg, R. J., & The Rainbow Project Collaborators (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence*, 34, 321–350.

Sternberg, R. J., & Zhang, L.-F. (Eds.). (2001). *Perspectives on thinking, learning, and cognitive styles*. Mahwah, NJ: Lawrence Erlbaum Associates.

Taylor, B. M., & Beach, R. W. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly*, *19*, 134–146.

Thorndike, E. L. (1911). *Animal intelligence*. New York: Hafner.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*, 192–212.

Wentzel, K. R., & Miele, D. B. (Eds.). (2016). *Handbook of motivation at school*. New York: Routledge.

Wolf, T. H. (1973). *Alfred Binet*. Chicago: University of Chicago Press.

Zwick, R. (2002). *Fair game: The use of standardized admissions tests in higher education*. New York: Routledge Falmer.

44 Intelligence and Motivation

Priyanka B. Carr and Carol S. Dweck

To understand intelligence, one must understand motivation. In the past, intelligence was often cast as an entity unto itself, relatively unaffected by motivation. The prevailing view in the study of cognition and intelligence was that intellectual ability and intellectual performance were simply a function of the individual's cognitive apparatus (as noted by Dai & Sternberg, 2004). As far as motivation was concerned, everyone agreed, of course, that the "motor" had to be turned on but, beyond that, there was no well-articulated view of how motivational factors ignited and shaped intellectual performance. In this chapter, we attempt to articulate such a view.

What do we mean by motivation? Motivational factors – which can include beliefs, nonintellectual skills, and affect – are those factors that influence the pursuit of goals. In the present case, these goals are related to the acquisition and display of intellectual skills. In our chapter, we spell out how motivational factors determine (1) whether individuals initiate goals relating to the acquisition and display of intellectual skills, (2) how persistently they pursue those goals, and (3) how effectively they pursue those goals, that is, how effectively they learn and perform in the intellectual arena. As will be seen, motivational factors can have a systematic and meaningful effect on such indices of intellectual ability as grades, achievement test scores, and outstanding accomplishment.

Background

For many years, the focus in the study of intelligence was on documenting stable individual differences in intelligence (e.g., Conley, 1984; Galton, 1883; Jensen, 1998; Terman, 1926) rather than understanding the factors that shape it. Where did this notion of pure intelligence, unaffected by context, experience, or motivation, come from? Much of the impetus for this view came from implications of Darwinian theory, in particular, the ideas of variation within species and the survival of the fittest (Darwin, 1859). These implications were developed by Sir Francis Galton, Darwin's cousin, who had a passion for measuring human variation in all its forms and whose studies of eminent men and twins led him to conclude that

This chapter is an edited version of the chapter that appeared in the previous edition of the handbook: Carr, P. B., & Dweck, C. S. (2011). Motivation and intelligence. In R. J. Sternberg and S. B. Feldman (Eds.), *The Cambridge handbook of intelligence* (pp. 748–770). New York: Cambridge University Press.

nature rather than nurture was behind intelligence (Galton, 1883, 1892; Jensen, 2002).

Inspired in part by Galton, Lewis Terman (1916) adopted the view that intelligence reflects differences in “original mental endowment” (p. 4) and is more or less unchanged by other factors inside or outside of the individual. He wrote, “children from successful and cultured parents test higher than children from wretched and ignorant homes for the simple reason that their heredity is better” (p. 115). Terman believed that using the intelligence test he adapted for the American population (the Stanford-Binet), he could uncover a child’s level of innate intellectual ability and then ascertain the position that that child should occupy in society later in life (Terman, 1916, p. 18). In this view, motivation had little role either in intellectual ability or in long-term achievement.

However, this was not the only view. Alfred Binet, the co-creator with Theodore Simon of what was later called an “intelligence test” (Binet & Simon, 1913) (and was the test that Terman later revised for use in America), conceptualized intelligence very differently. He saw it, within limits, as malleable and trainable through education (Siegler, 1992). In fact, Binet did not believe his test tapped fixed intelligence at all. He emphasized that intelligence could manifest itself differently in different children and was developed at different rates through teaching (Siegler, 1992). Indeed, Binet expressed his alarm at the emerging view of intelligence as a fixed entity that could be measured by his test: “A few modern philosophers . . . assert that an individual’s intelligence is a fixed quantity, a quantity which cannot be increased. We must protest and react against this brute pessimism . . . With practice, training, and above all, method, we manage to increase our attention, our memory, our judgment and literally to become more intelligent than we were before” (Binet, 1909/1975, pp. 106–107). Interestingly, even Terman, after thirty-five years of following the children he classified as intellectually gifted, began to change his mind. He saw that many of his high-IQ participants achieved relatively little in life. In an effort to understand how this could be, he was led to conclude that motivational variables such as “persistence in the accomplishment of ends” and “integration toward goals” played a role in intellectual performance and life achievement (Terman & Oden, 1959, p. 149).

Certainly, people differ in their genetic endowments and perhaps in their aptitudes to begin with. However, it is becoming increasingly clear that intellectual ability is not static and can be meaningfully affected by nongenetic factors (see Sternberg, 2005; Sternberg & Grigorenko, 2001). For example, Ramsden and colleagues (2011) have demonstrated the possibility of substantial changes in verbal and nonverbal IQ and in corresponding brain structures over a three-to-four–year period in the normal adolescents they studied. There is also much emerging evidence about the dynamic nature of intellectual functioning and its components (see Diamond et al., 2007; Rueda et al., 2005). And a sophisticated understanding of the fascinating interplay of genes and environment is emerging, one that shows how that interplay gives rise to both heritability and malleability (Sauce & Matzel, 2018). Such theories show clearly how heritability does not imply nonmalleability. Thus, one is led away from questions about how to measure and classify people and toward questions about the factors that foster or inhibit the growth of intellectual abilities: What can lead us to be more (or less) intellectually competent than we were before?

Our perspective is that motivational factors have an important role to play in answering this question. As suggested earlier, we conceptualize motivational factors as variables that foster or interfere with effective goal pursuit and, in the case of intelligence, the effective pursuit of intellectual goals. We argue that motivation is much more than simply a motor that turns actions on or off and more than simply a desire to do well. Motivation, importantly, also involves beliefs (e.g., beliefs about the nature of one's intelligence), nonintellectual skills (e.g., the ability to enforce self-discipline to achieve one's goals), and affect (e.g., how much one enjoys learning in a particular area) – all of which influence people's ability to pursue intellectual goals effectively. There are several important implications of this approach. One is that context can have a strong, consistent impact on the motivation-relevant beliefs and affects that are activated and hence on intellectual performance. The second is that motivation-relevant beliefs, skills, and affect can be changed. That is, once one pinpoints the specific factors that play a role in intellectual performance, one can take steps to foster them and thereby enhance intellectual performance.[1]

We present evidence from laboratory studies, field studies, and interventions showing that beliefs, nonintellectual skills, and affective factors play a key role in intellectual performance. For example, we show that individuals' beliefs about intelligence, beliefs about stereotypes, and beliefs about "belonging" in a setting can influence intellectual performance and that training that speaks to these beliefs can improve intellectual performance. We also discuss how the emerging view of intellectual abilities as dynamic and as influenced by motivation is changing the field's view of giftedness and talent. It is changing the conception of giftedness from an endowment that needs only to be measured to emerging abilities that need to be cultivated and nurtured. We turn now to motivational factors that have been shown to influence intellectual performance.

Beliefs About the Nature of Intelligence

Research has found that people differ in how they view their intelligence. Some people lean toward the belief that intelligence is fixed (an *entity theory* or *fixed mindset* of intelligence) and others lean toward the belief that intelligence is malleable and can be affected by such things as good strategies, effort, and mentoring (an *incremental theory* or *growth mindset* of intelligence). These different beliefs about intelligence can lead to different motivational frameworks (Blackwell, Dweck, & Trzesniewski, 2007; Robins & Pals, 2002) and to differences in intellectual or academic performance (e.g., Aronson, Fried, & Good, 2002; Blackwell et al., 2007; Cury et al., 2008; Cury et al., 2006; Good, Aronson, & Inzlicht, 2003).

[1] We define intellectual performance as not just scores on IQ tests but more broadly as performance in a variety of intellectual tasks and domains.

A fixed mindset orients people to see intellectual performances as tests of their fixed level of intellectual ability. People endorsing this mindset thus tend to adopt *performance* goals more often than people with a growth mindset, striving to validate their intelligence through their performance (or to avoid negative judgments of their intelligence by avoiding challenges). A growth mindset about intelligence, on the other hand, is more likely than a fixed mindset to give rise to *learning* goals. Those with a growth mindset, because they believe intelligence can be improved, tend to see intellectual tasks and challenges as opportunities to cultivate ability rather than simply as occasions on which to impress (or disappoint) through performance (Blackwell et al., 2007; Dweck & Leggett, 1988; Robins & Pals, 2002).

Motivation, as we have defined it, is about the pursuit of goals. And the theory of intelligence one holds can affect not only which goal – performance or learning – is pursued but also how persistently it is pursued. While both performance and learning goals can be important for intellectual performance, a predominant focus on performance goals rather than learning goals can have potentially detrimental effects on intellectual ability and its growth over time. We present evidence that a fixed mindset and the performance goals it engenders can actually lead to less challenge-seeking (exposing oneself to fewer opportunities for learning) and can lead to lowered intellectual performance, as indexed by grades and achievement test scores. As we present the research below, it is important to remember that, while a person's mindset about intelligence can remain relatively stable over time, these theories are amenable to change and can be influenced through targeted interventions.

Mindsets About Intelligence and Intellectual Performance

Across different ways of assessing intellectual performance, there is increasing evidence that the mindset one holds affects intellectual performance. The evidence also indicates that mindsets affect intellectual performance through a motivational pathway, that is, through their effects on goals.

Academic performance: grades and achievement tests. First, we consider research (Blackwell et al., 2007) that examined intellectual performance (grades) across a difficult academic transition period – the transition to junior high school. In this work, researchers assessed students' mindsets through the students' agreement with items such as "You have a certain amount of intelligence and you really can't do much to change it" (with higher agreement indicating a more entity belief about intelligence) and "You can always greatly change how intelligent you are" (with higher agreement indicating a more incremental belief about intelligence). Blackwell and colleagues (2007) found that mindsets about intelligence and their associated (performance or learning) goals were significant predictors of grades, above and beyond prior achievement. For example, in this study, although students with fixed and growth mindsets entered junior high at the same level of prior math achievement, those with more of a growth mindset saw their math grades steadily increase while entity theorists showed no improvement. Blackwell and colleagues also demonstrated that students' goals and motivations mediated the effects of beliefs about intelligence on improved intellectual performance. Possessing a growth mindset, compared with a fixed mindset, led to increased endorsement of learning goals and increased belief in the importance of effort. These motivational factors and their downstream effects (e.g., positive, effort-based study strategies in response to difficulty) mediated the positive effect of a belief that intelligence is malleable on intellectual growth. Motivation based on beliefs about intelligence, and not prior ability level, was critical in predicting intellectual growth.

In their second study, Blackwell and colleagues (2007) demonstrated that students' beliefs about intelligence are malleable and that changing these beliefs could produce meaningful effects on intellectual performance. In this research, seventh-graders, many with declining math grades, were assigned to receive either training in study skills (control group) or an intervention that combined study skills with a growth mindset about intelligence. The growth mindset part of the intervention taught students that intellectual abilities were malleable (that their brains formed new or stronger connections every time they stretched themselves to learn something new) and that one could become smarter over time. Whereas the control group continued their decline in grades after the intervention, the growth mindset group did not: The intervention stopped the decline in grades and students in this group showed the beginnings of a rebound in grades following the intervention. In addition, teachers, who did not know which group students were in, were three times more likely to spontaneously report increased motivation for the students who were taught that intelligence is malleable than for the control students. It is essential to note that the control group received eight sessions of training in important study skills, skills that are key to intellectual performance. Moreover, they learned these skills quite well. Nonetheless, without the motivation to put them into practice, the skills did not express themselves in improved grades.

In another study, Aronson, Fried, and Good (2002) found that the effect of changing mindsets on intellectual performance extends as late as college. An intervention affirming that intellectual abilities are malleable significantly improved the African American students' enjoyment of academic work, the perceived importance of academic work, and their grade point averages (GPAs) one quarter later. The two control groups, one of which learned that intelligence was multifaceted and one of which received no treatment, showed no change in their academic enjoyment, values, or performance.

Another important intervention examined the impact of theories of intelligence on achievement test performance. Good and colleagues (2003) assigned adolescents to receive a growth mindset intervention (teaching them to view intelligence as malleable) or antidrug-use training at the start of the seventh grade. At the end of the school year, students were administered standardized tests. Those who had received the growth mindset training scored significantly higher on the test of reading achievement than did those in the control condition and females in the growth mindset conditions also earned higher scores on the test of math achievement than those in the control condition. The studies, then, demonstrate that changing students' beliefs about their intelligence can change their academic performance significantly.

IQ test performance. Studies (Cury et al., 2006, 2008) are also showing that people's beliefs about intelligence can affect not only grades or achievement test scores but also performance on an IQ test (administered after a mindset-related induction) – an area that many might have considered a motivation-free assessment of cognitive abilities. In one of these studies (Cury et al., 2006), adolescents in France were administered a portion of an intelligence test (the Coding Test of the Wechsler Intelligence Scale for Children – Third Edition, or WISC-III; Wechsler, 1996). Then, they were taught either that intelligence was fixed (the fixed mindset condition) or that intelligence was malleable (the growth mindset condition). After this, all participants completed another portion of the same IQ test. The two groups in the experiment did not differ in their performance on the first portion of the IQ test, before their beliefs were influenced. However, they differed significantly on the second portion of the test. Those in the fixed mindset condition performed significantly worse than those in the growth mindset condition. Moreover, the researchers found that adoption of performance goals mediated the relationship between theories of intelligence and intellectual performance. A fixed mindset created a goal of avoiding performance failure, which, in turn, led to hampered intellectual performance.

Mueller and Dweck (1998) found similar effects of motivational frameworks on IQ test performance after an experience of difficulty. In their studies, students were given a set of moderately difficult items from a nonverbal IQ test (Raven's Progressive Matrices; Raven, Court, & Raven, 1977), were all told that they had performed well, and were praised for their performance. Some were given praise for being intelligent (intelligence praise), some for working hard (effort praise), and some were given no additional praise (control). These different types of praise oriented students toward different theories of intelligence, with intelligence praise leading to more of a fixed belief about intelligence compared with effort praise, which led to more of a growth-oriented belief. The students then experienced difficulty on a second, very challenging set of problems from the same IQ test, after which they received a third set of problems that was matched in difficulty to the first set. We might expect that the students would do better on this third set (given the practice they had accumulated) or at least just as well as the first time around. However, how students performed depended on the praise that they had been received. Those in the control group slightly improved their performance. Those given the effort praise improved their performance significantly. But, importantly, those who were given intelligence praise performed significantly *worse* on the third trial than the first trial and significantly worse than the other two groups on this third set of problems. Again, this speaks against the IQ-test as a motivation-free assessment.

Summary. There is evidence from laboratory studies and from real-world field studies that beliefs about intelligence and their concomitant goals can affect intellectual performance as reflected in grades (e.g., Aronson et al., 2002; Blackwell et al., 2007), achievement test scores (e.g., Good et al., 2003), and IQ scores administered after a mindset-related manipulation (e.g., Cury et al., 2006, 2008). These effects may be most pronounced for groups of people who are facing challenges, whether it is a difficult school transition or the experience of failure (e.g., Blackwell et al., 2007; Mueller & Dweck, 1998). When concerns about one's level of fixed intelligence predominate and the motivation to learn remains in the background, intellectual performance can suffer. The research suggests that differences among people that may have been assumed to arise from differences in underlying intelligence may instead at times arise from differences in motivation. Furthermore, it is critical to note that mindsets and the associated motivations can be changed, and interventions that promote a growth mindset of intelligence can contribute to increases in intellectual performance.

Mindsets and Opportunities for Intellectual Growth

In this section, we propose that theories of intelligence may also affect intellectual ability in the longer term by changing people's reactions to opportunities for intellectual growth. With their belief that intelligence is immutable and their goal of demonstrating their intelligence, those with more of a fixed mindset might give themselves fewer opportunities to experience challenges and intellectual growth than those who hold a growth mindset.

In the Blackwell and colleagues (2007) research described earlier, students with a growth mindset expressed a greater preference for difficult tasks they could learn from than did entity theorists, who tended to prefer tasks that would allow them to perform well (see also Robins & Pals, 2002. A study by Dweck and Leggett (1988) examined whether mindsets about intelligence also translated to actual behavioral choices about challenging tasks. Adolescents were given a choice between tasks that were either within their comfort zone or not. They could choose to do tasks that were "fairly easy, so I'll do well," "problems that are hard enough to show I'm smart," or "problems that are hard, new and different so that I could learn." The first two task options allowed students to remain in or near their comfort zone – at a level at which they knew they could succeed. The last task, however, presented a novel challenge with opportunity to stretch themselves in the service of learning. While 61 percent of those endorsing more of a growth mindset chose the novel, challenging task, only about 18 percent of those endorsing a fixed mindset did so (see also Mueller & Dweck, 1998). Thus, the vast majority of those with a belief that intelligence was fixed denied themselves an opportunity to experience intellectual growth through novel tasks that pushed them out of their comfort zone.

Hong and colleagues (1999) found that those holding a fixed mindset were less likely than those with more of a growth mindset to take steps to improve their performance. They manipulated people's mindsets about intelligence and gave them a set of items from an intelligence test. Some participants were then told that their performance had been unsatisfactory and were offered a choice between an unrelated task or a task that would help them improve their performance on intelligence tests. Of those given the growth mindset, 73 percent chose the remedial task that would allow them to grow and improve. However, only 13 percent of those in the fixed mindset condition chose this remedial task.

There is also electrophysiological evidence that people holding a fixed mindset are more affected by information about their performance and that they less effectively process information that might help them learn. In this research, Mangels and colleagues (2006; see also Moser et al., 2011) used electroencephalography (EEG) to determine how people with different mindsets process performance-relevant and learning-relevant information. Each participant took a long and difficult test of general knowledge. After the participants answered each question (e.g., What is the capital of Nepal?), participants learned whether they got the question right or wrong and then a short time later what the right answer was. Analysis of the EEG data indicated that those endorsing fixed vs. growth mindsets differed in how they appraised negative feedback (i.e., you got the answer wrong). Those with more of a fixed mindset, compared with those with more of a growth, found the negative performance information to be more affectively significant, perhaps viewing it more as a threat to their adequacy than as a simple indication of where they needed to improve.

Mangels and colleagues (2006) also found brainwave patterns indicating that, based on their mindset, participants responded very differently to learning-relevant information (e.g., "The correct answer is Kathmandu"). Those with more of a fixed mindset, compared with those with more of a growth mindset, processed the correct answer in a less sustained and deep manner, thus encoding it less well. Moreover, the more sustained and deeper processing of the growth-mindset participants predicted better performance for them than for entity theorists on a subsequent surprise test of questions that they had answered incorrectly.

Summary. Research supports the idea that a fixed mindset compared to a growth mindset leads people to expose themselves to fewer challenging learning environments (e.g., Blackwell et al., 2007; Dweck & Leggett, 1988; Hong et al., 1999; see also Mueller & Dweck, 1998). Their appraisal of performance feedback as an indicator of their fixed intelligence appears to interfere with their ability to attend to and take advantage of learning opportunities, resulting in poorer learning (Mangels et al., 2006). There is additional evidence that performance goals (stronger for those with more of a fixed mindset), compared with learning goals, lead to engaging with material at a less nuanced and deep level and can therefore also create a less effective learning experience (Grant & Dweck, 2003). Through their avoidance of opportunities for challenging learning and their less effective processing of challenging learning material, those with a fixed mindset might experience less intellectual growth and lose ground to those with more of a growth mindset over time.

Beliefs About Being Viewed Through the Lens of a Stereotype

Believing that you may be judged through the lens of a negative stereotype, one that questions your underlying ability, can also affect intellectual performance. Many stereotypes cast groups of people – African Americans, Latinos, those of lower socioeconomic status, and women – as inherently lower in intelligence or in particular kinds of intellectual ability. However, much research finds that group differences in intellectual performance are far from fixed. Perhaps the most well-known of this type of research is the research on stereotype threat (Steele, 1997; Steele & Aronson, 1995). Stereotype threat is triggered when people believe that their performance may fulfill a negative stereotype about their group's ability and it has been shown to hamper intellectual performance (e.g., Aronson et al., 1999; Brown & Josephs, 1999; Croizet & Claire, 1998; Davies et al., 2002; Gonzales, Blanton, & Williams, 2002; O'Brien & Crandall, 2003; Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995; for meta-analyses, see Nguyen & Ryan, 2008; Walton & Cohen, 2003; Walton & Spencer, 2009). We describe the effects of stereotype threat and review evidence that these effects occur for motivational reasons.

Understanding Stereotype Threat

In the original study on stereotype threat, Steele and Aronson (1995) administered a measure of intellectual performance, the Graduate Record Exam (GRE), to Black and White college students. Half of the students were told that the test was diagnostic of intellectual ability (diagnostic condition) and the other that the experimenters were not interested in diagnosing ability (nondiagnostic condition). The instructions that the test was diagnostic of intellectual ability made the negative stereotype of lower ability relevant for Black participants, leading them to believe they could be judged through the lens of that stereotype. The effects of this minor manipulation on performance were striking. In the diagnostic condition, that is, when stereotype threat was present for the Black participants, a race gap in performance appeared: The Black participants underperformed relative to the White participants. However, when this threat was lifted and the test was described as nondiagnostic, the race gap was closed. This means that simply changing the instructions in a way that made people believe stereotypes were relevant or not relevant significantly changed intellectual performance.

There have been many other studies demonstrating the same phenomenon for multiple groups, such as those of lower socioeconomic status (e.g., Croizet & Clare, 1998), Latinos (e.g., Gonzales et al., 2002), women in math and science (e.g., Spencer et al., 1999), and the elderly (e.g., Andreoletti & Lachman, 2004). Effects have been found not only for standardized tests of performance but also for other markers of intellectual ability such as working memory, cognitive flexibility, and speed of processing (e.g., Carr & Steele, 2009; Schmader & Johns, 2003).

Does not require a history of stigmatization. Stereotype threat effects do not arise simply because a group has been chronically stereotyped. It is a threat cued by the situation. Even groups who have no history of stigmatization can be made to believe that they could be viewed as inherently inferior to others and, when they are, they tend to display lowered intellectual performance (Aronson et al., 1999). White men are typically unburdened by negative stereotypes impugning their academic abilities.

Yet, when told they are participating in a study examining why Asians are superior to Whites in math, White male math majors then underperform on a test of math ability. The situation cuing the belief that your performance could confirm the notion that your group is inferior can subvert intellectual performance.

Does not arise merely from knowledge of a group difference. Women are stereotyped as less able in math compared to men and they can experience stereotype threat and exhibit underperformance on math tests when told that there are gender differences on the math test they will take. This underperformance does not manifest itself when they are told that there are no gender differences (e.g., Spencer et al., 1999).

However, stereotype threat is also not always triggered from just being reminded that there are group differences in performance and that you belong to the disadvantaged group. It is more reliably triggered when there is an implication about your underlying capacity for success. Dar-Nimrod and Heine (2006) found that women who were told that gender differences in math performance were due to experiential causes, such as treatment by teachers, did not experience stereotype threat and they performed at the same high level as women who were told there were no gender differences. In contrast, women who were told that sex differences in math were due to genetic differences between males and females experienced stereotype threat and performed substantially worse. Thus, it is not just knowing or being reminded that gender differences exist that creates underperformance; it is the threat of your inherent capacity being questioned.

Summary. We have presented evidence that stereotype threat can interfere with intellectual performance (e.g., Steele & Aronson, 1995). Stereotype threat is created in a situation that signals that you might be judged through the lens of a negative stereotype and does not require a history of stigmatization (e.g., Aronson et al., 1999). It is, moreover, not triggered simply by the knowledge that your group may have underperformed in the past (e.g., Dar-Nimrod & Heine, 2006). It appears to stem from the indication that your group may be viewed as inherently deficient and that your performance may confirm this deficiency. We will suggest that stereotype threat affects intellectual functioning through its impact on motivational frameworks and resources.

The Motivational Argument

Much research has tried to understand exactly when, how, and why stereotype threat undermines intellectual performance (e.g., Ben-Zeev, Fein, & Inzlicht, 2005; Bosson, Haymovitz, & Pinel, 2004; Cadinu et al., 2003; Davies et al., 2002; Krendl et al., 2008; Schmader & Johns, 2003). We propose that one can understand the process through a motivational lens. Stereotype threat triggers evaluative concerns, that is, concern that poor performance will confirm a stereotype that questions underlying ability. These concerns can lead to a goal of displaying high not low intelligence to others (a performance goal) and can divert the mental resources needed for effective goal-pursuit and the achievement of high performance.

Under the burden of a stereotype about their group's innate intellectual inferiority, people can be expected to become preoccupied not with maximizing learning and absorbing information but rather with negative stereotypes and their performance. We propose that, while experiencing stereotype threat, a person's principal focus is not to grow and cultivate ability (a learning goal) but to perform well and not confirm the stereotype (a performance goal). Preliminary evidence discussed below supports this hypothesis, finding that, when they experience stereotype threat, people become focused on the stereotype and do not focus on learning (e.g., Davies et al., 2002; Krendl et al., 2008). In addition, research finds that changing motivational frameworks – orienting people toward a growth mindset and the associated learning goals – reduces stereotype threat and its negative effects on intellectual performance (Aronson et al., 2002; Good et al., 2003).

Preoccupation with stereotypes and performance. Studies have found that, after experiencing stereotype threat, the self-relevant negative stereotype becomes activated and salient for the targets of the stereotype. One such study (Steele & Aronson, 1995) found that Black participants in the stereotype threat condition compared to all other unthreatened participants completed more word-stems (e.g., d_ _b) with words related to the negative stereotype questioning their ability, such as "dumb" and "inferior," indicating that they were thinking of the negative stereotype more. In another study, women's level of activation of such stereotype-relevant words predicted their underperformance on a math test (Davies et al., 2002), suggesting that thinking about the stereotype that questions your ability actually hampers your ability to perform intellectually.

There is also some direct evidence that stereotype threat triggers preoccupation with performance and ability. For example, Cadinu and colleagues (2005) found that those experiencing stereotype threat had more negative thoughts about their performance and ability in math (e.g., I am not good at math) and that these thoughts mediated the effects of stereotype threat on underperformance.

This research, which indicates a preoccupation with stereotypes that indict ability and with poor performance under stereotype threat, suggests a shift to a motivational framework driven by performance goals. Indeed, neuroimaging data also support the idea that, burdened by stereotype threat, people become focused on evaluation and rejection and not on learning and deep processing. Krendl and colleagues (2008) used functional magnetic resonance imaging (fMRI) to investigate brain activation during stereotype threat. In their study, women took a math test in the fMRI scanner and were then either reminded of the negative stereotype about women's abilities in math (threat condition) or not (no threat condition). They then took another math test. On the second test, those who had not experienced stereotype threat increased recruitment and engagement of brain areas associated with processing mathematical information and mathematical learning (such as the left prefrontal cortex). They appeared to be increasing their engagement with and learning of the math material. In contrast, those reminded about the negative stereotype did not increasingly recruit these mathematical learning areas. They, instead, increased recruitment of the area of the brain that processes social and emotional information such as stereotypes and social rejection. Those not reminded of the stereotype did not increase activation of this area. Thus, it appears that, under stereotype threat, concerns about how others might view you and your performance become salient, and learning and deep processing may take more of a back seat. In this way, preoccupation with thoughts about stereotypes, evaluation, and ability may contribute to intellectual underperformance.

Changing motivational frameworks reduces stereotype threat. Perhaps the most striking evidence that motivational frameworks are important in the effects of stereotype threat on intellectual performance come from interventions designed to reduce the impact of stereotype threat on intellectual performance. Good and colleagues (2003) conducted an intervention to eliminate achievement gaps created by stereotype threat, specifically, a gender gap in math scores in junior high school. One group in their study received an intervention that taught them a growth mindset of intelligence, which, as we have seen, is typically associated with a greater focus on learning rather than performance goals. The control group simply received antidrug-use training. In the control group, girls underperformed relative to boys on the standardized math test administered at the end of the year. In the growth mindset group, however, the gender difference in performance was substantially reduced. Although boys also tended to experience an improvement in performance in the incremental group compared with the control, the positive effect was even stronger for the stereotype-threatened participants – the girls. Drawing the focus away from performance as an index of intelligence and putting it on brain growth and learning was especially beneficial for the group burdened by the stereotype (see also Aronson et al., 2002 for similar findings).

Summary. In certain circumstances, stereotype threat can impair intellectual performance on standardized tests. The burden of contending with stereotypes that characterize your group as inherently deficient can shift people to a performance-focused motivational framework and interfere with the ability to effectively pursue intellectual goals. As they become preoccupied with proving their ability, it can become more difficult to focus on and engage with learning (e.g., Krendl et al., 2008), cognitive resources may be sapped (e.g., Schmader & Johns, 2003), and strategies may become more inflexible (e.g., Carr & Steele, 2009). This shift in motivational framework and the sapping of goal-pursuit resources likely combine to create the depression of intellectual performance seen in targets of stereotypes.

A Note on Stereotype Lift

While we have focused on how the motivational effects of negative stereotypes interfere with intellectual performance, positive stereotypes can also affect intellectual performance. Negative stereotypes that cast doubt on the ability of one group (e.g., of women in math) also indicate that another group (e.g., men) is considered superior. Moreover, as the negatively stereotyped group can experience stereotype threat, those in the positively stereotyped group can experience stereotype lift – a boost in intellectual performance on the stereotyped task (e.g., a math test) (Walton & Cohen, 2003).

Stereotype lift has recently been found to be a case in which a motivational framework based on a fixed mindset of intelligence leads to *better* intellectual performance (Mendoza-Denton, Kahn, & Chan, 2008). Individuals who were viewed favorably through the lens of a stereotype (males in math), when told that ability was determined by innate factors (a fixed mindset view) performed better on a subsequent math test. In other words, knowing that the ability was fixed, and that they had it, made performance easier and better. However, given that a fixed mindset does not serve people as well in the face of setbacks (cf. also the effects of intelligence praise; Mueller & Dweck, 1998), given that a fixed mindset may not promote the growth of intellectual skills over time (e.g., Hong et al., 1999), and given the cost of fixed-mindset beliefs for those who are negatively stereotyped (e.g., Aronson et al., 2002), we believe that a growth-mindset motivational framework is overall more beneficial for intellectual performance.

Beliefs About Belonging

The need to belong is a powerful human motivator (Baumeister & Leary, 1995). As social animals evolved in small groups that worked cooperatively, humans are driven to fit in and belong in their social settings. In this context, it is not surprising that, when people are not certain about whether they belong in an academic setting, their motivation and ability to learn can be compromised.

We present evidence that uncertainty about belonging, perhaps by causing a shift in motivational frameworks, can make people seem "less intelligent than they were before." The research we review shows that people's beliefs about their belonging can affect performance on an IQ test and that interventions and procedures that heighten an individual's sense of belonging affect intellectual performance and effort.

Lack of Belonging Subverts Intellectual Performance

Baumeister, Twenge, and Nuss (2002) examined whether social rejection, which calls belonging into question, could actually lower performance on an IQ test. Participants in their study took a personality test and received experimentally manipulated feedback. In the social belonging condition, participants were told they would have many friends. In the social exclusion condition, they were told that they might lose friends. The control condition provided negative information to participants that was not social in nature. All participants then took an IQ test (General Mental Abilities Test; Janda, 1996). The social exclusion condition significantly reduced intellectual performance compared with the social belonging or control conditions. Those in the social exclusion condition got 25 percent fewer answers correct than those in the social belonging condition. Concern about social fit made participants appear less intelligent on the test.

Creating Belonging Improves Intellectual Performance

Walton and Cohen (2007) asked the flip side of the question that Baumeister and colleagues (2002) asked: What would happen to intellectual performance if you bolstered a sense of belonging for students who are typically stereotyped in intellectual settings? These students (e.g., Black students) may be particularly vulnerable to worrying about whether people fully accept them in school; that is, they may experience uncertainty about their belonging in academic settings. Walton and Cohen (2007) developed an intervention to alleviate students' uncertainty about their belonging. In it, they taught university freshmen that uncertainty about belonging is very common across all ethnic groups and that such worries dissipate over time. Students in the control condition were taught that social and political views become more sophisticated over time. The researchers followed these students throughout their college career and recorded the effects of their intervention on intellectual performance. The effects were fascinating.

The White students, who were not expected to be experiencing concerns about belonging in an academic setting, did not benefit from the intervention, as predicted. However, the Black students did benefit. One semester after the intervention, Black students in the control condition and campus-wide saw their grades decline. In contrast, Black students who received the belonging intervention actually saw their grades significantly improve. Moreover, these effects persisted over the next three years of college. At the end of college, the Black-White achievement gap (the discrepancy in grades) decreased by almost 70 percent in the treatment condition.

Why does a boost in belonging increase intellectual achievement? It may do so because it frees students from concerns about proving themselves (a performance goal) and allows them to engage with learning. In fact, Black students in the intervention group were more far likely to exhibit learning-motivated behavior, such as going to office hours, attending review sessions, and asking questions in class. Walton, Cohen, and colleagues are currently finding similar effects of a belonging intervention for women in male-dominated fields and for middle-schoolers from stereotyped groups as well (Walton et al., 2015).

Summary. Research supports the idea that beliefs about belonging affect intellectual performance. Uncertainty about belonging can hamper performance on an IQ test and adversely affect grades in college (Baumeister et al., 2002; Walton & Cohen, 2007). Being freed from this uncertainty, it appears, allows individuals to focus on learning, increase their intellectual effort, and improve their pursuit of intellectual goals.

The Skill of Self-Regulation

To this point we have discussed how different motivation-relevant beliefs – about intelligence, about stereotypes, and about one's belonging – change intellectual performance. Now, we turn to another critical component of motivation – people's skill at self-regulation – and its impact on intellectual performance.

Self-regulation is the executive function process that directs cognitions, attention, and behaviors toward the attainment of an individual's goals in the face of other information (internal or external) that competes for the individual's attention (Baumeister & Heatherton, 1996, Baumeister et al., 1998; Engle, 2002; Kane et al., 2007). It is the resource we use when we undertake a challenging goal, when we choose to study instead of going out with friends, when we keep working when tired, and when we tune out an exciting conversation to stay focused on our work. It is a resource necessary for effective goal-pursuit.

In self-regulation, we see the intertwining of intellectual abilities and motivation. Attention-regulation and response-inhibition are considered to be part of executive function, but executive function also includes working memory (Engle, 2002), a more purely intellectual factor. In this section, we will focus on people's self-control skills to highlight the role they play in intellectual performance. These skills – specifically, delay of gratification, self-discipline, and behavioral control abilities – can have important effects on intellectual outcomes, affecting standardized test scores, academic success, professional success, and intellectual growth and learning.

Delay of Gratification and Self-Discipline

Walter Mischel and his colleagues (e.g., Mischel, Shoda, & Rodriguez, 1989) have long called attention to the important role of the ability to delay gratification and regulate oneself in service of one's goal. The importance of such regulation was underscored by Duckworth and Seligman (2005), who examined the effects of self-discipline and delay of gratification with eighth-graders. Using self-report, teacher reports, parent reports, and delay of gratification tasks (e.g., "Would you like $1 now or $2 next week?"), the researchers derived a self-discipline score for each student in the fall of the school year. These students were also given an IQ test. The researchers then tracked students' grades, their scores on standardized achievement tests, and their selection into a rigorous and competitive high school program – all intellectual performance variables – through the spring of that school year. They found that even after controlling for prior achievement, highly self-disciplined adolescents had higher grades than their less disciplined counterparts. In addition, they outperformed those lower in self-discipline on the other measures of intellectual performance. What was particularly impressive was that self-discipline predicted more variance in these intellectual outcomes than did the adolescents' IQ scores. What many people would consider a measure of pure intellectual ability – the IQ test – was not as effective in predicting intellectual success as was a motivational variable like self-discipline (see also Tangney, Baumeister, & Boone, 2004; Wolfe & Johnson, 1995).

It makes sense that self-discipline and delay of gratification would be important for intellectual success. Even the most gifted students may not get very far if they do not spend time pursuing challenging learning (see, e.g., Ericsson, Krampe, & Tesch-Römer, 1993).

Behavioral Regulation and Effortful Control

A closely related construct that has received a lot of attention recently is that of behavioral regulation and effortful control – the ability to follow instructions and inhibit inappropriate responses (Blair & Razza, 2007; McClelland et al., 2007). Behavioral regulation and effortful control, as well, have been found to affect intellectual performance. In one study, researchers (McClelland et al., 2007) measured preschooler's behavioral regulation ability in the fall and spring of their prekindergarten year using a "Head-to-Toes" game in which the children have to do the opposite of what the experimenter asks them to do (e.g., touch their toes when asked to touch their head). This task demands self-regulatory skill, as it requires the child to inhibit the dominant, inappropriate response and keep the task goal and rules salient in the face of distraction. Researchers also measured the children's math, vocabulary, and literacy abilities at both times. They found that children's behavioral control predicted their intellectual performance at both points in time. Furthermore, growth in a child's behavioral regulation ability predicted improvement in intellectual performance: Making great gains in behavioral regulation from fall to spring predicted making great gains in math, vocabulary, and literacy, even after controlling for prior achievement. In a similar study, Blair and Razza (2007) found that a teacher's reports of a child's effortful control ability in preschool (how able a child is to stay focused on activities, control responses when asked to, and not become frustrated) predicted math performance in kindergarten, even after controlling for IQ as measured by Raven's Progressive Matrices (Raven et al., 1977). Thus the degree to which a child can inhibit inappropriate responses and not succumb to distractions – can effectively self-regulate in the pursuit of his or her goals – predicts intellectual performance and intellectual growth (see also Bull & Scerif, 2001; Espy et al., 2004; Howse et al., 2003; Ponitz et al., 2009; St Clair-Thompson & Gathercole, 2006; Valiente et al., 2008).

Improving Self-Regulation

It is clear that self-regulation skill measured early in life can have an impact on intellectual outcomes even much later in life. However, that does not mean that self-regulation abilities are unchangeable or simply proxies for intelligence. In fact, research has shown that they can be trained. In one study (Diamond et al., 2007), researchers used the "Tools of the Mind" materials (which included training in inhibiting responses, sustaining attention, and keeping information in mind over time) to teach executive function to one group of preschool children. It was woven into the standard curriculum and the "Tools of the Mind" group was later compared to a similar group of children who received only the standard curriculum. At the end of one to two years of such training, their executive function abilities were measured on self-regulation tasks that were not familiar to any of the children. On these tasks that measured ability to tune out distracters and inhibit natural responses, particularly the challenging versions, the children who had received the "Tools of the Mind" training significantly outperformed the children who had received the standard curriculum. Thus, a curriculum focused on self-regulation had successfully increased self-regulation (executive function) capacity in young children (see also Dowsett & Livesey, 2000; Rueda et al., 2005). Moreover, as we have discussed, performance on tasks demanding self-regulation are predictive of academic achievement.

Summary. The evidence is clear in showing that self-regulation – people's skill at setting and maintaining their focus on their goals – is critical to short- and long-term intellectual performance (e.g., Duckworth & Seligman, 2005; McClelland et al., 2007). The effects of self-regulation on intellectual performance appear to be long-lasting and are obtained above and beyond the effects of prior achievement and IQ scores. Taken together with the recent success in training self-regulation (e.g., Diamond et al., 2007), these findings again support the idea that intelligence can be influenced by motivation-related factors.

Feelings of Intrinsic Motivation

We last consider the effects of affective components of motivation on intellectual performance. We first describe research that finds that the affective states of pleasure, enjoyment, and interest (that accompany and constitute "intrinsic" motivation for an activity) enhance intellectual performance, leading to higher grades and test scores. We then turn to a related definition of intrinsic motivation – engaging in an activity for its own sake rather than simply because of external demands and pressures (Sansone & Harackiewicz, 2000). Research finds that such internally driven motivation enhances intellectual performance.

Researchers have examined whether creating learning environments that enhance interest leads to better intellectual performance. In one study (Cordova & Lepper, 1996), researchers used several strategies to increase elementary school students' intrinsic interest in a game that taught arithmetic operations. The instructional content was identical in all conditions but, in some conditions, the researchers increased intrinsic motivation and interest by adding an element of fantasy (e.g., participants would advance a spaceship through solving math operations), creating personalization (e.g., participant's name and birthday was included in the game), or allowing participants choice (e.g., naming their character and the opponent's character). One to two weeks after the game was played, participants were given a written test of equations. Compared to the control condition, which was not designed to increase intrinsic interest, these strategies significantly improved performance on the math test. Thus, although all students received the same instruction, students who experienced greater intrinsic interest during the instruction exhibited better intellectual performance.

Another study investigated the effects of goals that were intrinsic in nature and contexts that were supportive of autonomy. Self-determination theory proposes and research finds that tasks that satisfy a need for autonomy are more intrinsically motivating (Deci, Koestner, & Ryan, 1999; Deci & Ryan, 1985; Ryan & Deci, 2000). Vansteenkiste and colleagues (2004) found that people performed significantly better on a test of new material when the material was framed in terms of intrinsic goals (e.g., material allowing personal growth) and not extrinsic goals (e.g., material allowing you to earn more) and when people were made to feel autonomous and volitional (e.g., by using phrases such as "you can" and "if you choose" in instructions) rather than controlled (e.g., by using phrases such as "you must" and "you have to" in instructions).

Iyengar and Lepper (1999) found that providing choice (by allowing students to pick which puzzles to work on) in contrast to not providing choice (by assigning students puzzles picked by authority figures) increased motivation for European American students. For the more interdependent, Asian American students, choices made by valued and trusted others (such as their mother or their in-group) produced high intrinsic motivation but choices made by lesser valued others (such as the out-group) undermined their motivation. And, across all cultures, situations that enhanced intrinsic motivation led to improved task performance. Thus, it appears that contexts that facilitate intrinsic motivation lead to better learning, comprehension, and intellectual performance.

We now turn to intrinsic motivation defined in a different but very related way – engaging in a task for its own sake or on your own terms. Of course, engaging in tasks for such reasons may also be accompanied by greater interest and enjoyment and the findings we discuss later may be mediated by such affective states. Several longitudinal studies have investigated whether children who possess higher intrinsic motivation for academics and learning – a desire to learn for learning's sake – actually perform better academically in school. In one such study (Lepper, Corpus, & Iyengar, 2005), students' intrinsic motivation was measured through agreement with items such as "I work on problems to learn how to solve them." The researchers found that higher intrinsic motivation for academics predicted higher grades and higher standardized test scores months later. In contrast, higher extrinsic motivation, motivation arising from external rewards or pressure (assessed by agreement with items such as "I work on problems because I'm supposed to"), was negatively correlated with future grades and standardized test scores. Other studies have found similar effects. Being intrinsically motivated for academics correlates with increased academic achievement (e.g., Gottfried, 1985; Gottfried, 1990; Gottfried, Fleming, & Gottfried, 2001; Harter, 1981). Though both extrinsic and intrinsic motivation may reflect a desire to do well, pursuing academic activities for their own sake appears to be associated with better intellectual performance.

Moreover, research also finds that interference with this desire to engage in an activity for its own sake through superfluous extrinsic rewards leads to worse performance (Lepper, Greene, & Nisbett, 1973). Researchers recruited children in a nursery school who had shown existing intrinsic interest in a drawing activity. They then either asked the children to simply engage in the drawing activity or asked them to engage in it in exchange for an extrinsic reward (a certificate with a gold star). Researchers found that the "over-justification" for the drawing activity created through the extrinsic reward actually lowered children's future interest in the activity and led to drawings of a lower quality.

Extrinsic rewards and extrinsic motivation may certainly "turn on the motor." However, as noted, research finds that intrinsic motivation – defined either as an affective state of interest and enjoyment or an internally driven motivation to engage with the material – is associated with greater academic achievement as reflected in intellectual performance.

It is important to note that extrinsic rewards may not always be detrimental to performance, especially if there was no intrinsic interest to begin with. There has been a recent push to pay students for academic performance and it is possible that such programs could jump-start engagement with academic work for some students. However, these programs must be seen in light of the decades of research on the benefits of intrinsic motivation and in the context of research on the beneficial impact of interventions that teach a growth mindset of intelligence and those that create a sense of belonging for these same groups of lower achieving or negatively stereotyped students. The implication is that such programs might be supplemented or replaced by programs in which students are motivated to learn in order to grow their brains and because school is a place where they belong and are valued.

Conclusions

In this chapter, we have presented research conducted in the laboratory and in field settings demonstrating the effects of motivational variables on intellectual outcomes as varied as grades, achievement on standardized tests, IQ test scores, and professional accomplishment (e.g., Aronson et al., 2002; Blackwell et al., 2007; Cury et al., 2006; Ericsson et al., 1993). And the research indicates that sometimes these dynamic motivational variables – individually and taken together – may be as important or more important than traditional measures of intellectual ability, like IQ, in predicting intellectual performance (e.g., Duckworth & Seligman, 2005). Moreover, the effects of motivation on intellectual performance emerge among individuals of equal cognitive ability and at equal levels of prior intellectual accomplishment.

Importantly, this research suggests motivational routes to enhancing intellectual accomplishment and has implications for our understanding of giftedness and intelligence, as it draws our attention to the importance of educational environments and cultures. Indeed, highlighting the point that motivation is amenable to change, we have described several empirically tested avenues for enhancing intellectual performance through affecting motivation (e.g., Aronson et al., 2002; Blackwell et al., 2007; Cordova & Lepper, 1996; Diamond et al., 2007; Good et al., 2003; Walton & Cohen, 2007).

The ability to change motivation and thereby change intellectual performance may also push us to alter the focus of intelligence and giftedness research. The focus in intelligence and giftedness research has long been on identifying those who are highly intelligent or gifted and tracking and supporting them (e.g., Colombo et al., 2009; Gagné, 2009; Jensen, 1998; Simonton, 2005; Terman, 1926). The research we have presented makes it evident that, while we may possibly come into the world with different aptitudes, our changeable beliefs, goals, skills, and interests meaningfully shape the expression of our emerging intellectual abilities. Given this evidence, it is no longer satisfactory to merely identify levels of intelligence – to test performance at one point in time, label children as gifted or not, or place them into enduring categories. In light of the research, the boundary between gifted and not gifted becomes fluid and fuzzy, something that can change with time and environments. Thus, instead of focusing on measurement and categorization, we are pushed to examine the factors that interfere with and that enhance intellectual accomplishment (e.g., Claxton & Meadows, 2009; Dweck, 2009a, 2009b; Hymer, 2009; Subotnik, 2009).

The research we have reviewed also gives us a different understanding of what it *means* to be "intelligent" or "gifted." Being intelligent or gifted over the long run seems to require not just initial ability but also the right motivation – a focus on learning and not just performance, freedom from stereotypes and belonging concerns, ability to pursue goals in a disciplined manner, and a pursuit of intrinsic goals. As Ericsson and colleagues (1993) noted, even the talented, without hard work and discipline to enhance their skills and address their weaknesses, may lose the giftedness race. Such hard work and accomplishment can be facilitated by environments that help build self-regulatory skills, that pique intrinsic interest, and that draw the focus to learning.

In conclusion, the research we have reviewed can change our understanding of intelligence and bring to light avenues through which motivation can enhance intellectual performance. While we are not arguing that motivation is a substitute for the learning of content and skills, we argue that it is the vehicle through which intellectual knowledge and skills are successfully acquired, expressed, and built on.

References

Andreoletti, C., & Lachman, M. E. (2004). Susceptibility and resilience to memory aging stereotypes: Education matters more than age. *Experimental Aging Research*, *30*(2), 129–148.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38*(2), 113–125.

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, *35*(1), 29–46.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, *7*(1), 1–15.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529.

Baumeister, R. F., Twenge, J. M., & Nuss, C. K. (2002). Effects of social exclusion on cognitive processes: Anticipated aloneness reduces intelligent thought. *Journal of Personality and Social Psychology*, *83*(4), 817–827.

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology*, *41*(2), 174–181.

Binet, A., & Simon, T. (1913). *A method of measuring the development of the intelligence of young children* (trans. C. H. Town). Lincoln, IL: Courier.

Binet, A. (1909/1975). *Modern ideas about children* (trans. S. Heisler). Menlo Park, CA: Suzanne Heisler.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663.

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology*, *40*(2), 247–255.

Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology*, *76* (2), 246–257.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*(3), 273–293.

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology*, *33*(2), 267–285.

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science*, *16*(7), 572–578.

Carr, P. B., & Steele, C. M. (2009). Stereotype threat and inflexible perseverance in problem solving. *Journal of Experimental Social Psychology*, *45*(4), 853–859.

Claxton, G., & Meadows, S. (2009). Brightening up: How children learn to be gifted. In T. Balchin, B. Hymer, & D. J. Matthews (Eds.), *The Routledge international companion to gifted education* (pp. 3–9). New York: Routledge.

Colombo, J., Shaddy, D. J., Blaga, O. M., Anderson, C. J., & Kannass, K. N. (2009). High cognitive ability in infancy and early childhood. In F. D. Horowitz, R. F. Subotnik, & D. Matthews (Eds.), *The development of gift-edness and talent across the lifespan* (pp. 23–42). Washington, DC: American Psychological Association.

Conley, J. J. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion. *Personality and Individual Differences*, *5*(1), 11–25.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology*, *88*(4), 715–730.

Croizet, J., & Claire, T. (1998). Extending the concept of stereotype and threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin*, *24*(6), 588–594.

Cury, F., Da Fonseca, D., Zahn, I., & Elliot, A. (2008). Implicit theories and IQ test performance: A sequential mediational analysis. *Journal of Experimental Social Psychology*, *44*(3), 783–791.

Cury, F., Elliot, A. J., Da Fonseca, D., & Moller, A. C. (2006). The social-cognitive model of achievement motivation and the 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *90*(4), 666–679.

Dai, D. Y., & Sternberg, R. J. (2004). Beyond cognitivism: Toward an integrated understanding of intellectual functioning and development. In D.Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum Associates.

Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science*, *314*(5798), 435–435.

Darwin, C. (1859). On *the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London: John Murray.

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, *28*(12), 1615–1628.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627–668.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*(2), 109–134.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, *318*(5855), 1387–1388.

Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of "executive skills" training. *Developmental Psychobiology*, *36* (2), 161–174.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, *16*(12), 939–944.

Dweck, C. S. (2009a). Foreword. In F. D. Horowitz, R. F. Subotnik, & D. Matthews (Eds.), *The development of giftedness and talent across the life-span* (pp. xi–xiv). Washington, DC: American Psychological Association.

Dweck, C. S. (2009b). Self-theories and giftedness: A reflective conversation. In T. Balchin, B. Hymer, & D. J. Matthews (Eds.), *The Routledge international companion to gifted education* (pp. 308–316). New York: Routledge.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256–273.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, *26*(1), 465–486.

Gagné, F. (2009). Talent development as seen through the differentiated model of talent and giftedness. In T. Balchin, B. Hymer, & D. J. Matthews (Eds.), *The Routledge international companion to gifted education* (pp. 32–41). New York: Routledge.

Galton, F. (1883). Inquiries into human faculty and its development. London: Macmillan.

Galton, F. (1892). *Hereditary genius: An inquiry into its laws and consequences*. London: Macmillan.

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin*, *28*(5), 659–670.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*(6), 645–662.

Gottfried, A. E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology*, *77*(6), 631–645.

Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology*, *82*(3), 525–538.

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology*, *93*(1), 3–13.

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, *85*(3), 541–553.

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology*, *17*(3), 300–312.

Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588–599.

Howse, R. B., Calkins, S. D., Anastopoulos, A. D., Keane, S. P., & Shelton, T. L. (2003). Regulatory contributors to children's kindergarten achievement. *Early Education and Development*, *14*(1), 101–119.

Hymer, B. J. (2009). Beyond compare? Thoughts towards an inclusional, fluid and non-normative understanding of giftedness. In T. Balchin, B. Hymer, & D. J. Matthews (Eds.), *The Routledge international companion to gifted education* (pp. 299–307). New York: Routledge.

Iyengar, S. S., & Lepper, M. R. (1999). Rethinking the value of choice: A cultural perspective on intrinsic motivation. *Journal of Personality and Social Psychology*, *76*(3), 349–366.

Janda, L. (1996). *The psychologists' book of self-tests*. New York: Berkley Books.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Jensen, A. R. (2002). Galton's legacy to research on intelligence. *Journal of Biosocial Science*, *34*, 145–172.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21–46). New York: Oxford University Press.

Krendl, A. C., Richeson, J. A., Kelley, W. M., & Heatherton, T. F. (2008). The negative consequences of threat: A functional magnetic resonance imaging investigation of the neural mechanisms underlying women's underperformance in math. *Psychological Science*, *19*(2), 168–175.

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology*, *97*(2), 184–196.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "over-justification" hypothesis. *Journal of Personality and Social Psychology*, *28*(1), 129–137.

Mangels, J. A., Butterfield, B., Lamb, J., Good, C., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective Neuroscience*, *1*(2), 75–86.

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*(4), 947–959.

Mendoza-Denton, R., Kahn, K., & Chan, W. (2008). Can fixed views of ability boost performance in the context of favorable stereotypes? *Journal of Experimental Social Psychology*, *44*(4), 1187–1193.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*(4907), 933–938.

Moser, J. S., Schroder, H. S., Heeter, C., Moran, T. P., & Lee, Y.-H. (2011). Mind your errors: Evidence for a neural mechanism linking growth mind-set to adaptive posterror adjustments. *Psychological Science*, *22*(12), 1484–1489. https://doi.org/10.1177/0956797611419520

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*(1), 33–52.

Nguyen, H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, *93*(6), 1314–1334.

O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin*, *29*(6), 782–789.

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, *45*(3), 605–619.

Ramsden, S., Richardson, F.M., Josse, G., Thomas, M., . . . & Price, C.J. (2011). Verbal and nonverbal intelligence changes in the teenage brain. *Nature*, *479*(7371), 113–116.

Raven, J. C., Court, J. H., & Raven, J. (1977). *Standard progressive matrices*. London: H. K. Lewis.

Robins, R. W., & Pals, J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity*, *1*(4), 313–336.

Rueda, R. M., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences*, *102*, 14931–14936.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78.

Sansone, C., & Harackiewicz, J. M. (2000). Looking beyond rewards: The problem and promise of intrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 1–13). San Diego: Academic Press.

Sauce, B., & Matzel, L.D. (2018). The paradox of intelligence: Heritability *and* malleability coexist in hidden gene-environment interplay. *Psychological Bulletin*, *144*(1), 26–47. http://dx.doi.org/10.1037/bul0000131

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*(3), 440–452.

Siegler, R. S. (1992). The other Alfred Binet. *Developmental Psychology*, *28*(2), 179–190.

Simonton, D. K. (2005). Giftedness and genetics: The emergenic-epigenetic mode and its implications. *Journal for the Education of the Gifted*, *28*, 270–286.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4–28.

St. Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology*, *59*(4), 745–759.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811.

Sternberg, R. J. (2005). Intelligence, competence and expertise. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 15–30). New York: Guilford Press

Sternberg, R. J., & Grigorenko, E. L. (Eds.). (2001). *Environmental effects on cognitive abilities*. Mahwah, NJ: Lawrence Erlbaum Associates.

Subotnik, R. F. (2009). Developmental transitions in giftedness and talent: Adolescence into adulthood. In F. D. Horowitz, R. F. Subotnik, & D. Matthews (Eds.), *The development of giftedness and talent across the life-span* (pp. 155–170). Washington, DC: American Psychological Association.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*(2), 271–322.

Terman, L. M. (1916). The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale. Boston: Houghton Mifflin.

Terman, L. M. (1926). *Genetic studies of genius*, Vol. 1. Stanford: Stanford University Press.

Terman, L. M., & Oden, M. H. (1959). *Genetic studies of genius: The gifted group at mid-life*, Vol. 5. Oxford: Stanford University Press.

Valiente, C., Lemery-Chalfant, K., Swanson, J., & Reiser, M. (2008). Prediction of children's academic competence from their effortful control, relationships, and classroom participation. *Journal of Educational Psychology*, *100*(1), 67–77.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology*, *87*(2), 246–260.

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, *39*(5), 456–467.

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*(1), 82–96.

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology*, *107*, 468–485.

Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, *20*(9), 1132–1139.

Wechsler, D. (1996) *Eschelle d'Intelligence de Wechsler pour enfants Troisieme Edition (Wechsler Intelligence Scale for Children – Third Edition)*. Paris: ECPA (Original work published 1971).

Wolfe, R. N., & Johnson, S. D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement*, *55*(2), 177–185.

45 Intelligence and Creativity

Jonathan A. Plucker, Maciej Karwowski, and James C. Kaufman

How are intelligence and creativity related? The question is of great interest because, in our schools and tests, we seem to value intelligence over creativity. In life, however, creativity is at least as important because it involves adapting to the novel situations that can lead people to either great success or stunning failure. Sternberg and O'Hara (1999) have argued that the relationship between creativity and intelligence "is theoretically important, and its answer probably affects the lives of countless children and adults" (p. 269).

Their point is well taken: Psychologists and educators frequently address issues related to either creativity *or* intelligence but they often ignore the interplay between the two – or worse, they feel that intelligence and creativity are inversely related. This may explain why research has consistently shown that teachers prefer intelligent students over creative students (e.g., Aljughaiman & Mowrer-Reynolds, 2005; Westby & Dawson, 1995), as though students are unlikely to exhibit evidence of high (or low) levels of both constructs. In addition, the nature of the relationship could help identify aspects of each construct that are ignored in traditional classroom settings.

For example, Wallach and Kogan (1965) suggested that students with high creativity but low intelligence are more disadvantaged in the traditional classroom setting than students with low creativity and low intelligence. If accurate, this observation has considerable implications for how instruction, the curriculum, and assessment are differentiated in classroom settings. Subsequent research has largely supported Wallach and Kogan's observations (e.g., Beghetto, 2006, 2007; Brandau et al., 2007).

Plucker and Renzulli (1999) concluded that it is a matter of uncovering not *whether* but *how* the two are related. Certainly, creativity has been an important part of many major theories of intelligence. For example, divergent thinking was an integral part of Guilford's (1967) Structure of Intellect (SOI) model. But, in general, the research on this topic is murky if not seemingly in outright conflict. As an example of research and theories that seem to contradict each other, the threshold theory suggests that intelligence is a necessary but not a sufficient condition of creativity (Barron, 1969; Yamamoto, 1964), certification theory proposes that there are environmental factors that allow people to display both creativity and intelligence (Hayes, 1989), and the interference hypothesis suggests that very high levels of intelligence may interfere with creativity (Simonton, 1994; Sternberg, 1996). The threshold theory has gotten the most empirical attention; we will go into some detail about the evidence in the section "Empirical Work on Intelligence and Creativity." The lack of clear conclusions about the nature of creativity-intelligence relationships is due, at least in part, to the dynamic, yet at times underdeveloped, constructs being studied. After all, we should not be surprised if conflicting results are observed when a notoriously ill-defined, complex construct (Plucker, Beghetto, & Dow, 2004), measured similarly for decades (J. C. Kaufman, Plucker, & Baer, 2008), is compared to another complex construct that has seen rapid theoretical and psychometric development (A. S. Kaufman, 2009). Researchers have often been aiming at two moving targets at the same time.

From an assessment perspective, the relationship of creativity to intelligence is of particular interest. First, the overlap (or lack thereof) between intelligence and creativity is an issue enduringly popular, controversial, and heavily dependent on psychometric issues. Second, creativity plays a major role in several theories of giftedness, and school districts struggle with the development of systems to identify gifted students, especially those with above-average creative abilities.

Roots of Creativity

The roots of creativity as a scientific discipline are planted in the intelligence literature. Many of the earlier scholars (such as Francis Galton, Lewis Terman, Alfred Binet, and Charles Spearman) who considered and discussed creativity were more primarily focused on intelligence. Indeed, it was an intelligence researcher, J. P. Guilford, who first publicly recognized the need for an independent study of creativity.

Guilford (1950, 1967) placed creativity into a larger framework of intelligence in his SOI model. He attempted to organize all of human cognition along three dimensions. The first dimension was called "operations" and simply meant the mental processes needed to complete almost any kind of task, such as cognition. The second dimension, "content," referred to the general subject matter, such as words. The third dimension, "product," represented the actual products that might result from different kinds of thinking in different kinds of subject matters, such as writing. With five operations, four contents, and six products, Guilford's (1967) model had 120 different possible mental abilities. Indeed, he later expanded the model to include 180 different abilities (Guilford, 1988), although the 120 abilities model is the one more often studied. This model was influential in educational circles (Meeker, 1969), and Renzulli (1973) developed an entire creativity curriculum based on the aspects of the SOI model involving divergent thinking.

One of Guilford's operations (or thought processes) was divergent thinking – analyzing one's response to questions with no obvious, singular answer. Such questions might include "What would happen if we didn't need sleep?" This work, followed up by other researchers (most notably Torrance, 1974, 2008), has often served as the basis for measures of creativity. Two of the most common ways of scoring these tests are fluency (the total number of responses given) and originality (how statistically rare are the responses).

A Framework for Exploring the Research

Sternberg (1999) has provided a framework for examining the research on this topic. We find this framework to be helpful because it emphasizes that one's conclusions about the creativity-intelligence relationship will largely be determined by one's theoretical conceptualization of each construct. The Sternberg framework includes five possible intelligence-creativity relationships: creativity as a subset of intelligence; intelligence as a subset of creativity; creativity and intelligence as overlapping sets; creativity and intelligence as coincident sets; and creativity and intelligence as disjoint sets. In the following sections, we provide examples of the two primary categories: those that portray creativity as a subset of intelligence and those that portray intelligence as a subset of creativity.

Theories of Intelligence Which Encompass Creativity

As already discussed, Guilford placed creativity within the context of an intellectual framework. In doing so, he was the first of many to consider creativity to be part of intelligence. Some theories of intelligence include creativity as a subcomponent. Undoubtedly, the theory of intelligence that is most often applied to IQ tests is the Cattell-Horn-Carroll (CHC) theory, a combination of two earlier theories. The Cattell-Horn theory (e.g., Horn & Cattell, 1966) initially proposed two types of intelligence, crystallized (G*c*) and fluid (G*f*). G*c* signifies what a person knows and has learned and G*f* represents how a person handles a new and different situation (i.e., problem-solving). Horn expanded the theory to include more dimensions (known as Broad Abilities). Carroll's (1993) theory proposed a hierarchy of intellectual abilities. At the top of the hierarchy is general ability; in the middle of the hierarchy are various broad abilities (including learning and memory processes and the effortless production of many ideas). At the bottom of the hierarchy are many narrow, specific abilities such as spelling ability and reasoning speed.

The combined CHC theory incorporates both the concept of a general intelligence (all of the different aspects of intelligence are considered to be related to a common "*g*," although this aspect is not often emphasized; see Flanagan & Ortiz, 2002) and the concept of many different aspects of intelligence. Ten different broad factors of intelligence are proposed. These include G*f* and G*c* from the initial Cattell-Horn theory. They also include G*q* (quantitative knowledge, typically math-related), G*rw* (reading and writing), G*sm* (short-term memory), G*v* (visual processing), G*a* (auditory processing), G*lr* (long-term storage and retrieval), G*s* (processing speed), and G*t* (decision speed/reaction time). Of these ten, only seven are directly measured by today's intelligence tests: G*q* and G*rw* are in the domain of academic achievement, and, therefore, are measured by achievement tests, and G*t* is not measured by any major standardized test. Intelligence tests may indirectly measure some of these other skills, however. In addition, some of the components of each broad factor may not be well measured by either ability or achievement tests.

Today, nearly every major intelligence test is founded either explicitly or implicitly on the current version of the theory, namely CHC. In addition, largely because of the influence of CHC theory, all current IQ tests (including the Wechsler Intelligence Scale for Children – Fifth Edition; WISC-V, Wechsler, 2014) have shifted the historical focus from a small number of part scores to a contemporary emphasis on anywhere from four to seven cognitive abilities (Sternberg, Kaufman, & Grigorenko, 2008).

Although in the early stages of the Cattell-Horn G*f*-G*c* theory, G*f* (fluid intelligence) was hypothesized to be strongly linked to creativity (Cattell & Butcher, 1968), such a relationship is no longer explicitly part of the CHC theory. The current model (McGrew, 2009; Schneider & McGrew, 2012) includes originality/creativity as a component of long-term storage and retrieval (G*lr*): "Some G_{lr} narrow abilities have been prominent in creativity research (e.g., production, ideational fluency, or associative fluency)" (McGrew, 2009, p. 6). In the detailed description of the model, this sentence is the only mention of creativity, originality, or divergent thinking. Fluid intelligence (G*f*) is discussed in terms of its relationship to problem-solving and coping with novel problems (both considered to be highly related to creativity), yet the emphasis is on G*lr*. Some studies have found that G*lr* is significantly associated with rated creativity (Avitia & Kaufman, 2014) and divergent thinking (Silvia, Beaty, & Nusbaum, 2013). There has been recent discussion of splitting G*lr* into two distinct abilities, G*l* and G*r*. With this split, the creativity-related components would be included with G*r* (Schneider & McGrew, 2018). Martindale (1999) proposed a differential relationship between G*s* (processing speed) and creativity. According to Martindale's theory, people who are creative are selective with their speed of information processing. Early in the creative problem-solving stage, they widen their breadth of attention, allowing a larger amount of information to be processed (and thereby lowering their speediness). Later, when the problem is better understood, their attention span is shortened and their reaction time is quicker. This theory is reminiscent of Sternberg's (1981) distinction between global and local planning: Brighter people spend more time in initial global planning so that later they do not have to spend as much time in local planning.

Some have argued that the current CHC model shortchanges creativity (J. C. Kaufman, 2015, 2016). Placing all references to creativity and originality under G*lr* (or G*r*) seems quite narrow. The ability to draw selectively on past experiences is essential for creating something new. But the connection between fluid intelligence and creativity is minimized in new conceptions of the model.

An intriguing perspective in this category is Sternberg's (1996, 1999; Sternberg et al., 2008) theory of successful intelligence. This theory comprises three "subtheories": a *componential subtheory*, which relates intelligence to the internal world of the individual; an *experiential subtheory*, which relates intelligence to both the external and the internal worlds of the individual; and a *contextual subtheory*, which relates intelligence to the external world of the individual. The componential subtheory specifies the mental mechanisms responsible for planning, carrying out, and evaluating intelligent behavior. The experiential subtheory expands on this definition by focusing on those important behaviors that involve either adjustment to relative novelty, automatization of information processing, or both. The contextual subtheory defines intelligent behavior as involving purposeful adaptation to, selection of, and shaping of real-world environments relevant to one's life (Sternberg et al., 2008).

The experiential subtheory is directly related to creativity. Sternberg's application of creativity assessments to admissions data increased prediction of college success beyond that obtained with standard admissions tests; in addition, ethnic-group differences were significantly reduced (Sternberg, 2006, 2010; Sternberg & The Rainbow Project Collaborators, 2006). Gardner's well-known theory of multiple intelligences (1999) does not specifically address creativity. However, his eight intelligences (interpersonal, intrapersonal, spatial, naturalistic, linguistic, logical-mathematical, bodily-kinesthetic, and musical) certainly seem to apply to creativity. Gardner (1993) used case studies of eminent creative individuals to argue that creative people can shine as a function of embodying different intelligences. For example, he selected Freud as an example of intrapersonal intelligence; Einstein to represent logical-mathematical intelligence; Picasso, spatial intelligence; Stravinsky, musical intelligence; T. S. Eliot, linguistic intelligence; Martha Graham, bodily-kinesthetic intelligence; and Gandhi, interpersonal intelligence (naturalistic intelligence had not been added at this time).

Theories of Creativity That Encompass Intelligence

Systems Theories

In recent years, there has been an emphasis on creativity theories that incorporate factors that are interrelated (Kozbelt, Beghetto, & Runco, 2010). Some of these theories emphasize issues such as the environment or evolution and are less relevant here because intelligence plays a more tangential role. Other theories emphasize a confluence of different elements and include intellectual and cognitive abilities in the equation. One such theory is Sternberg and Lubart's (1996) "investment" theory of creativity, in which the key to being creative is to buy low and sell high in the world of ideas. In this model, a creative person is like a talented Wall Street investor. A successful creator will generate ideas that may be initially unpopular or underappreciated (as in buying stocks with low price-earnings ratios) yet will persist and convince others of the ideas' merits. The creator will then know when to move on to pursue other ideas (as in selling high, when one divests oneself of stocks). The concept of defying the crowd was expanded in Sternberg's (2018) triangular theory of creativity, in which people can also defy themselves or the Zeitgeist. Someone who defies all three can be considered to have reached consummate creativity.

According to this model, six main elements contribute to creativity: intelligence, knowledge, thinking styles, personality, motivation, and the environment. Intelligence contributes using three elements drawn from Sternberg's triarchic theory (1988, 1996; later expanded into the theory of successful intelligence).

The first element is synthetic ability, which is the ability to generate ideas that are novel, high in quality, and high in task appropriateness. Because creativity is viewed as an interaction between a person, a task, and an environment, what is novel, high in quality, or task appropriate may vary from one person, task, or environment to another. Central to this ability is being able to redefine problems. Creative people may take problems that other people see, or they themselves may previously have seen, in one way and redefine the problems in a different way. This synthetic ability includes three knowledge-acquisition components. The first, selective encoding, involves distinguishing relevant from irrelevant information. Selective combination, the second, involves combining bits of relevant information in novel ways. Finally, selective comparison involves relating new information to old information in a novel way.

The second element, practical ability, is needed to communicate creative ideas to other people (i.e., "selling" an idea). Good ideas do not always sell themselves – the creative person needs to devise strategies for and expend effort in selling those ideas (see also Plucker, 2016).

The third component, analytical ability, is often measured by traditional intelligence tests. Yet this component is also related to creativity, as a successful creator must be able to judge the value of their own ideas and decide which ones to pursue. Such analytical ability can be used to evaluate the strengths and weaknesses of the idea and determine the best steps to improve on the idea. People who are high in synthetic ability but low in analytical ability may need someone else to evaluate and judge their work for them. People who are able incisively to evaluate their own work may be said to be high in metacognition.

The role of metacognition during the creative process has long been of interest to researchers and evidence suggests that metacognition plays a role in creative cognition (Feldhusen, 1995; Feldhusen & Goh, 1995; Pesut, 1990; Swanson, 1992). For example, Runco and colleagues (Runco & Dow, 2004; Runco & Smith, 1992) found that people who tended to produce more original responses also were better at rating their most original responses to a divergent thinking task. Silvia (2008a) asked people to pick their best responses to a similar divergent thinking task and then examined whether they were more likely to choose responses that outside raters considered creative. Silvia found that people were able to discern their more creative responses – and that people who were more open to experience were more likely to choose accurately. Researching the extremely creative end of the spectrum, Kozbelt (2007) analyzed Beethoven's self-critiques and found that the great composer was a reasonably accurate rater of his own work. Kaufman and Beghetto (2013) have suggested that the intersection of creativity and metacognition, or creative metacognition, is not only being aware of one's strengths and weaknesses but also knowing when to express one's creativity.

Another theory that views creativity as a mix of different abilities is Amabile's (1996; Amabile & Pratt, 2016) componential model of creativity. She argued that three variables were needed for creativity to occur: domain-relevant skills, creativity-relevant skills, and task motivation. Domain-relevant skills include knowledge, technical skills, and specialized talent (i.e., a creative mathematician should know basic algebra and geometry). Creativity-relevant skills are personal factors that are associated with creativity. These skills include tolerance for ambiguity, self-discipline, and risk-taking. Finally, Amabile singles out one's motivation toward the task at hand. Intelligence would primarily occur at the domain-relevant skill level.

A third theory that accounts for multiple variables and also takes a domain-specific approach is the Amusement Park theory (Baer & Kaufman, 2017; J. C. Kaufman & Baer, 2005). In an amusement park, there are *initial requirements* (e.g., a ticket) that apply to all areas of the park. Similarly, there are initial requirements that, to varying degrees, are necessary to creative performance in all domains. One such key initial requirement is intelligence. Amusement parks also have *general thematic areas* (e.g., at Disney World one might select among EPCOT or Disney-MGM Studios), just as there are several different general areas in which someone could be creative (e.g., the arts, science). Once in one type of park, there are sections (e.g., Fantasyland and Adventureland are all found in the Magic Kingdom), just as there are *domains* of creativity within larger *general thematic areas* (e.g., physics and biology are domains in the *general thematic area* of science). These domains in turn can be subdivided into *micro-domains* (e.g., in Fantasy-land one might visit Cinderella's Castle or It's a Small World; in the domain of psychology, one might specialize in cognitive psychology or social psychology).

Cognitive Theories of Creativity

The other group of theories that includes intellectual abilities as a key component is the set of cognitive theories of creativity. Guilford pioneered these ideas and his convergent versus divergent thinking dichotomy is still a key idea in creativity. Even before Guilford, however, Wallas (1926) proposed a model of the cognitive creative process. According to his five-stage model, you first use *preparation* to begin work on a problem. Next, there is *incubation*, in which you may work on other things while your mind thinks about the problem. In *intimation*, you realize you are about to have a breakthrough (this phase is sometimes dropped from the model) and then you actually have the insight in the *illumination* phase. Finally, with *verification*, you actually test, develop, and use your ideas.

Building off of these ideas, the Geneplore model has two phases, generative and explorative, that are comparable to Guilford's convergent and divergent thinking distinction. In the generative phase, someone constructs a preinventive structure, or a mental representation of a possible creative solution (Finke, Ward, & Smith, 1992). For example, Elias Howe was working on his invention of the modern sewing machine. He could not quite get the needle correctly designed. Howe had an odd dream in which he was chased by savages who threw spears at him. The spears had a circle loop at the end – and Howe realized that adding the circle (or an "eye") to the end of the needle was the solution he needed (Hartman, 2000). The image of a spear with a circle at the end – the image that preceded Howe's insight – would be an example of one of these preinventive structures. They do not need to be as dramatic or sudden as the realization based on Howe's dream. Indeed, the generation of preinventive structures is only one part of the creative process, according to the Geneplore model. The thinker must then explore these different preinventive structures within the constraints of the final goal. There may be several cycles before a creative work is produced.

Although the model focuses on the creative process, most tests of the model have actually measured the creative product. In an experiment testing the model, people were shown parts of objects (such as a circle or a cube). They were then asked to combine these parts together to produce a practical object or device. The creativity (and practicality) of the items was then assessed (e.g., Finke, 1990; Finke & Slayton, 1988). Interestingly, people produced more creative objects when they were told which parts had to be combined than when they could pick the parts to be combined.

Other theories have also focused on cognitive-oriented components of the creative process. Michael Mumford and his colleagues (Blair & Mumford, 2007; Mumford, Longergan, & Scott, 2002; Mumford et al., 1991) have argued for an eight-part model, focusing on problem construction, information encoding, category selection, category combination and reorganization, idea generation, idea evaluation, implementation planning, and solution monitoring. Reiter-Palmon and Robinson (2009) have specifically emphasized the value of problem construction – the ability to identify precisely which problem needs to be solved.

Mednick (1962, 1968) proposed the idea that creativity occurs when different elements are associated together to form new combinations. Creative individuals are assumed to be able to make meaningful, useful associations between disparate concepts and ideas to a greater extent than a relatively uncreative individual. The Remote Associates Test was developed based on this idea (Mednick & Mednick, 1967). The ideas behind associational theory are still being explored; for example, Kenett, Anaki, and Faust (2014) found that less creative people had more rigid semantic memory networks compared to more creative people.

In general, theories of creativity that include intelligence take one of the two approaches outlined above. Componential theories encompass multiple constructs that are a key component of creativity and they include intelligence as one such variable. Cognitive theories view creativity as occurring primarily in one's mind; as such, the processes used to be creative are comparable to those used to be intelligent.

Theories on How Intelligence and Creativity Are Related

The threshold theory argues that intelligence is a necessary but not a sufficient condition of creativity (Barron, 1969). According to this view, creativity and intelligence are positively correlated up until an IQ of approximately 120; in people with higher IQs, the two constructs are said to show little relationship (e.g., Barron, 1963; Getzels & Jackson, 1962). The interference hypothesis suggests that very high levels of intelligence may interfere with creativity (Simonton, 1994; Sternberg, 1996).

Empirical Work on Intelligence and Creativity

Most studies that investigate creativity and intelligence use divergent-thinking tests (such as the TTCT) or other related paper-and-pencil tests also scored for fluency, originality, or other divergent thinking-related methods of scoring (e.g., Plucker, 1999). The studies have generally found that levels of creativity are significantly associated with scores on psychometric measures of intelligence (especially verbally oriented measures, regardless of the type of creativity measured). This relationship is typically not a particularly strong one (Barron & Harrington, 1981; Kim, 2005; Wallach & Kogan, 1965), although Silvia (2008a, 2008b) argued that the relationship between the latent constructs of creativity and intelligence is underestimated because the analyses only look at observable scores (i.e., performance on an intelligence test). If it were possible to get a "true" measure of the constructs, there might be a higher relationship. Indeed, studies utilizing latent variables approach rather than relying on raw scores in creativity and intelligence tests tend to find higher latent correlations between these constructs, ranging from $r = 0.34$ (Benedek et al., 2014) to $r = 0.46$ (Karwowski, Czerwonka, & Kaufman, 2018) and $r = 0.51$ (Benedek et al., 2012). Thus, the data analysis approach does matter.[1]

Most of these studies reinforce the threshold theory (e.g., Fuchs-Beauchamp, Karnes, & Johnson, 1993; Getzels & Jackson, 1962) but the threshold theory has come under fire. Kim (2005), in a meta-analysis of twenty-one studies, found virtually no support for the threshold theory, with small positive correlations found at all levels of ability between several different measures of intelligence and creativity. Runco and Albert (1986) found that the nature of the relationship was dependent on the measures used and the populations tested. Preckel, Holling, and Wiese (2006) looked at measures of G*f* and creativity (as measured through divergent thinking tests) and found modest correlations across all levels of intellectual abilities. Wai, Lubinski, and Benbow (2005), in a longitudinal study of gifted (top 1 percent) thirteen-year-olds, found that differences in SAT scores – even within such an elite group – predicted creative accomplishments twenty years later. Park, Lubinski, and Benbow (2007) examined intellectual patterns of ability and eventual creativity in different domains. Using math and verbal SAT scores of people at age thirteen, they then tracked the accomplishments of these same people twenty-five years later. Unsurprisingly, early prowess was associated with eventual success. However, a person's specific strengths (in this case, math vs. verbal) predicted patents (math) and literary publications (verbal). Park, Lubinski, and Benbow (2008) further extended their findings to demonstrate this link in the fields of science and technology. It is important to note that family background variables (most notably socioeconomic status) may be responsible for some of these connections, given they were not controlled for in these studies.

Thus, after more than fifty years of interest in the threshold hypothesis, the evidence is equivocal. As Karwowski and Gralewski (2013) note, unsystematic findings may be caused by several unobvious decisions and pre-assumptions scholars make. First, previous tests of the threshold hypothesis tended to utilize different decisive criteria when assessing whether threshold theory is or is not supported. The most liberal, yet clearly wrong, criterion assumes that there is a significant correlation between creativity and intelligence below the threshold and a lack of significant

[1] Although we note that Plucker (1999), using a latent variable approach, found $r = 0.20$.

correlations above it. Obviously, however, such a pattern cannot convincingly demonstrate the threshold hypothesis, as statistical significance alone depends on the sample/s size, while samples above the threshold are usually much lower than those below it. Further, there is almost always a much smaller variance in intelligence scores above the threshold (restriction range) that artificially lowers observed correlations and – consequently – their standard errors. Hence, comparing the statistical significance of correlations below and above the hypothesized threshold cannot count as a proper test of the threshold theory.

The second, more conservative test of the threshold theory is an extension of the previous one (Karwowski & Gralewski, 2013). It entails demonstrating not only that one coefficient is significant and another is not but also that these two coefficients differ significantly in their strength. It leads to a third, the most general formal test of the threshold theory that assumes that the correlation below the threshold is statistically significant and, at the same time, significantly stronger than the correlation above the threshold. In this sense, it is not necessary that the correlation above the threshold is not higher than zero – according to the threshold theory, it could still be positive yet weaker than the correlation below the threshold. Importantly, the above-mentioned inconsistencies in testing the threshold hypothesis are not the only nuances that greatly influence researchers' conclusions. Another important point relates to the data analysis and estimation methods. If researchers utilize raw scores in intelligence and creativity tests, they implicitly assume that the reliability of their measurement is equal at all levels of the latent traits they measure. In other words, there is a hidden assumption that the precision of intelligence and creativity measurement is the same at very low, medium, or high levels of these characteristics. This assumption is based in one of the classic test theory premises, although more recent item response theory works (see Karwowski & Gralewski, 2013) clearly demonstrate that the reliability is curvilinearly related to the level of the latent trait. More specifically, at the ends of the continuum of the trait being measured, the quality of measurement is lower than in the case of the average level of the latent trait. Consequently, the conclusions regarding the links between intelligence and creativity above the hypothesized threshold may be biased (Karwowski & Gralewski, 2013).

It seems safe to conclude that the proper tests of the threshold hypothesis may require more sophisticated statistical approaches than simple comparisons of correlations below and above the threshold. We present examples of two such approaches. The first utilized segmented regression analysis. Unlike previous tests, segmented regression does not make any assumptions regarding the exact point of the threshold but instead estimates it based on the dataset. This is important, as the threshold of IQ of 120 initially postulated in the literature (Getzels & Jackson, 1962) was never sufficiently justified. Jauk and his colleagues (2013) used segmented regression to look for potentially different threshold in the relationship between intelligence and various aspects of creative thinking (fluency and originality) and creative achievement.

Interestingly, they indeed found different threshold values for different aspects of creative thinking. For fluency, the breakpoint was established at an IQ of 86, with the correlation below being not only significant and robust ($r = 0.56$) but also significantly stronger than the relationship above this point (r = 0.09, *ns*). For originality, the threshold obtained depended on the scoring method. If the researchers relied on a simpler yet less sensitive originality scoring based on the best two ideas as indicated by participants (so-called top two originality score; see Silvia et al., 2008), the threshold was established at an IQ of 104 points (correlation below $r = 0.38$, above $r = 0.14$; the difference between correlations was significant). When originality was scored using the more traditional way, by scoring all of the answers provided, the threshold was found at almost exactly 120 IQ points level (specifically 119.6 IQ points) with significant correlation below this threshold ($r = 0.35$), lack of correlation above it ($r = -0.01$), and a significant difference between these two. Interestingly, Jauk and colleagues did not find a threshold-like relationship between intelligence and creative achievement. Although these two characteristics were robustly correlated ($r = 0.28$), no nonlinear associations were found.

Segmented regression is more robust in examining the threshold hypothesis than looking at simple correlations below and above the threshold. Further, given that the threshold itself is identified, this method allows testing whether the classic hypothesis of the threshold being at an IQ of 120 finds confirmation in the current studies. Jauk and colleagues' study did suggest that there might be different thresholds for different creativity aspects. Whereas only basic intelligence level is required for ideational fluency (almost exactly one standard deviation below the average), the hypothesized 120 IQ threshold was observed for originality. Other studies that aimed at replicating these findings using the same techniques make the picture less clear. One large (N = 4,368) study conducted in Saudi Arabia (Mourgues et al., 2016) identified very mixed patterns of the links between creativity and intelligence that depended greatly on participants' grade level and the aspects of creativity being measured. In general, however, the correlations above the threshold tended to be higher than those below it – contrary to the assumptions of the threshold theory. A more recent Chinese study (Shi et al., 2017) did demonstrate the thresholds at an IQ of 109 for the fluency and flexibility (with correlations between these aspects of creativity and intelligence being significantly higher below the threshold than above it) and a threshold of an IQ of 117 for originality, with the correlation below threshold established at $r = 0.32$ and no links above the threshold ($r = -0.09$).

Another way of more robust tests of possibly nonlinear associations between intelligence and creativity is quantile regression (Yu, Lu, & Stander, 2003). Unlike the routinely used ordinary-least-squares (OLS) regression, which predicts the mean of the outcome, quantile regression predicts any specified quantile of the outcome. Dumas (2018) tested the threshold theory using quantile regression and demonstrated that relational reasoning (a proxy of intelligence) predicted originality but this pattern was clearly nonlinear – much stronger links were noted for lower quantiles of originality (10th and 30th – $\beta = 0.54$ and $\beta = 0.42$ respectively), smaller yet still significant when median originality was predicted ($\beta = 0.29$) and lack of relationships when above-median originality was predicted. Interestingly, no threshold-like relationships were observed in the case of fluency – relational reasoning was able to predict only the highest level (90th percentile) of fluency but not its lower levels. As quantile regression is based on different assumptions than segmented regression, both these methods could provide complementary insights. Although segmented regression tests for possible nonlinear relationships by *segmenting* the independent variable (in the case of intelligence-creativity links it refers to intelligence), quantile regression rather *segments* the dependent variable and tests what level of creativity, so the dependent variable – low, medium, or high (or any other decile) – is being predicted by intelligence. In that sense, quantile regression might be considered as a special case of so-called Necessary Condition Analysis (Dul, 2016) – a new approach we describe later in this section.

Sligh, Conners, and Roskos-Ewoldsen (2005) delved deeper into the intelligence-creativity relationship by specifically examining the relationship between G*f* (novel problem-solving) and G*c* (acquired knowledge) and a measure of actual creative innovation. G*c* showed the same moderate and positive relationship to creativity as past studies, mentioned previously; in contrast, G*f* showed the opposite pattern. Measured intelligence and creativity were significantly correlated for the high IQ group but they were not significantly correlated for people with average IQs. This finding implies that students who receive high G*f* scores may be more likely to be creative than students who receive high G*c* scores.

The Sligh and colleagues study also addresses a second major weakness in this line of research: the overreliance on divergent thinking measures as the sole assessment of creativity. Few studies have been conducted that include measures of creative personality, creative products, and creative processes (other than divergent thinking).

An interesting suggestion posed by Batey and Furnham (2006) is that the role of G*f* and G*c* in creativity may shift across the life span of a creative person. G*f*, they argue, might be more important in early stages of a career. Conversely, a later-career creator may rely more on G*c* – and, we might postulate, G*lr* (or G*r*). Batey and Furnham's (2006) hypotheses are yet to be tested but it seems likely that the role of different aspects of intelligence might vary for creativity analyzed at different levels of eminence (J. C. Kaufman & Beghetto, 2009) or across different domains (J. C. Kaufman, Glăveanu, & Baer, 2017; S. B. Kaufman et al., 2015). Is the role that intelligence plays for creative achievement the same as for divergent thinking? Is intelligence able to predict the likelihood of creative achievement?

A longitudinal study (Karwowski et al., 2017) conducted over a forty-year time span attempted to explore these questions. The intelligence of a large sample of Polish eleven-year-old children ($N = 1594$) was measured using Raven's Matrices at age eleven and a small percentage ($N = 255$) were subsequently individually tested using the Wechsler Intelligence Scale for Children (WISC) at age thirteen. Forty years later, at age fifty-two, participants' creative achievement was measured using the Creative Achievement Questionnaire (Carson, Peterson, & Higgins, 2005). Although the correlational links between intelligence measured in childhood and later creative achievement were negligible (similar to the results in Plucker's (1999) reanalysis of Torrance's longitudinal study), several interesting patterns were observed. First, intelligence in childhood (both at age eleven and thirteen) was nonlinearly related with creative achievements in more cognitively demanding domains (architecture, writing, humor, inventions, and science). In addition, an analysis of the WISC data with segmented regression demonstrated a breakpoint at an IQ of exactly 120, with the Raven's Matrices only slightly lower (114 IQ points). Second, an inspection of scatter plots, accompanied by a more formal analysis, demonstrated that, although the links were far from linear, it was highly unlikely to expect creative achievements among these participants whose intelligence in childhood was low. Therefore, although intelligence was by no means a sufficient condition for creative achievements, low intelligence made such achievements unlikely. In this sense, intelligence served as a necessary, yet not sufficient, condition.

The formal methodology to test if one variable serves as a necessary (yet not sufficient) condition of another was proposed by Dul (2016). Necessary Condition Analysis (or NCA) is a new statistical method allowing researchers to adequately test the hypothesis that one variable serves as a necessary (albeit not sufficient) condition of another. Unlike regression-based approaches, NCA quantifies the pattern of relationships between variables, not assuming linear or curvilinear links between them. Instead, NCA bases on scatter plot shape and estimates the relative size of the upper-left corner in the scatter plot, that is, the space expected to be empty if one variable is a necessary condition of another. Although recent and still being developed, to date, NCA was already used in a number of fields, including psychology (Karwowski et al., 2016, 2017; Shi et al., 2017), organizational behavior (Arenius, Engel, & Klyver, 2017), or buyer-supplier relationships (Valk et al., 2016). Within psychology, NCA was applied to test the hypothesis that intelligence is necessary for creativity and provided additional support for this claim (Karwowski et al., 2016, 2017; Shi et al., 2017). A recent overview of this method (Dul, Karwowski, & Kaufman, in press) provides a more detailed overview of the NCA, accompanied by a guide for the researchers interesting in using this method.

Given the existing studies, what do all of these results mean? Few studies contradict the idea that creative people tend to be fairly smart and smart people are usually somewhat creative. But some of the tested-and-true ideas about the specific relationship are still unclear. Whereas the threshold theory argues that there may be a certain point at which being smart stops helping creativity, the NCA approach instead sees intelligence as a necessary – yet not sufficient – condition of creative functioning. After all, creativity involves not only generating but also assessing ideas, thus its overlap with intelligence is not surprising. There are reasons to believe that the links between intelligence and creativity are much more complex and qualified by other characteristics. This idea was recently proposed as the "conditioned threshold hypothesis" (Harris, Williamson, & Carter, 2018), which argued that variables such openness to experience play a moderating role. Harris and colleagues, for example, found that the links between intelligence and creative achievement were not only stronger among those high on openness but gained strength with the intelligence level for such people.

Conclusions

Intelligence is strongly valued in schools and extensive and popular measures are often used to measure it. There are usually hundreds of empirical studies of each intelligence test. Creativity may be theoretically desired in school but it is often considered less important than intelligence (Westby & Dawson, 1995). Creativity assessment is murkier than intellectual assessment. The Torrance Tests remain the most-used creativity tests (Forgeard & Kaufman, 2016), despite extensive and multifaceted critiques (Baer, 2011; J. C. Kaufman et al., 2008).

From our perspective, the complexity of possible intelligence-creativity relationships is not surprising. Whenever one compares two constructs, the way in which each construct is conceptualized and assessed will have a significant impact on any empirical results. Researchers and theorists do not believe that intelligence and creativity are completely orthogonal but, even as we understand more about their intersection, the exact nature of that relationship continues to remain an open question. The basic need for both creativity and intelligence, however, remains undisputed.

References

Aljughaiman, A, & Mowrer-Reynolds, E. (2005). Teachers' conceptions of creativity and creative students. *Journal of Creative Behavior*, *39*, 17–34.

Amabile, T. M. (1996). *Creativity in context: Update to "The Social Psychology of Creativity."* Boulder, CO: Westview Press.

Amabile, T. M., & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior*, *36*, 157–183.

Arenius, P., Engel, Y., & Klyver, K. (2017). No particular action needed? A necessary condition analysis of gestation activities and firm emergence. *Journal of Business Venturing Insights*, *8*, 87–92.

Avitia, M. J., & Kaufman, J. C. (2014). Beyond g and c: The relationship of rated creativity to long-term storage and retrieval (Glr). *Psychology of Aesthetics, Creativity, and the Arts*, *8*, 293–302.

Baer, J. (2011). How divergent thinking tests mislead us: Are the Torrance Tests still relevant in the 21st century? *Psychology of Aesthetics, Creativity, and the Arts*, *5*, 309–313.

Baer, J., & Kaufman, J. C. (2017). The Amusement Park Theoretical Model of Creativity: An attempt to bridge the domain specificity/generality gap. In J. C. Kaufman, V. P. Glăveanu, & J. Baer (Eds.), *Cambridge handbook of creativity across domains* (pp. 8–17). New York: Cambridge University Press.

Barron, F. (1963). *Creativity and psychological health*. Princeton: D. Van Nostrand.

Barron, F. (1969). *Creative person and creative process*. New York: Holt, Rinehart & Winston.

Barron, F., & Harrington, D. M. (1981). Creativity, intelligence, and personality. *Annual Review of Psychology*, *32*, 439–476.

Batey, M., & Furnham, A. (2006). Creativity, intelligence and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, *132*, 355–429.

Beghetto, R. A. (2006). Creative justice? The relationship between prospective teachers' prior schooling experiences and perceived importance of promoting student creativity. *Journal of Creative Behavior*, *40*, 149–162.

Beghetto, R. A. (2007). Does creativity have a place in classroom discussions? Prospective teachers' response preferences. *Thinking Skills and Creativity*, *2*, 1–9.

Benedek, M., Franz, F., Heene, M., & Neubauer, A. C. (2012). Differential effects of cognitive inhibition and intelligence on creativity. *Personality and Individual Differences*, *53*, 480–485.

Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence*, *46*, 73–83.

Blair, C. S., & Mumford, M. D. (2007). Errors in idea evaluation: Preference for the unoriginal? *Journal of Creative Behavior*, *41*, 197–222.

Brandau, H., Daghofer, F., Hollerer, L., Kaschnitz, W., Kirchmair, G., Krammer, I., & Schlagbauer, A. (2007). The relationship between creativity, teacher ratings on behavior, age, and gender in pupils from seven to ten years. *Journal of Creative Behavior*, *41*, 91–113.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal*, *17*, 37–50.

Cattell, R. B., & Butcher, H. (1968). *The prediction of achievement and creativity*. Indianapolis, IN: Bobbs-Merrill.

Dul, J. (2016a). Necessary condition analysis (NCA): Logic and methodology of "necessary but not sufficient" causality. *Organizational Research Methods*, *19*, 10–52.

Dul, J., Karwowski, M., & Kaufman, J. C. (in press). Necessary condition analysis in creativity research. In V. Dörfler & M. Stierand (Eds.), *Handbook of research methods on creativity*. New York: Edward Elgar.

Dumas, D. (2018). Relational reasoning and divergent thinking: An examination of the threshold hypothesis with quantile regression. *Contemporary Educational Psychology*, *53*, 1–14.

Feldhusen, J. F. (1995). Creativity: A knowledge base, metacognitive skills, and personality factors. *Journal of Creative Behavior*, *29*, 255–268.

Feldhusen, J. F., & Goh, B. E. (1995). Assessing and accessing creativity: An integrative review of theory, research, and development. *Creativity Research Journal*, *8*(3), 231–247.

Finke, R. (1990). *Creative imagery: Discoveries and inventions in visualization*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Finke, R. A., & Slayton, K. (1988). Explorations of creative visual synthesis in mental imagery. *Memory and Cognition*, *16*, 252–257.

Finke, R. A., Ward T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA: MIT Press.

Flanagan, D. P., & Ortiz, S. O. (2002). Best practices in intellectual assessment: Future directions. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (Vol. 4, pp. 1351–1372). Washington, DC: National Association of School Psychologists.

Forgeard, M. J. C., & Kaufman, J. C. (2016). Who cares about imagination, creativity, and innovation, and why? A review. *Psychology of Aesthetics, Creativity, and the Arts*, 10, 250–269.

Fuchs-Beauchamp, K. D., Karnes, M. B., & Johnson, L. J. (1993). Creativity and intelligence in preschoolers. *Gifted Child Quarterly*, *37*, 113–117.

Gardner, H. (1993). *Creating minds*. New York: Basic Books.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century.* New York: Basic Books.

Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence: Explorations with gifted students*. New York: Wiley.

Guilford, J. P. (1950). Creativity. *American Psychologist*, *5*, 444–454.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Guilford, J. P. (1988). Some changes in the Structure-of-Intellect Model. *Educational and Psychological Measurements*, *48*, 1–4.

Harris, A. M., Williamson, R. L., & Carter, N. T. (2018). A conditional threshold hypothesis for creative achievement: On the interaction between intelligence and openness. *Psychology of Aesthetics, Creativity, and the Arts*, Advance online publication. http://dx.doi.org/10.1037/aca0000182

Hartman, E. (2000). *Dreams and nightmares: The origin and meaning of dreams*. New York: Perseus.

Hayes, J. R. (1989). Cognitive processes in creativity. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 135–145). New York: Plenum Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, *57*, 253–270.

Jauk, E., Benedek, M., Dunst, B., & Neubauer, A. C. (2013). The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. *Intelligence*, *41*, 212–221.

Karwowski, M., Czerwonka, M., & Kaufman, J. C. (2018). Does intelligence strengthen creative metacognition? *Psychology of Aesthetics, Creativity, and the Arts*, (Advance online publication). http://dx.doi.org/10.1037/aca0000208

Karwowski, M., Dul, J., Gralewski, J., Jauk, E., Jankowska, D. M., Gajda, A., Chruszczewski, M. H., & Benedek, M. (2016). Is creativity without intelligence possible? A Necessary Condition Analysis. *Intelligence*, *57*, 105–117.

Karwowski, M., & Gralewski, J. (2013). Threshold hypothesis: Fact or artifact? *Thinking Skills and Creativity*, *8*, 25–33.

Karwowski, M., Kaufman, J. C., Lebuda, I., Szumski, G., & Firkowska-Mankiewicz, A. (2017). Intelligence in childhood and creative achievements in middle-age: The necessary condition approach. *Intelligence*, *64*, 36–44.

Kaufman, A. S. (2009). *IQ Testing 101*. New York: Springer.

Kaufman, J. C. (2015). Why creativity isn't in IQ tests, why it matters, and why it won't change anytime soon . . .. Probably. *Journal of Intelligence 3*, 59–72.

Kaufman, J. C. (2016). *Creativity 101*, 2nd ed. New York: Springer.

Kaufman, J. C., & Baer, J. (2005). The amusement park theory of creativity. In J. C. Kaufman & J. Baer (Eds.), *Creativity across domains: Faces of the muse* (pp. 321–328). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, *13*, 1–12.

Kaufman, J. C., & Beghetto, R. A. (2013). In praise of Clark Kent: Creative metacognition and the importance of teaching kids when (not) to be creative. *Roeper Review*, *35*, 155–165.

Kaufman, J. C., Glăveanu, V. P., & Baer, J. (Eds.) (2017). *The Cambridge handbook of creativity across domains*. New York: Cambridge University Press.

Kaufman, J. C., Plucker, J. A., & Baer, J. (2008). *Essentials of creativity assessment*. New York: Wiley.

Kaufman, S. B., Quilty, L. C., Grazioplene, R. G., Hirsh, J. B., Gray, J. R., Peterson, J. B., & DeYoung, C. G. (2015). Openness to experience and intellect differentially predict creative achievement in the arts and sciences. *Journal of Personality*, *82*, 248–258.

Kenett, Y. N., Anaki, D., & Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, *8*, 407.

Kim, K. H. (2005). Can only intelligent people be creative? *Journal of Secondary Gifted Education*, *16*, 57–66.

Kozbelt, A. (2007). A quantitative analysis of Beethoven as self-critic: Implications for psychological theories of musical creativity. *Psychology of Music*, *35*, 147–172.

Kozbelt, A., Beghetto, R. A., & Runco, M. A. (2010). Theories of creativity. In J. C. Kaufman & R. J. Sternberg (Eds.), *Cambridge handbook of creativity* (pp. 20–47). New York: Cambridge University Press.

Martindale, C. (1999). Biological bases of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 137–152). New York: Cambridge University Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review*, *69*, 220–232.

Mednick, S. A. (1968). The Remote Associates Test. *Journal of Creative Behavior*, *2*, 213–214.

Mednick, S. A., & Mednick, M. T. (1967). *Examiner's manual: Remote Associates Test*. Boston, MA: Houghton Mifflin.

Meeker, M. N. (1969). *The structure of intellect: Its interpretation and uses*. Columbus, OH: Merrill.

Mourgues, C. V., Tan, M., Hein, S., Al-Harbi, K., Aljughaiman, A., & Grigorenko, E. L. (2016). The relationship between analytical and creative cognitive skills from middle childhood to adolescence: Testing the threshold theory in the Kingdom of Saudi Arabia. *Learning and Individual Differences*, *52*, 137–147.

Mumford, M. D., Lonergan, D. C., & Scott, G. M. (2002). Evaluating creative ideas: Processes, standards, and context. *Inquiry: Critical Thinking Across the Disciplines*, *22*, 21–30.

Mumford, M. D., Mobley, M. I., Uhlman, C. E., Reiter-Palmon, R., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal*, *4*, 91–122.

Park, G., Lubinski, D., & Benbow, C. P. (2007). Contrasting intellectual patterns predict creativity in the arts and sciences. *Psychological Science*, *18*, 948–952.

Park, G., Lubinski, D., & Benbow, C. P. (2008). Ability differences among people who have commensurate degrees matter for scientific creativity. *Psychological Science*, *19*, 957–961.

Pesut, D. J. (1990). Creative thinking as a self-regulatory metacognitive process: A model for education, training and further research. *Journal of Creative Behavior*, *24*(2), 105–110.

Plucker, J. A. (1999). Is the proof in the pudding? Reanalyses of Torrance's (1958 to present) longitudinal study data. *Creativity Research Journal*, *12*, 103–114.

Plucker, J. A. (2016). Creative articulation. In J. A. Plucker (Ed.), *Creativity and innovation: Theory, research, and practice* (pp. 151–163). Waco, TX: Prufrock Press.

Plucker, J. A., Beghetto, R. A., & Dow, G. (2004). Why isn't creativity more important to educational psychologists? Potential, pitfalls, and future directions in creativity research. *Educational Psychologist*, *39*, 83–96.

Plucker, J. A., & Renzulli, J. S. (1999). Psychometric approaches to the study of human creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 35–60). New York: Cambridge University Press.

Preckel, F., Holling, H., & Wiese, M. (2006). Relationship of intelligence and creativity in gifted and non-gifted students: An investigation of threshold theory. *Personality and Individual Differences*, *40*, 159–170.

Reiter-Palmon, R., & Robinson, E. J. (2009). Problem identification and construction: What do we know, what is the future? *Psychology of Aesthetics, Creativity, and the Arts*, *3*, 43–47.

Renzulli, J. S. (1973). *New directions in creativity*. New York: Harper & Row.

Runco, M. A., & Albert, R. S. (1986). The threshold theory regarding creativity and intelligence: An empirical test with gifted and nongifted children. *Creative Child and Adult Quarterly*, *11*, 212–218.

Runco, M. A., & Dow, G. T. (2004). Assessing the accuracy of judgments of originality on three divergent thinking tests. *Korean Journal of Thinking and Problem Solving*, *14*, 5–14.

Runco, M. A., & Smith, W. R. (1992). Interpersonal and intrapersonal evaluations of creative ideas. *Personality and Individual Differences*, *13*, 295–302.

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99–144). New York: Guilford Press.

Schneider, W. J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 99–144). New York: Guilford Press.

Shi, B., Wang, L., Yang, J., Zhang, M., & Xu, L. (2017). Relationship between divergent thinking and intelligence: An empirical study of the threshold hypothesis with Chinese children. *Frontiers in Psychology*, *8*, 254.

Silvia, P. J. (2008a). Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual Differences*, *44*, 1012–1021.

Silvia, P. J. (2008b). Creativity and intelligence revisited: A latent variable analysis of Wallach and Kogan (1965). *Creativity Research Journal*, 20, 34–39.

Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2013). Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. *Intelligence*, *41*, 328–340.

Simonton, D. K. (1994). *Greatness: Who makes history and why*. New York: Guilford Press.

Sligh, A. C., Conners, F. A., & Roskos-Ewoldsen, B. (2005). Relation of creativity to fluid and crystallized intelligence. *Journal of Creative Behavior*, *39*, 123–136.

Sternberg, R. J. (1981). Intelligence and nonentrenchment. *Journal of Educational Psychology*, *73*, 1–16.

Sternberg, R. J. (1988). A three-facet model of creativity. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 125–147). New York: Cambridge University Press.

Sternberg, R. J. (1996). *Successful intelligence*. New York: Simon & Schuster.
Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology*, *3*, 292–316.
Sternberg, R. J. (2006). Creating a vision of creativity: The first 25 years. *Psychology of Aesthetics, Creativity, and the Arts, S*, 2–12.
Sternberg, R. J. (2010). *College admissions for the 21st century*. Cambridge, MA: Harvard University Press.
Sternberg, R. J. (2018). A triangular theory of creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *12*, 50–67,
Sternberg, R. J., Kaufman, J. C., & Grigorenko, E. L. (2008). *Applied intelligence*. Cambridge: Cambridge University Press.
Sternberg, R. J., & Lubart, T. I. (1996). *Defying the crowd*. New York: Free Press.
Sternberg, R. J., & O'Hara, L. A. (1999). Creativity and intelligence. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 251–272). New York: Cambridge University Press.
Sternberg, R. J., & The Rainbow Project Collaborators. (2006). The Rainbow Project: Enhancing the SAT through assessment of analytical, practical and creative skills. *Intelligence*, *34*, 321–350.
Swanson, H. L. (1992). The relationship between metacognition and problem solving in gifted children. *Roeper Review*, *15*, 43–48.
Torrance, E. P. (1974). *Torrance Test of Creative Thinking: Directions manual and scoring guide. Verbal test booklet A*. Bensenville, IL: Scholastic Testing Service.
Torrance, E. P. (2008). *The Torrance Tests of Creative Thinking Norms-Technical Manual*. Bensenville, IL: Scholastic Testing Service.
Wai, J., Lubinski, D., & Benbow, C. P. (2005). Creativity and occupational accomplishments among intellectually precocious youths: An age 13 to age 33 longitudinal study. *Journal of Educational Psychology*, *97*, 484–492.
Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity-intelligence distinction*. New York: Holt, Rinehart and Winston.
Wallas, G. (1926). *The art of thought*. New York: Harcourt Brace.
Wechsler, D. (2014). *WISC-V Technical and interpretive manual*. Bloomington, MN: Pearson.
Westby, E. L., & Dawson, V. L. (1995). Creativity: Asset or burden in the classroom? *Creativity Research Journal*, *8*, 1–10.
Valk, W. van der, Sumo, R., Dul, J., & Schroeder, R. (2016). When are contracts and trust necessary for innovation in buyer-supplier relationships? A Necessary Condition Analysis. *Journal of Purchasing and Supply Management*, *22*, 266–277.
Yamamoto, K. (1964). Role of creative thinking and intelligence in high school achievement. *Psychological Reports*, *14*(3), 783–789.
Yu, K., Lu, Z., & Stander, J. (2003). Quantile regression: Applications and current research areas. *Journal of the Royal Statistical Society*, *52*, 331–350.

46 Intelligence and Rationality

Keith E. Stanovich, Maggie E. Toplak, and Richard F. West

Intelligence tests are often treated as if they encompassed all cognitive abilities. Our goal in this chapter is to challenge this assumption by showing that an important class of cognitive skills is missing from commonly used intelligence tests. We accomplish this by showing that intelligence, operationally defined by what current intelligence tests measure, fails to encompass rational thinking.

One way of understanding the difference between rationality and intelligence is to do a little analysis of a phenomenon we have all observed: smart people acting stupidly. We get surprised when someone whom we consider to be smart acts foolishly. But why should we be so surprised? It seems that smart people do foolish things all the time. Wasn't the financial crisis of 2008 just littered with smart people doing dumb things – from the buyers and sellers of the toxic mortgage securities to the homebuyers who seemed to think their house price would double every three years?

So, if it is not rare for smart people to act foolishly, then why the surprise? In fact, the confusion here derives from being caught up in the inconsistencies and incoherence of folk language. The folk terms being used in this discussion are in dire need of some unpacking. Consider the title of an edited book to which we contributed a chapter: *Why Smart People Can Be So Stupid* (Sternberg, 2002). A typical dictionary definition of the adjectival form of the word "smart" is "characterized by sharp quick thought; bright" or "having or showing quick intelligence or ready mental capacity." Thus, being smart seems a lot like being intelligent, according to the dictionary. Dictionaries also tell us that a stupid person is "slow to learn or understand; lacking or marked by lack of intelligence." Thus, if a smart person is intelligent and stupid means a lack of intelligence, then the "smart person being stupid" phrase seems to make no sense.

However, a secondary definition of the word stupid is "tending to make poor decisions or careless mistakes" – a phrase that attenuates the sense of contradiction. A similar thing happens if we analyze the word "dumb" to see if the phrase "smart but acting dumb" makes sense. The primary definition describes "dumb" as the antonym of intelligence, again leading to a contradiction. But, in phrases referring to decisions or actions such as "what a dumb thing to do!," we see a secondary definition like that of stupid: tending to make poor decisions or careless mistakes. These phrases pick out a particular meaning of "stupid" or "dumb" – albeit not the primary one.

It is likewise with the word foolish. A foolish person is a person "lacking good sense or judgment; showing a lack of sense; unwise; without judgment or discretion." This picks out the aspect of "stupid" and "dumb" that we wish to focus on here – the aspect that refers not to intelligence (general mental "brightness") but instead to the tendency to make judicious decisions (or, rather, injudicious ones).

However we phrase it – "smart but acting dumb," "smart but acting foolish," or whatever – we have finally specified the phenomenon: intelligent people taking injudicious actions or holding unjustified beliefs. Folk psychology is picking out two different traits: mental "brightness" (intelligence) and making judicious decisions (*rational* thinking). If we were clear about the fact that the two traits were different, the sense of paradox or surprise at the "smart but acting foolish" phenomenon would vanish. What perpetuates the surprise is that we tend to think of the two traits as one, or at least that they should be strongly associated. The confusion is fostered because psychology has a measurement device (the intelligence test) for the first but not, until recently (see Stanovich, West, & Toplak, 2016), the second. Psychology has a long and storied history (more than a hundred years old) of measuring the intelligence trait. Although, there has been psychological work on rational thinking, this research started much later (Tversky & Kahneman, 1974) and it was not focused on individual differences.

The novice psychology student might be a bit confused at this point – thinking that somewhere along the line they have heard definitions of intelligence that included rationality. Such a student would be right. Many *theoretical* definitions of intelligence incorporate the rationality concept by alluding to judgment and decision-making in the definition. Other definitions emphasize behavioral adaptiveness and thus also fold rationality into intelligence. The problem here is that *none* of these components of rationality – adaptive responding, good judgment, and decision-making – are assessed on *actual tests* of intelligence.

In short, many treatments of the intelligence concept could be characterized as *permissive* conceptualizations rather than *grounded* conceptualizations. Permissive theories include in their definitions of intelligence aspects of functioning that are captured by the vernacular term "intelligence" (adaptation to the environment, showing wisdom, creativity, etc.) *whether or not* these aspects are actually measured by existing tests of intelligence. Grounded theories, in contrast, confine the concept of intelligence to the set of mental abilities actually tested on IQ tests. Adopting permissive definitions of the concept of intelligence serves to obscure what is absent from extant IQ tests. Instead, in order to highlight the *missing* elements in IQ tests, we adopt a thoroughly grounded notion of the intelligence concept in this chapter – one that anchors the concept in what actual IQ tests measure. Likewise, we ground the concept of rationality in operationalizations from current cognitive science.

A Grounded Theory of Intelligence

The closest thing to a consensus, grounded theory of intelligence in psychology is the Cattell-Horn-Carroll (CHC) theory of intelligence (Carroll, 1993; Cattell, 1963, 1998; Horn & Cattell, 1967). It yields a scientific concept of general intelligence, usually symbolized by *g*, and a small number of broad factors, of which two are dominant. Fluid intelligence (G*f*) reflects reasoning abilities operating across of variety of domains – including novel ones. It is measured by tests of abstract thinking such as figural analogies, Raven's Matrices, and series completion (e.g., what is the next number in the series 1, 4, 5, 8, 9, 12, __). Crystallized intelligence (G*c*) reflects declarative knowledge acquired from acculturated learning experiences. It is measured by vocabulary tasks, verbal comprehension, and general knowledge measures. Although substantially correlated, Gf and Gc reflect a long history of considering two aspects of intelligence: intelligence-as -process and intelligence-as-knowledge (Ackerman, 1996, 2014; Duncan, 2010; Hunt, 2011; Nisbett et al., 2012). In addition to G*f* and G*c*, other broad factors represent things like memory and learning, auditory perception, and processing speed (for a full account, see Carroll, 1993).

There is a large literature on the CHC theory and on the processing correlates of Gf and Gc (see Duncan, 2010; Duncan et al., 2008; Engle, 2018; Geary, 2005; Gignac, 2005; Hunt, 2011; Mackintosh & Bennett, 2003; McGrew, 2009; Nisbett et al., 2012). The constructs in the theory have been validated in studies of brain injury, educational attainment, cognitive neuroscience, developmental trends, and information processing. There are, of course, alternative models to the CHC conception (Deary, 2013; Hunt, 2011). For example, Hunt (2011) discusses Johnson and Bouchard's (2005) g-VPR model as an alternative model that is empirically differentiable from the CHC theory. However, for the purposes of the theoretical contrast with rationality, it makes no difference which of the currently viable grounded theories of intelligence we choose. All of them ignore a critical level of cognitive analysis that is important for rationality.

Rationality in Cognitive Science

Rationality is a torturous and tortured term in intellectual discourse. It is contentious and has a multitude of definitions. The term is claimed by many disciplines and parsed slightly differently by each discipline. Philosophy, economics, decision theory, psychology – all claim the term and have their own definitions. For example, animal behaviorists claim to measure degrees of rationality in animals (Kacelnik, 2006); yet, by some of the definitions used in other disciplines, animals couldn't have rationality at all!

Many philosophical notions of rationality are crafted so as to equate all humans – thus, by fiat, defining away the very individual differences that a psychologist wishes to study. For example, some definitions of rationality derive from a categorical notion of rationality tracing to Aristotle that posits humans as the only animals who base actions on reason. As de Sousa (2007) has pointed out, such a notion of rationality as "based on reason" has as its opposite not irrationality but *arationality.* Aristotle's characterization is categorical – the behavior of entities is either based on thought or not rational. In this conception, humans are rational; other animals are not. There is no room for individual differences in rational thinking *among* humans in this view.

In contrast, rationality – in the sense employed in cognitive science and in this chapter – is a normative notion. Normative models of optimal judgment and decision-making define rationality in the noncategorical manner employed in cognitive science. Rationality thus comes in degrees defined by the distance of the thought or behavior from the optimum defined by a normative model (Etzioni, 2014). Thus, when cognitive scientists term a behavior less than rational, they mean that the behavior departs from the optimum prescribed by a particular normative model. The scientist is not implying that no thought or reasoning was behind the behavior, as in the categorical sense.

One reason why psychologists do not adopt categorical definitions of rationality is that, under such definitions, there is no motivation for cognitive reform or cognitive change. Continuous definitions of rationality motivate cognitive reform, because most people are less than optimally rational, and thus most people can improve their rational thinking tendencies.

Rationality: Instrumental and Epistemic

We follow many cognitive science theorists in recognizing two types of rationality, instrumental and epistemic (Manktelow, 2004; Over, 2004) – roughly mapping into the terms practical and theoretical that philosophers prefer. The simplest definition of instrumental rationality is: behaving in the world so that you get exactly what you most want, given the resources (physical and mental) available to you. Epistemic rationality concerns how well beliefs map onto the actual structure of the world. Manktelow (2004) has emphasized the practicality of both types of rationality by noting that they concern two critical things: what is true and what to do. For our beliefs to be rational they must correspond to the way the world is – they must be true (epistemic rationality). For our actions to be rational, they must be the best means toward our goals – they must be the best things to do (instrumental rationality).

More formally, economists and cognitive scientists define instrumental rationality as the maximization of expected utility. To be instrumentally rational, a person must choose among options based on which option has the largest expected utility. Decision situations can be broken down into three components: (1) possible actions; (2) possible states of the world; and (3) evaluations of the consequences of possible actions in each possible state of the world. Expected utility is calculated by taking the utility of each outcome and multiplying it by the probability of that outcome occurring and then summing those products over all of the possible outcomes.

In practice, assessing rationality in this computational manner can be difficult because eliciting personal probabilities can be tricky. Also, getting measurements of the utilities of various consequences can be experimentally difficult. Fortunately, there is another useful way to measure the rationality of decisions and deviations from rationality. It has been proven through several formal analyses that, if people's preferences follow certain consistent patterns (the so-called axioms of choice), then the people are behaving as if they are maximizing utility (Dawes, 1998; Edwards, 1954; Jeffrey, 1983; Luce & Raiffa, 1957; Savage, 1954; von Neumann & Morgenstern, 1944). These analyses have led to what has been termed the axiomatic approach to whether people are maximizing utility. It is what makes people's degrees of rationality more easily measurable by the experimental methods of cognitive science. The deviation from the optimal choice pattern according to the axioms is an (inverse) measure of the degree of rationality.

The axiomatic approach to choice defines instrumental rationality as adherence to certain types of consistency and coherence relationships. For example, one such axiom is that of *transitivity*: If you prefer A to B and B to C, then you should prefer A to C. All of the axioms of choice (independence of irrelevant alternatives, transitivity, independence, and reduction of compound lotteries, etc.), in one way or another, ensure that decisions are not influenced by irrelevant context (Stanovich, 2013). There is considerable empirical evidence in cognitive science indicating that people sometimes violate the axioms of utility theory (Kahneman & Tversky, 2000; Thaler, 2015). It is also known that there is considerable variability in the tendency to adhere to the basic axioms of choice that define instrumental rationality.

An axiomatic approach can be applied to assessing epistemic rationality as well. Recall that the expected utility of an action involves multiplying the probability of an outcome by its utility and summing across possible outcomes. Thus, determining the best action involves estimating the probabilities of various outcomes. Rationality of belief is assessed by looking at a variety of probabilistic reasoning skills, evidence evaluation skills, and hypothesis testing skills. In order for a person to be epistemically rational, their probability estimates must follow the rules of objective probabilities. That is, their estimates must follow the so-called probability calculus.

A substantial research literature – one comprising literally hundreds of empirical studies conducted over several decades – has firmly established that people's responses sometimes deviate from the performance considered normative on many reasoning tasks. For example, people assess probabilities incorrectly, they test hypotheses inefficiently, they violate the axioms of utility theory, they do not properly calibrate degrees of belief, their choices are affected by irrelevant context, they ignore the alternative hypothesis when evaluating data, and they display numerous other information-processing biases (Baron, 2008, 2014; Evans, 2014; Kahneman, 2011; Koehler & Harvey, 2004; Stanovich, 1999, 2011; Stanovich, West, & Toplak, 2016; Thaler, 2015). Much of the operationalization of rational thinking in cognitive science comes from the heuristics and biases tradition where these thinking errors were first discovered in the 1970s (Kahneman & Tversky, 1972, 1973; Tversky & Kahneman, 1974). The term biases refers to the systematic errors that people make in choosing actions and in estimating probabilities and the term heuristic refers to *why* people often make these errors – because they use mental shortcuts (heuristics) to solve many problems.

To this point, using grounded theories of intelligence and rationality, we have established that there are individual differences in performance in both of these domains. In the next sections, we will outline the functional cognitive theory that will situate both concepts. Even more specifically, we will show that rationality is actually a more encompassing mental construct than is intelligence. Thus, as measures of rationality, the tasks in the heuristics and biases literature, while tapping intelligence in part, actually encompass more cognitive processes and knowledge than are assessed by IQ tests.

A Dual-Process Cognitive Architecture

There is a wide variety of evidence that has converged on the conclusion that some type of dual-process model of the mind is needed in a diverse set of specialty areas not limited to cognitive psychology, economics, social psychology, naturalistic philosophy, and decision theory (Alós-Ferrer & Strack, 2014; Chein & Schneider, 2012; De Neys, 2018; Evans, 2008, 2010, 2014; Evans & Frankish, 2009; Evans & Stanovich, 2013; Lieberman, 2009; McLaren et al., 2014; Sherman, Gawronski, & Trope, 2014; Stanovich, 1999, 2004). Evolutionary theorizing and neurophysiological work also have supported a dual-process conception (Corser & Jasper, 2014; Frank, Cohen, & Sanfey, 2009; Lieberman, 2009; McClure & Bickel, 2014; McClure et al., 2004; Prado & Noveck, 2007; Rand et al., 2017; Toates, 2005, 2006).

Because there is now a plethora of dual-process theories (for a list of the numerous versions of such theories, see Stanovich, 2011, 2012), there is currently much variation in the terms for the two processes. For the purposes of this chapter, we will most often adopt the Type 1/Type 2 terminology discussed by Evans and Stanovich (2013) and occasionally use the similar System 1/System 2 terminology of Stanovich (1999) and Kahneman (2011). The defining feature of Type 1 processing is its autonomy – the execution of Type 1 processes is mandatory when their triggering stimuli are encountered and they are not dependent on input from high-level control systems. Autonomous processes have other correlated features – their execution tends to be rapid, they do not put a heavy load on central processing capacity, they tend to be associative – but these other seventy-seven correlated features are not defining (Stanovich & Toplak, 2012). The category of autonomous processes would include processes of emotional regulation; the encapsulated modules for solving specific adaptive problems that have been posited by evolutionary psychologists; processes of implicit learning; and the automatic firing of overlearned associations (see Barrett & Kurzban, 2006; Carruthers, 2006; Evans, 2008, 2009; Moors & De Houwer, 2006; Samuels, 2005, 2009; Shiffrin & Schneider, 1977).

Type 1 processing encompasses many rules, stimulus discriminations, and decision-making principles that have been practiced to automaticity (e.g., Kahneman & Klein, 2009; Shiffrin & Schneider, 1977). These processes can lead to correct or incorrect responding on rational thinking tasks (Evans & Stanovich, 2013). Some participants provide correct immediate responses with high confidence (Bago & De Neys, 2017), which is not surprising if these participants have practiced and consolidated skills in this particular domain. Alternatively, this learned information can sometimes be just as much a threat to rational behavior as are evolutionary modules that fire inappropriately in a modern environment. Rules learned to automaticity can be overgeneralized – they can autonomously trigger behavior when the situation is an exception to the class of events they are meant to cover (Arkes & Ayton, 1999; Hsee & Hastie, 2006).

In contrast with Type 1 processing, Type 2 processing is nonautonomous. It is relatively slow and computationally expensive. Many Type 1 processes can operate at once in parallel but Type 2 processing is largely serial. One of the most critical functions of Type 2 processing is to override Type 1 processing. This is because Type 1 processing heuristics depend on benign environments providing obvious cues that elicit adaptive behaviors. In hostile environments, reliance on heuristics can be costly (see Hilton, 2003; Over, 2000; Stanovich, 2004). A benign environment means one that contains useful (that is, diagnostic) cues that can be exploited by various heuristics (e.g., affect-triggering cues, vivid and salient stimulus components, convenient and accurate anchors). Additionally, for an environment to be classified as benign, it must also contain no other individuals who will adjust their behavior to exploit those relying only on Type 1 processing. In contrast, a hostile environment for heuristics is one in which there are few cues that are usable by autonomous processes or there are misleading cues (Kahneman & Klein, 2009). Another way that an environment can turn hostile for a user of Type 1 processing occurs when other agents discern the simple cues that are being used and arrange them for their own advantage (e.g., advertisements, or the strategic design of supermarket floor space in order to maximize revenue).

All of the different kinds of Type 1 processing (processes of emotional regulation, Darwinian modules, associative and implicit learning processes) can produce responses that are suboptimal in a particular context if not overridden. For example, humans often act as cognitive misers by engaging in attribute substitution (Kahneman & Frederick, 2002) – the substitution of an easy-to-evaluate characteristic in place of a harder one, even if the easier one is less accurate. For example, the cognitive miser will substitute the less effortful attributes of vividness or affect as a replacement for the more effortful retrieval of relevant facts (Slovic & Peters, 2006; Slovic & Slovic, 2015). But, when we are evaluating important risks – such as the risk of certain activities and environments – we do not want to substitute vividness for careful thought about the situation. In such situations, we want to employ Type 2 override processing to block the attribute substitution of the cognitive miser.

Once detection of the conflict between the normative response and the response triggered by System 1 has taken place (on detection, see Bago & De Neys, 2017; Pennycook et al., 2015; Stanovich, 2018; Stanovich et al., 2016; Thompson et al., 2011), Type 2 processing must display at least two related capabilities in order to override Type 1 processing. One is the capability of interrupting Type 1 processing. Type 2 processing thus involves inhibitory mechanisms of the type that have been the focus of work on executive functioning (Kovacs & Conway, 2016; Miyake & Friedman, 2012; Nigg, 2017). But the ability to suppress Type 1 processing gets the job only half done. Suppressing one response is not helpful unless there is a better response available to substitute for it. Where do these better responses come from? One answer is that they can come from processes of hypothetical reasoning and cognitive simulation that are a unique aspect of Type 2 processing (Evans, 2007, 2010; Evans & Stanovich, 2013).

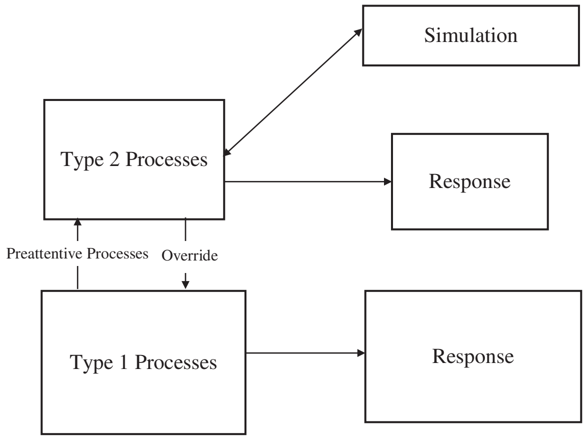

Figure 46.1 *A preliminary dual-process model.*
Reprinted from *What Intelligence Tests Miss: The Psychology of Rational Thought* by Keith E. Stanovich, courtesy of Yale University Press.

When we reason hypothetically, we create temporary models of the world and test out actions (or alternative causes) in that simulated world. In order to reason hypothetically we must, however, have one critical cognitive capability – we must be able to prevent our representations of the real world from becoming confused with representations of imaginary situations. The so-called cognitive decoupling operations (Stanovich, 2011; Stanovich & Toplak, 2012) are the central feature of Type 2 processing that make this possible and they have implications for how we conceptualize both intelligence and rationality, as we shall see. The important issue for our purposes is that decoupling secondary representations from the world and then maintaining the decoupling while simulation is carried out is a Type 2 processing operation. It is computationally taxing and greatly restricts the ability to conduct any other Type 2 operation simultaneously.

A preliminary dual-process model of mind, based on what we have outlined thus far, is presented in Figure 46.1. The figure shows the Type 2 override function we have been discussing, as well as the Type 2 process of simulation. Also rendered in the figure is an arrow indicating that Type 2 processes receive inputs from Type 1 computations. These so-called preattentive processes (Evans, 2008) establish the content of most Type 2 processing.

Three Kinds of Minds and Two Kinds of Individual Differences

In this section, we will explain why rational thinking stresses a level in the hierarchical control system of the brain that is only partly tapped by IQ tests. This is because the override mechanism depicted in Figure 46.1 needs to be conceptualized in terms of two levels of processing. To understand the two levels in a vernacular way, consider two imaginary stories.

Both stories involve a lady walking on a cliff. The stories are both sad – the lady dies in each. The purpose of this exercise is to get us to think about how we explain the death in each story. In incident A, a woman is walking on a cliffside by the ocean and goes to step on a large rock but what appears to be a rock is not a rock at all. Instead, it is actually the side of a crevice and she falls down the crevice and dies. In incident B, a woman attempts suicide by jumping off an ocean cliff and dies when she is crushed on the rocks below.

In both cases, at the most basic level, when we ask ourselves for an explanation of why the woman died, we might say that the answer is the same. The same laws of physics operating in incident A (the gravitational laws that describe why the woman will be crushed on impact) are also operating in incident B. However, we feel that the laws of gravity and force somehow do not provide a complete explanation of what has happened in either incident. When we attempt a more fine-grained explanation, incidents A and B seem to call for a different level of explanation if we wish to zero in on the *essential* cause of death.

In analyzing incident A, a psychologist would be prone to say that when processing a stimulus (the crevice that looked somewhat like a rock), the woman's information-processing system malfunctioned – sending the wrong information to response decision mechanisms, which then resulted in a disastrous motor response. We will refer to this level of analysis as the algorithmic level (on terminology, see Stanovich, 1999). The cognitive psychologist works largely at this level by showing that human performance can be explained by positing certain information-processing mechanisms in the brain (input coding mechanisms, perceptual registration mechanisms, short- and long-term memory storage systems, etc.). In the case of the woman in incident A, the algorithmic level is the right level to explain her unfortunate demise. Her perceptual registration and classification mechanisms malfunctioned by providing incorrect information to response decision mechanisms, causing her to step into the crevice.

Incident B, on the other hand, does not involve such an algorithmic-level information-processing error. The woman's perceptual apparatus accurately recognized the edge of the cliff and her motor command centers quite accurately programmed her body to jump off the cliff. The computational processes posited at the algorithmic level of analysis executed quite perfectly. No error at this level of analysis explains why the woman is dead in incident B. Instead, this woman died because of her overall goals and how these goals interacted with her beliefs about the world in which she lived.

We will present our model of cognitive architecture (building on Stanovich, 2011) in the spirit of Dan Dennett's (1996) book *Kinds of Minds*, where he used that title to suggest that within the brain of humans are control systems of very different types – different kinds of minds. In our terms, the woman in incident A had a problem with the algorithmic mind and the woman in incident B had a problem with the reflective mind. In short, the reflective mind is concerned with the goals of the system, beliefs relevant to those goals, and the choice of action that is optimal given the system's goals and beliefs. All of these characteristics implicate the reflective mind in many issues of rationality. High computational efficiency in the algorithmic mind is not a sufficient condition for rationality.

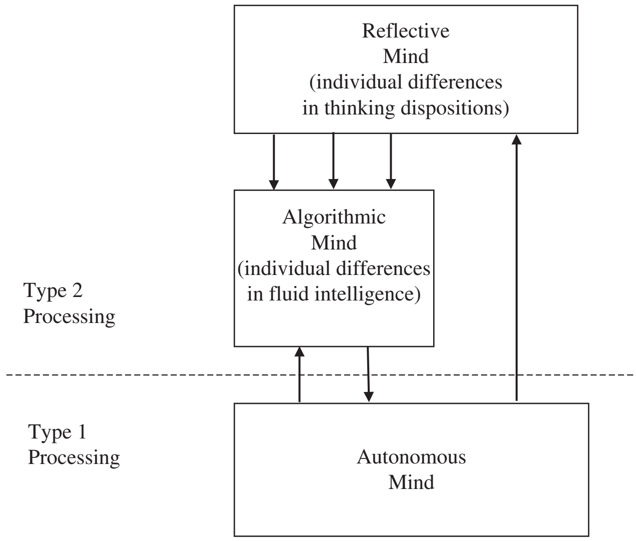

Figure 46.2 *The tripartite structure and the locus of individual differences.* Reprinted from *What Intelligence Tests Miss: The Psychology of Rational Thought* by Keith E. Stanovich, courtesy of Yale University Press.

Our attempt to differentiate the two levels of control involved in Type 2 processing is displayed in Figure 46.2. The psychological literature provides much converging evidence and theory to support such a structure. First, psychometricians have long distinguished typical performance situations from optimal (sometimes termed maximal) performance situations (Ackerman & Kanfer, 2004; Cronbach, 1949; Sternberg, Grigorenko, & Zhang, 2008). Typical performance measures implicate, at least in part, the reflective mind – they assess goal prioritization and epistemic regulation. In contrast, optimal performance situations are those where the task interpretation is determined externally. The person performing the task is told the rules that maximize performance. Thus, optimal performance tasks assess questions of the *efficiency* of goal pursuit – they capture the processing efficiency of the algorithmic mind. All tests of intelligence or cognitive aptitude are optimal performance assessments, whereas measures of critical or rational thinking are often assessed under typical performance conditions.

The difference between the algorithmic mind and the reflective mind is captured in another well-established distinction in the measurement of individual differences – the distinction between cognitive ability and thinking dispositions. The former are, as just mentioned, measures of the efficiency of the algorithmic mind. The latter travel under a variety of names in psychology – thinking dispositions or cognitive styles being the two most popular. Many thinking dispositions concern beliefs, belief structure, and, importantly, attitudes toward forming and changing beliefs. Other thinking dispositions that have been identified concern a person's goals and goal hierarchy. Examples of some thinking dispositions that have been investigated by psychologists are actively open-minded thinking, need for cognition, consideration of future consequences, reflection/intuition, and dogmatism (Baron et al., 2015; Cacioppo et al., 1996; Phillips et al., 2016; Stanovich, 1999, 2011; Sternberg, 2003; Strathman et al., 1994).

In short, measures of individual differences in thinking dispositions are assessing variation in people's goal management, epistemic values, and epistemic self-regulation – differences in the operation of reflective mind. People have indeed come up with *definitions* of intelligence that encompass the reflective level of processing but, nevertheless, *the actual measures of intelligence in use assess only algorithmic-level cognitive capacity.*

Figure 46.2 represents the classification of individual differences in the tripartite view. The broken horizontal line represents the location of the key distinction in older, dual-process views. The figure identifies variation in fluid intelligence (Gf) with individual differences in the efficiency of processing of the algorithmic mind. To a substantial extent Gf measures the ability to cognitively decouple – to suppress Type 1 activity and to enable hypothetical thinking. The raw ability to sustain such simulations while keeping the relevant representations decoupled is one key aspect of the brain's computational power that is being assessed by measures of fluid intelligence. This is becoming clear from converging work on executive function and working memory, which both display correlations with fluid intelligence that are quite high (Chuderski, 2015; Duncan, et al., 2008; Hicks, Harrison, & Engle, 2015; Jastrzębskia, et al., 2018; Kane, Hambrick, & Conway, 2005). This is because most measures of executive function, such as working memory, are direct or indirect indicators of a person's ability to sustain decoupling operations (Feldman Barrett, Tugade, & Engle, 2004). Thus, Type 2 processes are strongly associated with Gf (Burgess et al., 2011; Chuderski, 2014; Engel de Abreu, Conway, & Gathercole, 2010; Kovacs & Conway, 2016; McVay & Kane, 2012; Mrazek et al., 2012; Salthouse et al., 2003). Finally, the reflective mind is identified with individual differences in thinking dispositions related to beliefs and goals.

Why Rationality Is a More Encompassing Construct Than Intelligence

Figure 46.2 highlights an important sense in which rationality is a more encompassing construct than intelligence. As previously discussed, to be rational, a person must have well-calibrated beliefs and must act appropriately on those beliefs to achieve goals – both of these depend on the thinking dispositions of the reflective mind. The types of cognitive propensities that these thinking disposition measures reflect are the tendency to collect information before making up one's mind, the tendency to seek various points of view before coming to a conclusion, the disposition to think extensively about a problem before responding, the tendency to calibrate the degree of strength of one's opinion to the degree of evidence available, the tendency to think about future consequences before taking action, the tendency to explicitly weigh pluses and minuses of situations before making a decision, and the tendency to seek nuance and avoid absolutism.

In order to achieve both epistemic and instrumental rationality, individuals must also, of course, have the algorithmic-level machinery that enables them to carry out the actions and to process the environment in a way that enables the correct beliefs to be fixed and the correct actions to be taken. Thus, individual differences in rational thought and action can arise because of individual differences in fluid intelligence (the algorithmic mind) or because of individual differences in thinking dispositions (the reflective mind) or from a combination of both.

To put it simply, the concept of rationality encompasses thinking dispositions and algorithmic-level capacity, whereas the concept of intelligence (at least as it is commonly operationalized) is largely confined to algorithmic-level capacity. Intelligence tests do not attempt to measure aspects of epistemic or instrumental rationality, nor do they examine any thinking dispositions that relate to rationality. It is clear from Figure 46.2 why rationality and intelligence are separable. Rational thinking depends on thinking dispositions as well as algorithmic efficiency. Thus, as long as variation in thinking dispositions is not perfectly correlated with fluid intelligence, there is the statistical possibility of rationality and intelligence explaining at least partially separable variance.

In fact, substantial empirical evidence indicates that individual differences in thinking dispositions and intelligence are far from perfectly correlated. Studies (e.g., Ackerman & Heggestad, 1997; Cacioppo et al., 1996; Kanazawa, 2004; Zeidner & Matthews, 2000) have indicated that measures of intelligence display only moderate to weak correlations with some thinking dispositions (e.g., actively open-minded thinking, need for cognition) and near zero correlations with others (e.g., conscientiousness, curiosity, diligence). Other important evidence supports the conceptual distinction made here between algorithmic cognitive capacity and thinking dispositions. For example, across a variety of tasks from the heuristics and biases literature, it has consistently been found that rational thinking dispositions will predict variance after the effects of general intelligence have been controlled (Bruine de Bruin, Parker, & Fischhoff, 2007; Finucane & Gullion, 2010; Klaczynski & Lavallee, 2005; Kokis et al., 2002; Macpherson & Stanovich, 2007; Parker & Fischhoff, 2005; Stanovich & West, 1997, 1998; Toplak et al., 2007; Toplak & Stanovich, 2002; Toplak, West, & Stanovich, 2011, 2014a, 2014b).

Rationality and Intelligence in a Fleshed-Out Tripartite Cognitive Architecture

The functions of the different levels of control are illustrated more completely in Figure 46.3. There, it is clear that the override capacity itself is a property of the algorithmic mind and it is indicated by the arrow labeled A. However, previous dual-process theories have tended to ignore the higher-level cognitive operation that initiates the override function in the first place. This is a dispositional property of the reflective mind that is related to rationality. In the model in Figure 46.3, it corresponds to arrow B, which represents the instruction to the algorithmic mind to override the Type 1 response by taking it offline. This is a different mental function

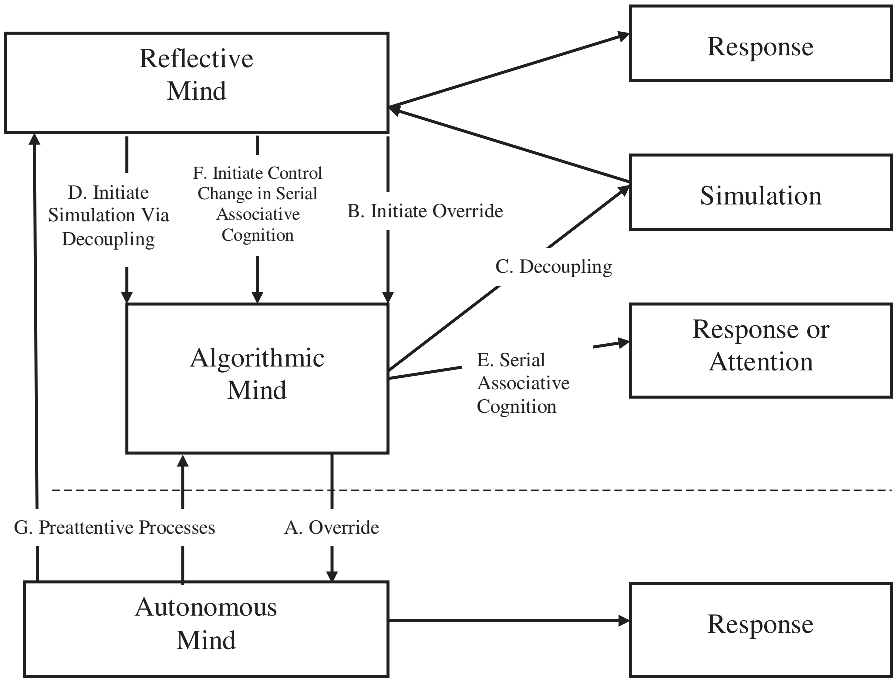

Figure 46.3 *A more complete model of the tripartite structure.*
Reprinted from *What Intelligence Tests Miss: The Psychology of Rational Thought* by Keith E. Stanovich, courtesy of Yale University Press.

than the override function itself (arrow A) and the evidence cited above indicates that the two functions are indexed by different types of individual differences.

The override function has loomed so large in dual-process theory that it has somewhat overshadowed the simulation process that computes the alternative response that makes the override worthwhile. Thus, Figure 46.3 explicitly represents the simulation function as well as the fact that the instruction to initiate simulation originates in the reflective mind. The decoupling operation (indicated by arrow C) itself is carried out by the algorithmic mind. The instruction to initiate simulation (indicated by arrow D) is carried out by the reflective mind. Again, two different types of individual differences are associated with the initiation call and the decoupling operator – specifically, thinking dispositions with the former and fluid intelligence with the latter. Also represented is the fact that the higher levels of control receive inputs from the autonomous mind (arrow G) via so-called preattentive processes (Evans, 2006, 2009).

The arrows labeled E and F reflect the decoupling and higher-level control of a kind of Type 2 processing (serial associative cognition) that does not involve fully explicit cognitive simulation (see Stanovich, 2011). There are types of slow, serial cognition that do not involve simulating alternative worlds and exploring them exhaustively. Their existence points to an important fact: All hypothetical thinking involves Type 2 processing (Evans & Over, 2004) but not all Type 2 processing involves hypothetical thinking. Serial associative cognition represents this latter category. This kind of Type 2 processing is not a full-blown cognitive simulation of alternative world models. It is thinking of a shallower type – cognition that is inflexibly locked into an associative mode that takes as its starting point a model of the world that is given to the subject.

Thus, serial associative cognition is defined by its reliance on a single focal model that triggers all subsequent thought. Framing effects, for instance, are clear examples of serial associative cognition with a focal bias. As Kahneman (2003) notes, "the basic principle of framing is the passive acceptance of the formulation given" (p. 703). The frame presented to the subject is taken as focal and all subsequent thought derives from it rather than from alternative framings because the latter would necessitate more computationally expensive simulation operations. In short, serial associative cognition is sequential and analytic (as opposed to holistic) in style but it relies on a single focal model that triggers all subsequent thought.

Returning to Figure 46.3, we can now identify a third function of the reflective mind – initiating an interrupt of serial associative cognition (arrow F). This interrupt signal alters the next step in a serial associative sequence that would otherwise direct thought. This interrupt signal might stop serial associative cognition altogether in order to initiate a comprehensive simulation (arrow C). Alternatively, it might start a new serial associative chain (arrow E) from a different starting point by altering the temporary focal model that is the source of a current associative chain.

Although taking the Type 1 response priming offline might itself be procedural, the process of synthesizing an alternative response often utilizes stored knowledge of various types. During the simulation process, declarative knowledge and strategic rules (linguistically-coded strategies) are used to transform a decoupled representation. The knowledge, rules, procedures, and strategies that can be retrieved and used to transform decoupled representations have been referred to as mindware, a term coined by Perkins (1995; Clark, 2001, uses the term in a slightly different way from Perkins' original coinage). The mindware available for use during cognitive simulation is in part the product of past learning experiences. This means that there will be individual differences in the ability to simulate better alternatives to a Type 1 response based on variation in the mindware available (Frey, Johnson, & De Neys, 2018; Stanovich, 2018).

Because the CHC theory of intelligence is one of the most comprehensively validated theories of intelligence available, it is important to see how two of its major components miss critical aspects of rational thought. Fluid intelligence will, of course, have some relation to rationality because it indexes the computational power of the algorithmic mind to sustain decoupling. Because override and simulation are important operations for rational thought, Gf will definitely facilitate rational action in some situations. Nevertheless, the tendency to initiate override (arrow B in Figure 46.3) and to initiate simulation activities (arrow D in Figure 46.3) are both aspects of the reflective mind not assessed by intelligence tests, so the tests will miss these components of rationality. Such propensities are instead indexed by measures of typical performance (cognitive styles and thinking dispositions) as opposed to measures of maximal performance such as IQ tests. Measures of wisdom likewise try to tap typical performance (Grossmann, 2017; Sternberg, 2003).

The situation with respect to crystallized intelligence is a little different. Rational thought depends critically on the acquisition of certain types of knowledge (Stanovich, 2018). That knowledge would, in the abstract, be classified as crystallized intelligence. But is it the kind of crystallized knowledge that is assessed on actual tests of intelligence? The answer is "no." The knowledge structures that support rational thought are specialized. They cluster in the domains of probabilistic reasoning, causal reasoning, and scientific reasoning (Stanovich et al., 2016). In contrast, the crystallized knowledge assessed on IQ tests is deliberately designed to be nonspecialized. The designers of the tests, in order to make sure the sampling of vocabulary and knowledge is fair and unbiased, explicitly attempt to *broadly* sample vocabulary, verbal comprehension domains, and general knowledge. In short, crystallized intelligence, as traditionally measured, does not assess individual differences in rationality.

Finally, it may have seemed from our discussion so far that only the algorithmic and reflective minds are implicated in rational thought. Such an interpretation would represent a mistaken implication. In fact, the autonomous mind, as well as the algorithmic and reflective minds, often operates to support rational thought. There is one particular way that the autonomous mind supports rationality that we would like to emphasize. It is the point mentioned previously, that the autonomous mind contains rational rules and normative strategies that have been tightly compiled and that are automatically activated due to overlearning and practice. This means that, for some people, in some instances, the normative response emanates directly from the autonomous mind rather than from the more costly Type 2 process of simulation (see Bago & De Neys, 2017).

Figure 46.4 illustrates more clearly the point we wish to make here. This figure has been simplified by the removal of all the arrow labels and the removal of the boxes representing serial associative cognition, as well as the response boxes. In the upper right is represented the accessing of mindware that is most discussed in the literature. In the case represented there, a nonnormative response from the autonomous mind has been interrupted and the computationally taxing process of simulating an alternative response is underway. That simulation involves the computationally expensive process of accessing mindware for the simulation.

In contrast to this type of normative mindware access, indicated in the lower left of the figure is a qualitatively different way that mindware can determine the normative response. The figure indicates the point we have stressed earlier, that within the autonomous mind can reside normative rules and rational strategies that have been practiced to automaticity and that can *automatically* compete with (and often immediately defeat) any alternative nonnormative response that is also stored in the autonomous mind (Bago & De Neys, 2017; Pennycook et al., 2015).

So it should be clear from Figure 46.4, that it does not follow from the output of a normative response that System 2 was necessarily the genesis of the rational responding. Neither does it necessarily follow (as has been wrongly inferred in some recent research on dual-process theory) that a rapid response should necessarily

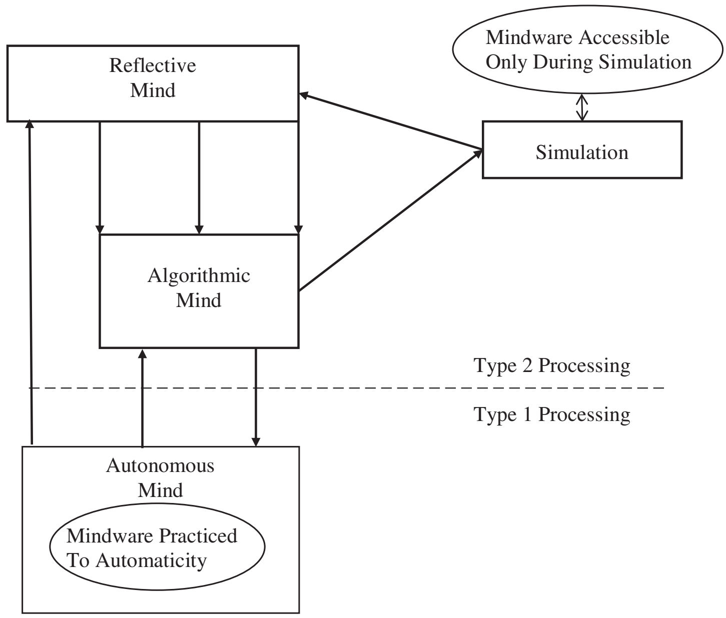

Figure 46.4 *A simplified model showing both automatized mindware and mindware accessible during simulation.*

be an incorrect one (Stanovich, 2018). The main purpose of Figure 46.4 is to concretize the idea that the normative mindware of rational responding is not exclusively retrieved during simulation activities, but can become implicated in performance directly and automatically from the autonomous mind if it has been practiced enough.

According to the model just presented, rationality requires three different classes of mental characteristics. First, algorithmic-level cognitive capacity (Gf) is needed for override and sustained simulation activities. Second, the reflective mind must be characterized by the tendency to initiate the override of suboptimal responses generated by the autonomous mind and to initiate simulation activities that will result in a better response. Finally, the mindware that allows the computation of rational responses needs to be available and accessible during simulation activities or be accessible from the autonomous mind (see Figure 46.4) because it has been highly practiced. Intelligence tests primarily assess only the first of these three characteristics that determine rational thought and action. This is why rationality requires more than just intelligence.

Rationality and the Heuristics and Biases Literature

There exists a nascent literature on the assessment of rational thinking (Bruine de Bruin et al., 2007; Halpern, 2008, 2010; Stanovich, 2009, 2011, 2016; Stanovich & West, 1998, 2008; Stanovich et al., 2016). All of these efforts have drawn on the vast literature in cognitive psychology that has demonstrated violations of the normative models of instrumental and epistemic rationality, especially the heuristics and biases literature (Baron, 2008, 2014; Kahneman, 2011; Kahneman & Tversky, 2000). This is certainly true of our own rational thinking assessment instrument, the Comprehensive Assessment of Rational Thinking (CART; Stanovich et al., 2016).

As measures of rationality, the tasks in the heuristics and biases literature, while tapping intelligence in part, actually encompass more cognitive processes and knowledge than are assessed by IQ tests. Heuristics and biases tasks are often conceptualized within dual-process architectures because most of the tasks in the heuristics and biases literature were deliberately designed to pit an automatically triggered response against a normative response generated by more controlled types of processing (Kahneman, 2011; Kelman, 2011). These tasks, interpreted within a dual-process framework, end up being diagnostic of the dominance of Type 1 versus Type 2 processing in determining the final response. As mentioned previously, for a person who defaults often to Type 1 processing, environments can be either benign or hostile. We have argued (Stanovich, 2004; Stanovich & West, 2000) that the modern world is becoming increasingly hostile to responses derived from Type 1 processing, thus making it important to assess rational thinking tendencies via the logic of heuristics and biases tasks.

It is appropriate here to emphasize another way in which intelligence tests fail to tap important aspects of rational thinking. The novice reader might have thought at this point that it seems that intelligence tests clearly measure Type 2 reasoning – that is, conscious, serial simulation of imaginary worlds in order to solve problems. This is all true, but there is a critical difference. Intelligence tests contain salient warnings that Type 2 reasoning is necessary. Most tests of rational thinking do not strongly cue the subject in this manner. Instead, many heuristics and biases tasks suggest a compelling intuitive response that happens to be wrong. In heuristics and biases tasks, unlike the case for intelligence tests, the subject must detect the inadequacy of the Type 1 response and then must use Type 2 processing to both suppress the Type 1 response and to simulate a better alternative.

Most of the tasks in the heuristics and biases literature have both processing and knowledge requirements. From a processing standpoint, the necessity of overriding Type 1 processing must be detected (unless the relevant normative response is highly automated). Then, the intuitive response primed by Type 1 processing must be inhibited and the normative response must be retrieved or synthesized and then substituted by Type 2 processing. In addition to these processing requirements, successful performance on heuristics and biases tasks requires the presence of several important knowledge bases – the mindware discussed previously. The mindware available for use during cognitive simulation is in part the product of past learning experiences. This means that there will be individual differences in the ability to simulate better alternatives to a Type 1 response based on variation in the mindware available.

The fact that many items on the CART tap process as well as knowledge is specifically intended (as it was in the original heuristics and biases literature) and is not a flaw. It is a design feature, not a drawback. In the domain of rational thinking, we are interested in individual differences in the *sensitivity* to probabilistic reasoning principles, for example. People can have knowledge of these principles without the propensity to use them. They can have the knowledge but not the propensity to see situations in terms of probabilities. A typical item on the CART will pit a statistical way of viewing a problem against a nonstatistical way of viewing a problem in order to see which kind of thinking dominates in the situation. So, for example, we would not design an item where the subject chooses between a nine out of ten chance of winning versus a three out of ten chance of winning, with no other context provided. Instead, on most of the Probabilistic Reasoning subtest items on the CART, statistical information is presented but also a *nonstatistical* way of thinking about the problem. People who may get the pure mathematics of statistical reasoning correct might well not see certain problems themselves as probabilistic. Rational thinking assessment taps variance in sensitivity to *seeing* a problem as probabilistic.

The CART Tasks and Framework

It is important to stress that knowledge and process are intertwined in most heuristics and biases tasks but that it is not the case that the dependence on knowledge and the dependence on process are the same for each and every task. Some heuristics and biases tasks are more process dependent than knowledge dependent. Others are more knowledge dependent than process dependent. Still others seem to stress knowledge and process both quite strongly.

Table 46.1 presents the overall framework for the CART, as well as some indication of the tasks used for assessment. The left column of Table 46.1 serves to represent tasks saturated with processing requirements. The second column from the left represents tasks that are relatively saturated with knowledge from specific rational thinking domains. The first two domains of rational thinking represented in the upper left – probabilistic and statistical reasoning and scientific reasoning – have process and knowledge so intertwined that they span both columns in Table 46.1 to emphasize this point.

Working down the left column, Table 46.1 next identifies some tasks that have heavy processing requirements. The first set of tasks are indicators of the tendency to avoid miserly information processing. That humans are cognitive misers has been a major theme throughout the past forty years of research in psychology and cognitive science (see Dawes, 1976; Kahneman, 2011; Simon, 1955, 1956; Taylor, 1981; Tversky & Kahneman, 1974; for the evolutionary reasons why, see Stanovich, 2004, 2009). When approaching any problem, our brains have available various computational mechanisms for dealing with the situation. These mechanisms embody a tradeoff, however. The tradeoff is between power and expense. Some mechanisms have great computational power – they can solve a large number of novel problems with great accuracy. However, this power comes with a cost. These mechanisms (Type 2 processing) take up a great deal of attention, tend to be slow, tend to interfere with other thoughts and actions we are carrying out, and they require great concentration that is often experienced as aversive. Humans are cognitive

Table 46.1 *Framework for classifying the types of rational thinking tasks and subtests on the CART.*

Tasks Saturated with Processing Requirements (Detection, Sustained Override, Hypothetical Thinking)	Rational Thinking Tasks Saturated with Knowledge	Avoidance of Contaminated Mindware	Thinking Dispositions that Foster Thorough and Prudent Thought, Unbiased Thought, and Knowledge Acquisition
Probabilistic and Statistical Reasoning Subtest		Superstitious Thinking Subtest	Actively Open-Minded Thinking Scale
Scientific Reasoning Subtest		Anti-Science Attitudes Subtest	Deliberative Thinking Scale
Avoidance of Miserly Information Processing Subtests: – Reflection versus Intuition – Belief Bias Syllogisms – Ratio Bias – Disjunctive Reasoning	Probabilistic Numeracy Subtest	Conspiracy Beliefs Subtest	Future Orientation Scale
Absence of Irrelevant Context Effects in Decision Making Subtests: – Framing – Anchoring – Preference Anomalies	Financial Literacy and Economic Knowledge Subtest	Dysfunctional Personal Beliefs Subtest	Differentiation of Emotions Scale
Avoidance of Myside Bias: – Argument Evaluation Subtest	Sensitivity to Expected Value Subtest		
Avoiding Overconfidence: – Knowledge Calibration Subtest	Risk Knowledge Subtest		
Rational Temporal Discounting Subtest			

misers because their basic tendency is to default to other less-accurate processing mechanisms of low computational expense (the Type 1 processes).

The CART contains several subtests that assess a person's ability to avoid miserly information processing. One, the Reflection Versus Intuition subtest, was inspired by a famous problem introduced into the literature by Kahneman and Frederick (2002): A bat and a ball cost $1.10 in total. The bat costs $1 more than the ball. How much does the ball cost?

When they answer this problem, many people give the first response that comes to mind – 10 cents – without thinking further and realizing that this cannot be correct. The bat would then have to cost $1.10 and the total cost would be $1.20 rather than the required $1.10. People often do not think deeply enough to realize their error and cognitive ability is no guarantee against making the error. Frederick (2005) found that large numbers of highly select university students at MIT, Princeton, and Harvard were cognitive misers – they responded that the cost was 10 cents rather than the correct answer: 5 cents. The other direct measures of miserly processing are overcoming belief bias in a syllogistic reasoning task, the ability to overcome ratio bias, and the ability to engage in fully disjunctive reasoning.

Continuing down the left column of Table 46.1 are some other tasks that are best viewed as *indirect* measures of the avoidance of miserly processing. All are heavy in their processing requirements. All of these tasks and their associated effects, although involving miserly processing, are still quite complex tasks. More than miserly processing is going on when someone answers suboptimally in all of them. In any case, they are all are important measures of rational thinking in their own right whether or not they are due to miserly information processing. The left-hand column of Table 46.1 shows several of these important categories of our assessment battery: the absence of irrelevant context effects in decision-making; the avoidance of myside bias; the avoidance of overconfidence in knowledge calibration; and rational temporal discounting of future rewards.

In the second column from the left in Table 46.1 are four components of the CART that represent components that are particularly heavily dependent on knowledge bases. This is not to say that these components are completely independent of the degree of miserly processing, just that variation on them is considerably less dependent on processing considerations and much more dependent on the presence of certain specific types of declarative knowledge than other tasks. These subtests of the CART tap the following: probabilistic numeracy; financial literacy and economic knowledge; sensitivity to expected value; and risk knowledge.

The third column in Table 46.1 reflects the fact that irrational thinking is potentially caused by two different types of mindware problems. Missing mindware, or mindware gaps, reflect the most common type – where a person does not have access to adequately compiled declarative knowledge from which to synthesize a normative response to use in the override of Type 1 processing. However, not all mindware is helpful or useful in fostering rationality (Stanovich, 2004, 2009, 2011). Indeed, the presence of certain kinds of mindware is often precisely the problem. We use the category label *contaminated mindware* for the presence of declarative knowledge bases that foster irrational rather than rational thinking.

There are probably dozens of different kinds of contaminated mindware if one looks very specifically at narrow domains of knowledge. It would obviously be impossible for a test of rational thinking to encompass all of these. Instead, we have focused on just a few of the broader categories of contaminated mindware that might have more general implications and might have some domain generality in their effects. Of course, rational thinking as indicated by CART performance, is defined as the *avoidance* or *rejection* of these domains of contaminated mindware. The third column from the left in Table 46.1 lists the four categories of contaminated mindware that are assessed on the CART: superstitious thinking; anti-scientific attitudes; conspiracy beliefs; and dysfunctional personal beliefs.

Superstitious thinking is measured with twelve items such as "a person's thoughts can influence the movement of a physical object" and "astrology can be useful in making personality judgments." Anti-scientific attitudes are measured with thirteen items such as "I don't place great value on scientific facts, because scientific facts can be used to prove almost anything." Dysfunctional personal beliefs are measured with nine items such as "I worry a lot that I am unlikable." Finally, there are twenty-four items on the Conspiracy Beliefs subtest of the CART. We drew on a large number of conspiracies that have been studied in the literature and added a few new ones of our own. Our subtest covered a wide range of conspiratorial beliefs. Most importantly, it contained conspiracies of both the political left and the political right. Unlike some previous measures, it was not just a proxy for political attitudes. Some of the commonly studied conspiracies that we assessed were the assassination of President John F. Kennedy, the 9/11 attacks, fluoridation, the moon landing, pharmaceutical industry plots, the spread of AIDS, oil industry plots, and Federal Reserve conspiracies.

Finally, the far right column of Table 46.1 shows a set of supplementary measures that are included in the CART but are not part of the overall rational thinking score on the test itself. Column four lists some thinking dispositions that we measure by self-report questionnaires. There are many different thinking dispositions studied in psychology. However, we have chosen those specifically relevant to rational thinking. The four thinking dispositions that we assess are actively open-minded thinking; deliberative thinking; future orientation; and the differentiation of emotions. These self-report measures are different from the other performance measures on the CART, which is why they are not part of the overall score on the test but instead provide supplementary information. They are not part of the total score on the test because, among other things, the maximum score on a thinking disposition measure is not to be equated with the maximal rationality. Optimal functioning on these measures is traced instead by an inverted U-shaped function. Maximizing these dispositions is not the criterion of rational thought itself. Thinking dispositions such as these are a means to rationality, not ends in themselves. For this reason, thinking dispositions subscales are segregated in the CART and not treated as direct measures of rational thinking themselves.

The Unique Features of Rationality Assessment:CART Subtests ≠ IQ Test Components

With the construction of the CART, we now have an instrument designed to assess the types of cognitive skills that have been studied for forty years in the heuristics and biases literature. It is amazing that until now we have not had a battery that comprehensively assesses these cognitive skills, given their epic influence on cognitive science. The 1974 *Science* paper by Tversky and Kahneman had, by the year 2018, received more than 46,000 citations according to Google Scholar. Kahneman's recent (2011) book had received more than 18,000 citations by the same time. These numbers, along with the 2002 Nobel Prize to Kahneman, represent an unprecedented scientific influence. Yet, until the CART and the work that preceded it (Bruine de Bruin et al., 2007; Stanovich & West, 1998), psychologists had completely neglected to develop assessment devices for these unique cognitive skills.

Of course, small sets of heuristics and biases tasks have been examined together before (Cokely & Kelley, 2009; Liberali et al., 2012; Stanovich & West, 1998). Nevertheless, our collection is unique in its comprehensiveness. However, it is important to stress that the issue of measuring rationality goes far beyond the comprehensiveness of the heuristics and biases battery that is involved. Instrumental and epistemic rationality, as defined in this chapter, both implicate important knowledge bases when their definitions are operationalized. The CART is unique in this particular respect, that is, in explicitly encompassing important declarative knowledge bases in its assessment model. Beyond the measurement of the important probabilistic reasoning tendencies and reflective reasoning tendencies that are well captured by the heuristics and biases tasks, the CART taps knowledge bases that importantly facilitate rational thought and behavior as well as knowledge bases that importantly impede normative responding.

The emphasis on heuristics and biases tasks (e.g., Probabilistic and Statistical Reasoning subtest) and subtests composed more purely of knowledge assessment (e.g., Financial Literacy subtest) in the CART highlights the two most important ways in which the CART is different from IQ tests. Regarding knowledge, the important point to note is that the knowledge bases assessed on the CART are domain-specific (financial literacy; avoidance of conspiracy beliefs) and not like the broad-based vocabulary assessments of IQ tests.

Regarding the parts of the CART that are composed of heuristics and biases tasks, the logic of these tasks makes it possible for the CART to measure the *propensity* to use a cognitive skill in a way that IQ tests do not. In the domain of rational thinking, we are interested in individual differences in the *sensitivity* to probabilistic reasoning principles and in the *tendency* to apply scientific principles when seeking causal explanations. People who can answer an explicit probability question on a test, or who can accurately define "control-group" when asked, may not show the sensitivity to invoke these principles when their relevance to a problem is partially disguised. In contrast, the cognitive skills assessed by IQ tests are explicit ones. The respondent does not have to recognize their applicability – and does not have to overcome an intuitive response that the problem deliberately activates. On IQ tests, people are not tempted to engage in miserly processing due to the presence of an intuitively compelling alternative.

The results that we have obtained with our Actively Open-Minded Thinking (AOT) thinking disposition scale are consistent with these differences between the CART and IQ tests. Although the AOT scale is correlated with both, it correlated significantly more strongly with CART performance than with cognitive ability (Stanovich et al., 2016). The AOT also remains a significant predictor of most of the subtests after cognitive ability has been partialled out. Our AOT results indicate a startlingly tight linkage between a particular thinking disposition and rational thinking. A generic style of thought – one characterized by the cultivation of reflectiveness rather than impulsivity, the seeking and processing of information that disconfirms one's beliefs (as opposed to confirmation bias in evidence seeking), and the willingness to change one's beliefs in the face of contradictory evidence – has been linked in our data to a very comprehensive measure of rational thought. The results from the AOT show that there is a global mental attitude that pervades these tasks. It is certainly not a specific cognitive skill but instead is best characterized as a generic mental attitude toward cognitive tasks – one of openness, full engagement, mental caution, exhaustiveness of thought, humility, and willingness to encompass new evidence. Rationality is multifarious, involving knowledge and process in complex and changing proportions across tasks and situations. Nevertheless, despite the multifariousness of the rationality construct itself, a global thinking style – actively open-minded thinking – does permeate almost all of the subtests of the CART.

Rational Thinking Subsumes Critical Thinking

We saw in a previous discussion that the concept of rationality – in encompassing both the reflective mind and the algorithmic mind – can be said to be a superordinate construct to intelligence. The study of rational thinking is a normative/evaluative endeavor (Lee, 2006). Specifically, if one's goal is to *aid* people in their thinking, then it is essential that one have some way of *evaluating* thinking. The admonition to educators to "teach critical thinking skills" contains implicit evaluative assumptions. The students *already* think. Educators are charged with getting them to think *better* (Adams, 1993; Baron, 1993). This of course implies normative models of what we mean by better thinking (Baron, 1993, 2008). The assessment of rational thinking explicitly uses those normative models.

Normative issues arise when thinking dispositions are discussed in the educational literature of critical thinking. Why do we want people to think in an actively open-minded fashion? Why do we want people to be reflective? It can be argued that the superordinate goal we are actually trying to foster is that of rationality (Stanovich, 2004, 2009). That is, much of what educators are ultimately concerned about is rational thought in both the epistemic sense and the instrumental sense. We value certain thinking dispositions because we think that they will at least aid in bringing belief in line with the world (epistemic rationality) and in achieving our goals (instrumental rationality).

In short, a large part of the rationale for educational interventions to change thinking dispositions derives from a tacit assumption that actively open-minded critical-thinking dispositions make the individual a more rational person. Thus, the normative justification for fostering critical thought is that it is the foundation of rational thought. Our view is consistent with that of many other theorists who have moved toward conceptualizing critical thinking as a subspecies of rational thinking, or at least as closely related to rational thinking (Kuhn, 2005; Moshman, 2004; Siegel, 1997). Additionally, theory in cognitive science differentiates rationality from intelligence and explains why rationality and intelligence sometimes dissociate. This finding confirms a long-standing belief in education that intelligence does not guarantee critical thinking.

The Context of Rational Thinking Assessment

For many years, we have argued that professional inertia and psychologists' investment in IQ testing have prevented us from realizing that our science had developed enough to allow us to develop a parallel RQ test (Stanovich, 1993, 2009; Stanovich, Toplak, & West, 2008). With the development of the CART, our research group has turned this prediction into reality (Stanovich et al., 2016). Although our initial effort should be viewed more as a prototype, it accomplishes the task of showing that there is nothing conceptually or theoretically preventing us from developing such a test. We know the types of thinking processes that would be assessed by such an instrument and we have in hand prototypes of the kinds of tasks that would be used in the domains of both instrumental rationality and epistemic rationality – both of which are represented on the CART. The existence of the CART demonstrates that there is no practical limitation to constructing a rational thinking test.

The current version of the CART, which has twenty subtests and four supplementary thinking dispositions scales, takes less than three hours to complete. The total CART score has a reliability of 0.86 but the reliability of the individual subtests varies widely. Nevertheless, most of the subtests are themselves reliable enough for research purposes. In a study of the full CART employing more than 700 subjects (discussed in Stanovich, West, & Toplak, 2016), we found that the amalgamated total score on the CART displayed substantial correlations, in the range of 0.50–0.70, with various measures of cognitive ability. Importantly, though, the subcomponents of the CART display quite variable correlations with cognitive ability. Components such as scientific reasoning and reflective thinking display moderate correlations with cognitive ability (0.51 and 0.54), whereas other components show much lower associations with cognitive ability – for example, avoiding overconfidence (0.38), optimal temporal discounting (0.06), and argument evaluation (0.37). Importantly, the cognitive disposition of actively open-minded reasoning predicts performance on virtually all of the CART subtests after cognitive ability has been partialled out. This is true of the total CART score as well. Finally, a very short form of the CART consisting of just two of the subtests – scientific reasoning and probabilistic reasoning – can be used in many situations.

Unlike many such lists of thinking skills in textbooks, the conceptual components of the CART are each grounded in a task or paradigm in the literature of cognitive science. In fact, many (e.g., probabilistic reasoning) have generated enormous empirical literatures. For example, there are many paradigms that have been used to measure the avoidance of miserly information processing (left column of Table 46.1, third row). Another part of the CART that is richly populated by work in cognitive science is a set of tasks that collectively define the mental tendency to not be affected by irrelevant context in decision-making (left column of Table 46.1, fourth row). All three paradigms that assess the latter tendency have each generated enormous literatures. Resistance to framing has been measured with countless tasks (e.g., Levin et al., 2002; Maule & Villejoubert, 2007), as has the resistance to irrelevant anchoring in decisions (e.g., Epley & Gilovich, 2004, 2006; Jacowitz & Kahneman, 1995). Lichtenstein and Slovic (2006) summarized several decades' worth of work on preference anomalies that followed their seminal research in the 1970s (Lichtenstein & Slovic, 1971, 1973).

The existence of the CART is our attempt to follow through on a claim made years ago (Stanovich, 2009) – that there is no *conceptual* barrier to creating a prototype of a test of rational thinking. This does not of course mean that there is not substantial work to be done in turning the prototype into an easily usable test. We have given a book-length treatment (Stanovich et al., 2016) of the twenty years of work on individual differences in rational thinking that went into the development of our prototype. Rationality is a multifarious concept and reporting outcomes across all of its components (see Table 46.1) is complex. Nevertheless, a reasonable amount of research has already been conducted linking rational thinking tendencies to real-life decision-making (Baron, Bazerman, & Shonk, 2006; Bruine de Bruin et al., 2007; Camerer, 2000; Fenton-O'Creevy, et al., 2003; Hilton, 2003; Milkman, Rogers, & Bazerman, 2008; Parker et al., 2015; Thaler, 2015; Thaler & Sunstein, 2008). In our book (Stanovich et al., 2016), we include a table indicating how each of the thinking skills assessed on the CART have been linked to real-life outcomes in the work of other investigators. Performance on several CART subtests has been shown to be related to several real-world behaviors, including secure computing and making prudent financial choices (Toplak, West, & Stanovich, 2017).

Implications of Rational Thinking Assessment

When a layperson thinks of individual differences in reasoning, they think of IQ tests. It is quite natural that this is their primary association, because IQ tests are among the most publicized products of psychological research. This association is not entirely inaccurate either, because intelligence is correlated with performance on a host of reasoning tasks and real-life outcomes (Carroll, 1993; Deary, 2000; Hunt, 2011; Schmidt & Hunter, 2004). Nonetheless, certain very important classes of individual differences in thinking are ignored if only intelligence-related variance is the primary focus. A number of these ignored classes of individual differences are those relating to rational thought.

We tend not to notice the mental processes that are missing from IQ tests because, as we noted at the beginning of this chapter, many theorists have adopted permissive conceptualizations of what intelligence is rather than a grounded conceptualization. In contrast, in this chapter, we have stressed that the operationalization of rationality is different from that of intelligence and thus, as every introductory psychology student is taught, the concepts must be treated as different. Our comprehensive test of rational thinking will go a long way toward grounding the rationality concept – a concept that captures aspects of thought that have heretofore gone unmeasured in assessment devices.

Now that we have the CART, we could, in *theory*, begin to assess rationality as systematically as we do IQ. We could choose tomorrow to more formally assess rational thinking skills, focus more on teaching them, and redesign our environment so that irrational thinking is not so costly. Whereas just thirty years ago we knew vastly more about intelligence than we knew about rational thinking, this imbalance has been redressed in the last few decades because cognitive scientists have developed laboratory tasks and real-life performance indicators to measure rational thinking.

References

Ackerman, P. L. (1996). A theory of adult development: Process, personality, interests, and knowledge. *Intelligence*, *22*, 227–257.

Ackerman, P. L. (2014). Adolescent and adult intellectual development. *Current Directions in Psychological Science*, *23*, 246–251.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.

Ackerman, P. L., & Kanfer, R. (2004). Cognitive, affective, and conative aspects of adult intellect within a typical and maximal performance framework. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 119–141). Mahwah, NJ: Lawrence Erlbaum Associates.

Adams, M. J. (1993). Towards making it happen. *Applied Psychology: An International Review*, *42*, 214–218.

Alós-Ferrer, C., & Strack, F. (2014). From dual processes to multiple selves: Implications for economic behavior. *Journal of Economic Psychology*, *41*, 1–11.

Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals? *Psychological Bulletin*, *125*, 591–600.

Bago, B., & De Neys, W. (2017). Fast logic? Examining the time course assumption of dual process theory. *Cognition*, *158*, 90–109.

Baron, J. (1993). Why teach thinking?–An essay. *Applied Psychology: An International Review*, *42*, 191–214.

Baron, J. (2008). *Thinking and deciding* (4th ed.). Cambridge: Cambridge University Press.

Baron, J. (2014). Heuristics and biases. In E. Zamir & D. Teichman (Eds.), *The Oxford handbook of behavioral economics and the law* (pp. 3–27). Oxford: Oxford University Press.

Baron, J., Bazerman, M. H., & Shonk, K. (2006). Enlarging the societal pie through wise legislation. A psychological perspective. *Perspectives on Psychological Science*, *1*, 123–132.

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, *4*, 265–284.

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, *113*, 628–647.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, *92*, 938–956.

Burgess, G. C., Gray, J. R., Conway, A. R., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology: General*, *140*, 674–692.

Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197–253.

Camerer, C. F. (2000). Prospect theory in the wild: Evidence from the field. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 288–300). Cambridge: Cambridge University Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Carruthers, P. (2006). *The architecture of the mind*. New York: Oxford University Press.

Cattell, R. B. (1963). Theory for fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1–22.

Cattell, R. B. (1998). Where is intelligence? Some answers from the triadic theory. In J. J. McArdle & R. W. Woodcock (Eds.), *Human cognitive abilities in theory and practice* (pp. 29–38). Mahwah, NJ: Lawrence Erlbaum Associates.

Chein, J., & Schneider, W. (2012). The brain's learning and control architecture. *Current Directions in Psychological Science*, *21*, 78–84.

Chuderski, A. (2014). The relational integration task explains fluid reasoning above and beyond other working memory tasks. *Memory and Cognition*, *42*, 448–463.

Chuderski, A. (2015). The broad factor of working memory is virtually isomorphic to fluid intelligence tested under time pressure. *Personality and Individual Differences*, *85*, 98–104.

Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science*. New York: Oxford University Press.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, *4*, 20–33.

Corser, R., & Jasper, J. D. (2014). Enhanced activation of the left hemisphere promotes normative decision making. *Laterality: Asymmetries of Body, Brain and Cognition*, *19*(3), 368–382.

Cronbach, L. J. (1949). *Essentials of psychological testing*. New York: Harper.

Dawes, R. M. (1976). Shallow psychology. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior* (pp. 3–11). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dawes, R. M. (1998). Behavioral decision making and judgment. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology*, Vol. 1 (pp. 497–548). Boston: McGraw-Hill.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford: Oxford University Press.

Deary, I. J. (2013). Intelligence. *Current Biology*, *23*, R673–R676.

De Neys, W. (Ed.) (2018). *Dual process theory 2.0*. London: Routledge.

de Sousa, R. (2007). *Why think? Evolution and the rational mind*. Oxford: Oxford University Press.

Dennett, D. C. (1996). *Kinds of minds: Toward an understanding of consciousness*. New York: Basic Books.

Duncan, J. (2010). *How intelligence happens*. New Haven, CT: Yale University Press.
Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., Bishop, S., & Nimmo-Smith, I. (2008). Goal neglect and Spearman's g: Competing parts of a complex task. *Journal of Experimental Psychology-General*, *137*, 131–148.
Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*, 380–417.
Engel de Abreu, P. M. J., Conway, A. R. A., & Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence*, *38*, 552–561.
Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*, 190–193.
Epley, N., & Gilovich, T. (2004). Are adjustments insufficient? *Personality and Social Psychology Bulletin*, *30*, 447–460.
Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science*, *17*, 311–318.
Etzioni, Amitai. (2014). Treating rationality as a continuous variable. *Society*, *51*, 393–400.
Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review*, *13*, 378–395.
Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgment*. New York: Psychology Press.
Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology*, *59*, 255–278.
Evans, J. St. B. T. (2009). How many dual-process theories do we need? One, two, or many? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 33–54). Oxford: Oxford University Press.
Evans, J. St. B. T. (2010). *Thinking twice: Two minds in one brain*. Oxford: Oxford University Press.
Evans, J. St. B. T. (2014). *Reasoning, rationality and dual processes*. London: Psychology Press.
Evans, J. St. B. T., & Frankish, K. (Eds.). (2009). *In two minds: Dual processes and beyond*. Oxford: Oxford University Press.
Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.
Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8, 223–241.
Feldman Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, *130*, 553–573.
Fenton-O'Creevy, M., Nicholson, N., Soane, E., & Willman, P. (2003). Trading on illusions: Unrealistic perceptions of control and trading performance. *Journal of Occupational and Organizational Psychology*, *76*, 53–68.
Finucane, M. L., & Gullion, C. M. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychology and Aging*, *25*, 271–288.
Frank, M. J., Cohen, M., & Sanfey, A. G. (2009). Multiple systems in decision making. *Current Directions in Psychological Science*, *18*, 73–77.
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42.
Frey, D., Johnson, E., & De Neys, W. (2018). Individual differences in conflict detection during reasoning. *Quarterly Journal of Experimental Psychology*, *71*(5), 1188–1208.
Geary, D. C. (2005). *The origin of the mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: American Psychological Association.

Gignac, G. E. (2005). Openness to experience, general intelligence and crystallized intelligence: A methodological extension. *Intelligence*, *33*, 161–167.

Grossmann, I. (2017). Wisdom in context. *Perspectives on Psychological Science*, *12*, 956–972.

Halpern, D. (2008). *Halpern Critical thinking assessment: Background and scoring standards*. Unpublished manuscript. Claremont, CA: Claremont McKenna College.

Halpern, D. (2010). *Halpern Critical Thinking Assessment: Manual Version 21*. Mödling: Schuhfried.

Hicks, K. L., Harrison, T. L., & Engle, R. W. (2015). Wonderlic, working memory capacity, and fluid intelligence. *Intelligence*, *50*, 186–195.

Hilton, D. J. (2003). Psychology and the financial markets: Applications to understanding and remedying irrational decision-making. In I. Brocas & J. D. Carrillo (Eds.), *The psychology of economic decisions*, Vol. 1, *Rationality and well-being* (pp. 273–297). Oxford: Oxford University Press.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 1–23.

Hsee, C. K., & Hastie, R. (2006). Decision and experience: Why don't we choose what makes us happy? *Trends in Cognitive Sciences*, *10*, 31–37.

Hunt, E. (2011). *Human intelligence*. Cambridge: Cambridge University Press.

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, *21*, 1161–1167.

Jastrzębskia, J., Ciechanowska, I., & Chuderski, A. (2018). The strong link between fluid intelligence and working memory cannot be explained away by strategy use. *Intelligence*, *66*, 44–53.

Jeffrey, R. C. (1983). *The logic of decision* (2nd ed.). Chicago: University of Chicago Press.

Johnson, W., & Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 393–416.

Kacelnik, A. (2006). Meanings of rationality. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 87–106). Oxford: Oxford University Press.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697–720.

Kahneman, D. (2011). *Thinking, fast and slow*. New York:Farrar, Straus & Giroux.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York: Cambridge University Press.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, *64*, 515–526.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Cambridge: Cambridge University Press.

Kanazawa, S. (2004). General intelligence as a domain-specific adaptation. *Psychological Review*, *111*, 512–523.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66–71.

Kelman, M. (2011). *The heuristics debate*. New York: Oxford University Press.
Klaczynski, P. A., & Lavallee, K. L. (2005). Domain-specific identity, epistemic regulation, and intellectual ability as predictors of belief-based reasoning: A dual-process perspective. *Journal of Experimental Child Psychology*, *92*, 1–24.
Koehler, D. J., & Harvey, N. (Eds.). (2004). *Blackwell handbook of judgment and decision making*. Oxford: Blackwell.
Kokis, J., Macpherson, R., Toplak, M., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, *83*, 26–52.
Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, *27*, 151–177.
Kuhn, D. (2005). *Education for thinking*. Cambridge, MA: Harvard University Press.
Lee, C. J. (2006). Gricean charity: The Gricean turn in psychology. *Philosophy of the Social Sciences*, *36*, 193–218.
Levin, I. P., Gaeth, G. J., Schreiber, J., & Lauriola, M. (2002). A new look at framing effects: Distribution of effect sizes, individual differences, and independence of types of effects. *Organizational Behavior and Human Decision Processes*, *88*, 411–429.
Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, *25*, 361–381.
Lichtenstein, S., & Slovic, P. (1971). Reversal of preferences between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*, 46–55.
Lichtenstein, S., & Slovic, P. (1973). Response-induced reversals of preference in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology*, *101*, 16–20.
Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference*. Cambridge: Cambridge University Press.
Lieberman, M. D. (2009). What zombies can't do: A social cognitive neuroscience approach to the irreducibility of reflective consciousness. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 293–316). Oxford: Oxford University Press.
Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York: Wiley.
Mackintosh, N. J., & Bennett, E. S. (2003). The fractionation of working memory maps onto different components of intelligence. *Intelligence*, *31*, 519–531.
Macpherson, R., & Stanovich, K. E. (2007). Cognitive ability, thinking dispositions, and instructional set as predictors of critical thinking. *Learning and Individual Differences*, *17*, 115–127.
Manktelow, K. I. (2004). Reasoning and rationality: The pure and the practical. In K. I. Manktelow & M. C. Chung (Eds.), *Psychology of reasoning: Theoretical and historical perspectives* (pp. 157–177). Hove: Psychology Press.
Maule, J., & Villejoubert, G. (2007). What lies beneath: Reframing framing effects. *Thinking and Reasoning*, *13*, 25–44.
McClure, S. M., & Bickel, W. K. (2014). A dual-systems perspective on addiction: contributions from neuroimaging and cognitive training. *Annals of the New York Academy of Sciences*, *1327*, 62–78.
McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*, 503–507.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10.

McLaren, I. P., Forrest, C. L., McLaren, R. P., Jones, F. W., Aitken, M. R., & Mackintosh, N. J. (2014). Associations and propositions: the case for a dual-process account of learning in humans. *Neurobiology of Learning and Memory*, *108*, 185–195.

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, *141*, 302–320.

Milkman, K. L., Rogers, T., & Bazerman, M. H. (2008). Harnessing our inner angels and demons. *Perspectives on Psychological Science*, 3, 324–338.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions four general conclusions. *Current Directions in Psychological Science*, *21*, 8–14.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297–326.

Moshman, D. (2004). From inference to reasoning: The construction of rationality. *Thinking and Reasoning*, *10*, 221–239.

Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The role of mind-wandering in measurements of general aptitude. *Journal of Experimental Psychology: General*, *141*, 788–798.

Nigg, J. T. (2017). On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 58, 361–383.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, *67*, 130–159.

Over, D. E. (2000). Ecological rationality and its heuristics. *Thinking and Reasoning*, *6*, 182–192.

Over, D. E. (2004). Rationality and the normative/descriptive distinction. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 3–18). Malden, MA: Blackwell Publishing.

Parker, A. M., Bruine de Bruin, W., & Fischhoff, B. (2015). Negative decision outcomes are more common among people with lower decision-making competence: an item-level analysis of the Decision Outcome Inventory (DOI). *Frontiers in Psychology*, *6*, 363. https://doi.org/10.3389/fpsyg.2015.00363

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual differences approach. *Journal of Behavioral Decision Making*, *18*, 1–27.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72.

Perkins, D. N. (1995). *Outsmarting IQ: The emerging science of learnable intelligence*. New York: Free Press.

Phillips, W. J., Fletcher, J. M., Marks, A. D., & Hine, D. W. (2016). Thinking styles and decision making: A meta-analysis. *Psychological Bulletin*, *142*, 260–290.

Prado, J., & Noveck, I. A. (2007). Overcoming perceptual features in logical reasoning: A parametric functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *19*, 642–657.

Rand, D. G., Tomlin, D., Bear, A., Ludvig, E. A., & Cohen, J. D. (2017). Cyclical population dynamics of automatic versus controlled processing: An evolutionary pendulum. *Psychological Review*, 124(5), 626–642.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*, 566–594.

Samuels, R. (2005). The complexity of cognition: Tractability arguments for massive modularity. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind* (pp. 107–121). Oxford: Oxford University Press.

Samuels, R. (2009). The magical number two, plus or minus: Dual-process theory as a theory of cognitive kinds. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 129–146). Oxford: Oxford University Press.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162–173.

Sherman, J. W., Gawronski, B., & Trope, Y. (2014). *Dual-process theories of the social mind*. New York: Guilford Publications.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.

Siegel, H. (1997). *Rationality redeemed? Further dialogues on an educational ideal*. New York: Routledge.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*, 99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138.

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, *15*, 322–325.

Slovic, S., & Slovic, P. (2015). *Numbers and nerves: Information, emotion, and meaning in a world of data*. Corvallis: Oregon State University Press.

Stanovich, K. E. (1993). Dysrationalia: A new specific learning disability. *Journal of Learning Disabilities*, *26*, 501–515.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago: University of Chicago Press.

Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT: Yale University Press.

Stanovich, K. E. (2011). *Rationality and the reflective mind*. New York: Oxford University Press.

Stanovich, K. E. (2012). On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning. In K. Holyoak & R. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 343–365). New York: Oxford University Press.

Stanovich, K. E. (2013). Why humans are (sometimes) less rational than other animals: Cognitive complexity and the axioms of rational choice. *Thinking and Reasoning*, *19*, 1–26.

Stanovich, K. E. (2016). The Comprehensive Assessment of Rational Thinking. *Educational Psychologist*, *51*, 23–34.

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override, and mindware. *Thinking and Reasoning*, *24*, 423–444.

Stanovich, K. E., & Toplak, M. E. (2012). Defining features versus incidental correlates of Type 1 and Type 2 processing. *Mind and Society*, *11*, 3–13.

Stanovich, K. E., Toplak, M. E., & West, R. F. (2008). The development of rational thought: A taxonomy of heuristics and biases. *Advances in child development and behavior*, *36*, 251–285.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*, 342–357.

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*, 161–188.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645–726.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*, 672–695.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2016). *The Rationality Quotient: Toward a test of rational thinking*. Cambridge, MA: MIT Press.

Sternberg, R. J. (Ed.). (2002). *Why smart people can be so stupid*. New Haven, CT: Yale University Press.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. Cambridge: Cambridge University Press.

Sternberg, R. J., Grigorenko, E. L., & Zhang, L. F. (2008). Styles of learning and thinking matter in instruction and assessment. *Perspectives on Psychological Science*, *3*, 486–506.

Strathman, A., Gleicher, F., Boninger, D. S., & Scott Edwards, C. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology*, *66*, 742–752.

Taylor, S. E. (1981). The interface of cognitive and social psychology. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 189–211). Hillsdale, NJ: Lawrence Erlbaum Associates.

Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. New York: Norton.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Thompson, V. A., Prowse Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, *63*, 107–140.

Toates, F. (2005). Evolutionary psychology: Towards a more integrative model. *Biology and Philosophy*, *20*, 305–328.

Toates, F. (2006). A model of the hierarchy of behavior, cognition, and consciousness. *Consciousness and Cognition*, *15*, 75–118.

Toplak, M., Liu, E., Macpherson, R., Toneatto, T., & Stanovich, K. E. (2007). The reasoning skills and thinking dispositions of problem gamblers: A dual-process taxonomy. *Journal of Behavioral Decision Making*, *20*, 103–124.

Toplak, M. E., & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, *94*, 197–209.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics and biases tasks. *Memory and Cognition*, *39*, 1275–1289.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014a). Assessing miserly processing: An expansion of the Cognitive Reflection Test. *Thinking and Reasoning*, *20*, 147–168.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014b). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology*, *50*, 1037–1048.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2017). Real-world correlates of performance on heuristics and biases tasks in a community sample. *Journal of Behavioral Decision Making*, *30*, 541–554.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton: Princeton University Press.

Zeidner, M., & Matthews, G. (2000). Intelligence and personality. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 581–610). New York: Cambridge University Press.

47 Intelligence and Wisdom

Judith Glück

How is wisdom related to intelligence? Is wisdom a form of intelligence? Is intelligence a form of wisdom? The answers to these questions obviously depend on how one conceptualizes both wisdom and intelligence. In this chapter, I review both the empirical evidence and the prevalent theoretical positions on the relationship between wisdom and intelligence. In short, I suggest that wisdom includes aspects of intelligence but is far more than "just" intelligence. Wisdom integrates the ability to think about complex issues in a complex way with certain personality facets such as openness to experience and empathy with others, motivational facets such as a deep curiosity about the fundamental questions of the human existence as well as a willingness to critically reflect on oneself, and ethical facets such as a concern for a greater good that is strong enough to overrule more egoistical or self-enhancing goals (e.g., Ardelt, 2003; Sternberg, 1998; Weststrate & Glück, 2017a). Together, these different components lead individuals to acquire a strong body of broad and deep knowledge about human existence that makes these individuals a source of advice on how to live a good life (Baltes & Staudinger, 2000).

Obviously, the question of how wisdom is related to intelligence cannot be answered independently of how one defines wisdom – and, of course, of how one defines intelligence but I am not going to give a definition of intelligence in a handbook of intelligence. Rather, I am going to refer to various forms and facets of intelligence, including cross-references to the respective chapters. Different conceptions of wisdom differ in the emphases they put on cognitive aspects and, as one would expect, more cognition-oriented measures of wisdom turn out to have stronger empirical relationships with intelligence. This chapter is structured roughly chronologically, showing how concepts of intelligence have played a central role in the development of psychological models of wisdom and how our thinking about the relationship between intelligence and wisdom has evolved as wisdom research has become a larger and more diverse field. On the way, I also review the (surprisingly scarce) evidence on empirical relationships between wisdom and intelligence. At the end, I note some general issues of importance in the study of the complex relationship between wisdom and intelligence.

A Brief History of Wisdom and Intelligence

What is wisdom? A first glance at the psychological literature may leave readers with the impression that there are as many conceptions of wisdom as there are wisdom researchers. Wisdom has been defined, for example, as a form of expertise (Baltes & Staudinger, 2000), a personality type (Ardelt, 2003), self-transcendence (Levenson et al., 2005), or a type of practical intelligence combined with an ethical orientation (Sternberg, 1998). After twenty years of studying wisdom, however, I believe that the various definitions have more in common than one might think – they focus on different aspects of wisdom but essentially they may all be looking at the same thing. Wisdom is a complex, multifaceted construct that consists of interrelated cognitive and noncognitive components.

The Berlin Wisdom Model: Wisdom as Expertise About Life

The historically first large-scale endeavor to define and measure wisdom emerged from a domain of intelligence research, as it defined wisdom as expert knowledge. The Berlin Wisdom Model was developed by Paul Baltes and his colleagues at the Max Planck Institute for Human Development in Berlin in the 1980s (e.g., Baltes & Smith, 1990; Smith & Baltes, 1990). At that time, many intelligence researchers were interested in expertise, studying how experts differ from novices in the ways they reason about problems in domains such as chess or mathematics (e.g., Ericsson, Krampe, & Tesch-Römer, 1993; for an overview, see Ackerman, Chapter 48, this volume). Expertise is in-depth knowledge that develops through intense long-term training aimed at optimizing one's performance in a domain. An often-cited rule of thumb says that at least 10,000 hours of deliberate practice in any field are necessary to make a person an expert (Ericsson et al., 1993). Obviously, investing this much time and effort into one activity also requires certain preconditions such as sufficient motivation, self-confidence, and material and social resources, such as competent coaches and supportive family members. Recent research suggests that aspects of both fluid and crystallized intelligence are also quite relevant to the development of expertise (e.g., Hambrick et al., 2014; see Ackerman, Chapter 48, this volume).

Applying this logic to wisdom, Baltes and colleagues argued that wisdom is expertise in the fundamental pragmatics of life. The fundamental pragmatics of life are those issues of human existence that many of us struggle with sometimes. For example, as the only animal species that is consciously aware of its own mortality, how do we deal with knowing that we are going to die? How can we balance our own needs and those of others or intimacy and autonomy in our close relationships? How can we negotiate conflicts between individuals, groups, or societies that result from fundamentally different worldviews? How can we overcome aspects of human nature that evolved to make us fit for a very different environment from the one we live in today? The Berlin group argued that people who are deeply interested in such questions and have enough opportunities to "practice" dealing with them are likely to develop expert knowledge in life matters over time and that this expertise is equivalent to wisdom.

Based on this conception, Baltes and colleagues devised the first method to actually measure wisdom. The approach is still one of the most established and respected approaches today. Expertise researchers often used think-aloud tasks to study how experts differed from novices in their reasoning about problems. In the

Berlin Wisdom Paradigm (BWP), participants are asked to think aloud about difficult life problems such as "Somebody gets a phone call from a good friend. The friend says that he cannot go on any more, and that he has decided to commit suicide" or "A 15-year old girl absolutely wants to move out of her family home immediately" (Glück & Baltes, 2006, p. 682). Study participants are asked to think aloud about what one could consider and do in these situations. Participants' thoughts are recorded and transcribed and then evaluated with respect to five criteria derived from the conceptualization of wisdom as expertise. First, experts in any field have extensive *factual and procedural knowledge* about their area of expertise. Thus, these two forms of knowledge became the first two criteria. How much does the participant know, for example, about possible reasons for suicide and about ways to deal with suicidal individuals? How much do they know about the life situations and needs of fifteen-year-old girls? A wise person would be expected to have accumulated more knowledge about these issues than someone whose area of expertise is quantum mechanics or sports. The other three criteria were called meta-criteria because they are less about the specific content of knowledge than about the way participants thought about the problems. *Value relativism* means that the participant is aware of the different values, worldviews, and priorities that people can have, both within and across cultures, and that they accept and value these differences. In other words, wise individuals do not believe that their own beliefs are right and those of others are wrong – they see how different backgrounds and experiences can lead people to have different and equally legitimate worldviews. A wise individuals might, for example, have strong religious beliefs about suicide but they would be aware and accepting of the fact that their friend does not share these beliefs. At the same time, a wise person would, of course, believe that certain ethical values are nonnegotiable. *Life span contextualism* means that the participant interprets others' behavior in the light of their particular context – life phases, cultural and historical background, social and material resources, and so on. For example, a wise person would not simply interpret the fifteen-year-old girl's desire to leave her home as one of the typical stupidities of adolescents but consider the possibility of a different cultural background, the needs of people in this life phase, and possible reasons concerning her family or her friends. In other words, wise people are aware of the extent to which people's acts are the product of their situations and not just their personality. *Recognition and management of uncertainty* means that a participant is fully aware of the limited knowledge and control we have about future developments. Wise individuals know how much they do not know. Therefore, a wise person will not simply state what needs to be done in a situation – they will be aware of the unpredictability and uncontrollability of human life. The three meta-criteria are conceptually interesting in that they are, on the one hand, knowledge – wise people have learned that values and contexts differ and that many things are more uncertain than we know. On the other hand, they are relatively general ways of experiencing life and thinking about it – knowledge translated into a general mindset.

To evaluate the amount of wisdom in participants' responses to the BWP vignettes, a total of ten independent raters (two per criterion) are carefully trained to judge the transcripts on a scale from 1 = "very little similarity to an ideal response" to

7 = "very strong similarity to an ideal response." The average across the ten ratings is the participant's wisdom score. The BWP has been shown to have adequate reliability and validity in several studies (for an overview, see Baltes & Staudinger, 2000; Kunzmann, 2019).

The BWP: Relationships with Intelligence. Staudinger, Lopez, and Baltes (1997) investigated how BWP scores were correlated to various measures of personality, intelligence, and what they called the intelligence-personality interface. The study included two measures of fluid intelligence (a matrices-based measure of figural inductive reasoning and a digit-symbol substitution test measuring processing speed) and two measures of crystallized intelligence (a vocabulary test and a test of practical knowledge). The intelligence-personality interface includes capacities that combine cognitive facets with aspects of personality and motivation: In this study, they were social intelligence, creativity, and Sternberg's (1997) thinking styles. In an age-diverse sample of 125 participants, BWP scores were significantly correlated with inductive reasoning (r = 0.29), vocabulary (r = 0.34), and practical knowledge (r = 0.24) but not processing speed. Wisdom was also positively correlated with creativity (r = 0.37) and a judicial thinking style (r = 0.25) and negatively with progressive, conservative, and oligarchic thinking styles. The authors concluded that intelligence was a necessary but not sufficient condition for wisdom – that is, wise people are likely to be quite intelligent but not all highly intelligent people are also wise.

Mickler and Staudinger (2008) found similar relationships between the BWP and intelligence – $r = 0.17$ with inductive reasoning and $r = 0.33$ with vocabulary. Interestingly, in a study in my own lab, we did not find any significant relationships between the BWP and intelligence, with correlations of only $r = 0.12$ for vocabulary and $r = 0.10$ for inductive reasoning (Glück et al., 2013). One possible reason is that the variance in that study was somewhat restricted because half of the participants were wisdom nominees. In our most recent study with a mostly unselected sample, we did find significant correlations of about the same magnitude as in the Berlin studies ($r = 0.28$ with verbal analogies, $r = 0.18$ with figural inductive reasoning, and $r = 0.22$ with vocabulary).

Sternberg's Conception of Wisdom as Practical Intelligence Plus Ethics

Another researcher who took up the concept of wisdom early on was Robert J. Sternberg. In 1985, he published a highly influential study that looked at laypeople's conceptions, so-called implicit theories, of wisdom (Sternberg, 1985a). In a series of studies, he investigated people's implicit theories of wisdom, creativity, and intelligence. He found that wisdom and intelligence were more similar to one another in people's minds than both were to creativity. Wisdom and intelligence had some overlap in that both involved aspects of problem-solving and knowledge but wisdom involved a far broader range of attributes including concern for others, self-reflectivity, openness, and a general orientation at learning from life, including from one's own mistakes. Another interesting finding from Sternberg's study was that people from different professional fields differed considerably in their conceptions. For example, philosophy professors were the only group that considered intelligence as more similar to creativity than to wisdom and business professors were the only group for whom wisdom was negatively correlated to creativity. As I discuss in the section "Issues for Future Research," other research has since confirmed the important idea that wisdom conceptions have a common core but also marked differences both between cultures and between subcultures within our "Western" societies.

In 1998, Sternberg first published his "Balance Theory of Wisdom" (see also Sternberg, 2019), which became quite influential especially in applied fields such as leadership and education. Similar to the Berlin Wisdom Model, Sternberg's theory originated from a conception of intelligence. According to Sternberg (1998), the core element of wisdom is tacit knowledge "about oneself, others, and situational contexts" (Sternberg, 1998, p. 351). Tacit knowledge is procedural, action-oriented knowledge that is not necessarily put into words. It helps people to attain the goals that they value and it is typically acquired without help from others – that is, it is not learned at school or university but in the concrete contexts of people's personal and professional lives. Tacit knowledge in professional contexts comes with time spent on a job but it seems to be more dependent on what the person concludes from experiences than on having had the experiences per se. Tacit knowledge predicts job performance beyond conventional academic intelligence. Tacit knowledge is a central part of practical intelligence (see Hedlund, Chapter 30, this volume), which Sternberg and colleagues (e.g., Sternberg & Wagner, 1993; Sternberg, Wagner, & Okagaki, 1993) defined as the ability to utilize basic information-processing capacities in real-life contexts in order to adapt to the environment, shape it, or select a new environment. In his triarchic theory of intelligence, Sternberg (1985b) distinguishes practical intelligence from analytic and creative intelligence. Analytic intelligence is typically used in academic contexts or intelligence tests – it is what we need to perform decontextualized and abstract but relatively familiar tasks. Creative intelligence is needed in novel situations where one's set routines are not applicable. Practical intelligence is required to solve the highly contextualized problems that we encounter in our personal or professional lives – How do I deal with a conflict between my parents? How do I deal with a journal editor who seems to hate my paper although the reviewers like it?

As a form of practical intelligence, wisdom is called for in complex real-life contexts. Typically, wisdom-requiring situations are uncertain, complex, and characterized by conflicting interests. For example, someone might be faced with a decision between job opportunities that have different implications concerning her professional goals, her desire to balance family and work, her partner's career, her partner's and children's needs, and her desire to achieve something positive for a larger group of people. Wise individuals apply their tacit knowledge by balancing multiple intrapersonal, interpersonal, and extrapersonal interests. The overarching goal they pursue in searching for a balance is a common good – that is, they do not aim to achieve the optimal outcome for themselves or for another person but to strike an optimal balance between all interests involved. In this sense, wisdom is a special form of practical intelligence: While practical intelligence can be used toward any goal, including manipulating people for one's own profit, wisdom involves an ethical dimension in that it puts the common good over one's own interests. Wisdom also balances different ways of dealing with the environment, as explained earlier in this section. For example, the woman's job decision might be to try to change the requirements of her current job somewhat to better suit her family's needs and to enable her to support a charity in her free time and, if that fails, to change jobs and start working for a nongovernmental organization. Of course, it is a difficult question how the common good can be maximized in a particular context, as it may require weighing the needs of different individuals or groups against one another. The main characteristic of wisdom, however, is that the wise person is *oriented* toward that common good and at least tries to find an optimal balance. In later research, Sternberg has taken the concept of wisdom to applied fields such as leadership (e.g., Sternberg, 2007) and education (e.g., Sternberg, 2001). With respect to education, one important claim he has made is that schools tend to teach for intelligence and knowledge but not for wisdom (e.g., Sternberg, Reznitskaya, & Jarvin, 2007). He and his colleagues have developed approaches to teach students at different ages competencies that are likely to lead to wisdom (e.g., Reznitskaya & Sternberg, 2004).

In sum, Sternberg's balance theory is quite compatible with the BWP. One could argue that tacit knowledge is a form of expert knowledge, as both are acquired over time as the individual accumulates experience in a specific domain. One important difference is that Sternberg conceptualizes wisdom as manifest in the relation between a person and a context and that he focuses on the outcome or product – a balanced solution to a complex, contextualized problem – rather than on the process of thinking about the problem. In this sense, the contextualized, relativistic way of thinking about life problems that the Berlin model describes may lead to well-balanced solutions as described by Sternberg. One important difference is that the balance theory puts an explicit focus on ethical aspects, as achieving a common good is viewed as the main outcome of wise reasoning. The Berlin model is largely silent about ethical aspects of wise thinking, although Kunzmann and Baltes (2003) found significant correlations between BWP scores and values oriented toward a common good (e.g., societal engagement and environmental protection). Still, Sternberg's model stands out as the one that most clearly includes an ethical orientation as crucial for wisdom (see also Sternberg & Glück, 2019).

"Noncognitive" Conceptualizations of Wisdom

For a long time, the Berlin Wisdom Model and the measurement paradigm associated with it essentially dominated empirical wisdom research. In 2004, however, Monika Ardelt published a rather fundamental critique of the Berlin approach. Ardelt did not say that the Berlin Wisdom Model was wrong with respect to wisdom but she argued that it did not capture the essence of wisdom. According to Baltes and colleagues (e.g., Baltes & Staudinger, 2000), wisdom, in the sense of the Berlin model, was not an attribute of individuals – rather, there was a body of collective, culturally based knowledge "out there" of which an individual could acquire more or less through various forms of learning, including reading philosophical books or observing others' wise behavior. Ardelt, however, argued that wisdom is deeply situated inside the individual. She distinguished between theoretical knowledge that a person intellectually understands and experiential knowledge that a person has internalized because they have acquired it through personal, transformative experiences. "Wisdom needs to be realized through a reflection on personal experiences that transform the individual in the process" (Ardelt, 2004, p. 260). In other words, we all have access to a lot of knowledge available in the form of proverbs, sayings, or, nowadays, websites but there is an enormous difference between sagely saying "There are things in life that you simply have to accept" to a friend who has lost the love of his life and accepting the loss of your own love.

As an alternative to knowledge-based models of wisdom, Ardelt (2003, 2004) has suggested conceptualizing wisdom as a certain specific personality constellation that integrates three personality dimensions. These dimensions were first identified in a seminal early study of laypeople's wisdom conceptions by Clayton and Birren (1980). Ardelt defined the cognitive dimension of wisdom not as a certain way of thinking or a certain type of knowledge but rather as a deep-seated desire to understand life and to gain insights about the meanings of experiences, especially with respect to intra- and interpersonal aspects. This intense curiosity is likely to lead people to acquire the type of knowledge conceptualized in the Berlin model, including acceptance of the difficult sides and the uncertainties of the human existence. However, Ardelt believes the essence of wisdom is not in the knowledge itself but in the personality that drives its acquisition. The reflective dimension is the willingness and ability to look at phenomena from multiple perspectives. Notably, this includes distancing oneself from one's own personal viewpoint and taking a self-examining perspective, which helps the individual overcome subjectivity and projections. The affective component is defined as "sympathetic and compassionate love for others" (Ardelt, 2004, p. 275). Wise people care about others and their concern for the well-being of others is not limited to their personal friends or family.

To measure wisdom, Ardelt (2003) developed a self-report scale that assesses all three of the personality dimensions. The Three-Dimensional Wisdom Scale (3D-WS) consists of statements such as "There is only one right way to do anything" for the cognitive dimension, "I would feel much better if my present circumstances changed" for the reflective dimension, and "I am annoyed by unhappy people who just feel sorry for themselves" for the affective dimension. As in most self-report scales, participants indicate their agreement to each of these statements on five-point scales. As the examples show, many items are reverse-coded – a wise person is expected to disagree with them.

Although the 3D-WS has been used in many studies, few have looked at relationships with intelligence. Glück and colleagues (2013) found a correlation of $r = 0.22$ with inductive reasoning and no significant correlation with vocabulary. Specifically, inductive reasoning was related to the cognitive ($r = 0.31$) and the reflective ($r = 0.29$) dimensions but not the affective dimension. Thus, the willingness to think in a deep and complex way about issues and to question one's own perspectives is related to fluid intelligence, although the causality of this relationship is unclear. That study also found a correlation of 0.25 between the 3D-WS and the BWP, suggesting that the two measures share some but not a lot of variance.

In sum, Ardelt's approach conceptualizes wisdom as a lot more than cognition. Wisdom clearly involves certain ways of thinking – trying to get beyond the surface and understand complex issues in depth, trying to take different perspectives on issues and to understand one's own role in events – and these ways of thinking certainly require a certain amount of intelligence. But other aspects are more important in Ardelt's conception – particularly a personality that is more interested in understanding the deep truths of life than in protecting the person's self-esteem or making a positive impression on others and that cares deeply about the well-being of others.

Other authors have also developed conceptions and measures of wisdom that emphasize a certain attitude toward life experiences and other people more than knowledge or reasoning. Jeffrey Dean Webster (2003, 2007, 2019) defines wisdom as the motivation and ability to learn from critical life experiences and to apply this knowledge so as to enhance the well-being of oneself and others. Thus, again, wisdom lies not primarily in the knowledge that has been acquired but rather in the motivation that leads people to acquire it and to use it for positive purposes. According to Webster, wisdom entails five facets: critical life experiences that are necessary for wisdom to be gained, openness to explore new and different experiences, reminiscence and reflectiveness in order to think deeply about experiences and learn from them, emotion regulation competencies to identify and manage emotions as difficult situations require, and humor, as a form of mature, nondefensive coping with difficult experiences. Webster (2003, 2007) developed the Self-Assessed Wisdom Scale (SAWS), which measures his five components of wisdom by self-report. In our study of different wisdom measures, the SAWS had correlations of $r = -0.15$ with inductive reasoning and $r = 0.18$ with vocabulary. It also had a correlation of $r = 0.23$ with the BWP and $r = 0.26$ with the 3D-WS (Glück et al., 2013).

Michael R. Levenson and colleagues (Levenson et al., 2005; Aldwin, Igarashi, & Levenson, 2019) based their model of wisdom on theories of identity development and on transcultural philosophical conceptions of wisdom (Curnow, 1999). They define wisdom as self-transcendence – a mature sense of self that is not dependent on external enhancements. There are four main facets of self-transcendence that may form a developmental sequence: self-knowledge, or an awareness of the core facets as well as the more external sources of one's sense of self; nonattachment, or awareness of the transient nature of external sources of self; integration and acceptance of unwanted or conflicting characteristics as part of one's self rather than a threat to one's self-esteem; and self-transcendence itself, which means being independent of external sources of self-enhancement. Self-transcendence enables people to truly and nonselfishly care about others and to feel that they are part of something larger than themselves, that is, to overcome rigid boundaries between the self and the external world. Levenson et al. (2005; see also Aldwin et al., 2019; Koller, Levenson & Glück, 2017) developed the Adult Self-Transcendence Inventory (ASTI) as a self-report measure of self-transcendence. In our 2013 study, the ASTI had insignificant correlations with both inductive reasoning and vocabulary but it had the highest correlations with other measures of wisdom (BWP: $r = 0.30$, 3D-WS: $r = 0.58$, SAWS: $r = 0.50$). These findings suggest that the ASTI captures common variance across measures of wisdom that is unrelated to intelligence (Glück et al., 2013).

Like Ardelt, both Webster and Levenson and colleagues primarily define wisdom as an attitude – a way of being in the world, of experiencing events and reflecting on these experiences – and not primarily as a cognitive competence as in the Berlin Wisdom Model or as in Sternberg's balance theory. It seems likely, however, that wisdom is both – that the noncognitive and cognitive aspects of wisdom are interrelated. The desire to learn from life that Ardelt describes and the reflective, open attitude that both Ardelt and Webster consider as central are likely to lead to actual learning, that is, to the acquisition of tacit and nontacit, factual, and procedural knowledge and the meta-knowledge that Baltes and colleagues or Sternberg and colleagues have described. Self-transcendence is related to the acceptance of differences between people and of the inherent uncertainty of the human existence included in the Berlin Wisdom Model. One facet that clearly goes beyond learning and knowledge is the ethical dimension of concern for others and the world at large that is evident in Sternberg's and Ardelt's conceptions and also related to self-transcendence. In other words, wisdom involves aspects of thinking and knowledge – I will discuss later, in the section "An 'Investment Theory' of Wisdom and Intelligence," how these aspects fit with conceptions of intelligence – but it clearly goes beyond intelligence in that it includes a certain attitude toward life and other people.

The Distinction Between Self-Related and General Wisdom

As explained earlier, Ardelt (2004) argued that conceptions of wisdom as general, abstract life knowledge do not capture the experiential, personal qualities of wisdom. Ursula Staudinger, a collaborator in the Berlin wisdom project from early on, followed a similar line of thinking when she proposed a distinction between general and personal wisdom (Mickler & Staudinger, 2008; Staudinger, Dörner, & Mickler, 2005; Staudinger, 2019). She noted that people often think and act much more wisely when they are dealing with other people's problems than when the same problems occur in their own life. Staudinger and colleagues proposed a conception of personal wisdom that conceptualizes wisdom as "sound judgment and deep insight with regard to difficult and uncertain matters of one's own life" (Staudinger, 2019). The so-called Bremen Wisdom Paradigm (BrWP; Mickler & Staudinger, 2008) measures personal wisdom by means of an interview about the participants' strengths and weaknesses as a friend. Utilizing a procedure parallel to that of the BWP, five criteria are used to assess personal wisdom. As in the BWP, there are two basic criteria – rich self-knowledge and heuristics of growth and self-regulation. The three meta-criteria are interrelating the self (being able to reflect on and understand the internal or external causes of one's feelings or behavior), self-relativism (being able to take a distanced, unbiased view on oneself without being overly self-critical), and tolerance of ambiguity (acceptance and management of the uncertainties in one's life). Mickler and Staudinger (2008) found correlations of $r = 0.30$ with inductive reasoning and $r = 0.24$ with vocabulary for the BrWP. Its correlation with the BWP was $r = 0.48$. (Correlations with self-report measures of wisdom are not available for the BrWP yet.) In sum, Staudinger's work on the BrWP suggests that there is indeed a difference between wisdom-related knowledge about life in general and about one's own life, although the two are related. Staudinger and colleagues have argued that it is easier to gain general wisdom than to apply the insights one has about other people and life in general to oneself and one's own life (Mickler & Staudinger, 2008; Staudinger, 2019). This idea underscores the importance of noncognitive facets of wisdom: When a problem concerns our own life, we may not be able to access or apply knowledge or thinking strategies that we utilize perfectly well in other situations.

Grossmann's Research on Wise Reasoning in Context

This idea is supported by recent research by Igor Grossmann and colleagues, who explicitly study wise reasoning. Rather than measuring wisdom as a trait-like quality of persons, they look at how different contexts influence the way people think about difficult life problems. Based on the Berlin Wisdom Model and other relevant research, Grossmann identified characteristics of wise reasoning such as intellectual humility (recognizing the limitations of one's knowledge), searching for compromise, awareness of uncertainty and change, and consideration of broader contexts and different perspectives (Grossmann, 2017, p. 236) and developed a Wise Reasoning Paradigm (WRP) to measure them (for an overview, see Oakes et al., 2019). In a series of experimental studies, Grossmann and colleagues have shown that people reason more wisely if they take a decentered perspective on a problem than if they are strongly focused on their personal viewpoint. For example, people show higher levels of wise reasoning if they imagine that a friend has been cheated on by their partner than if they imagine being in that situation themselves (Grossmann & Kross, 2014). They reason more wisely about US politics if they imagine living in Iceland than living in the United States (Kross & Grossmann, 2012). Even thinking about a problem in third-person language seems to make people reason more wisely than thinking in first-person language (Grossmann & Kross, 2014).

In addition to the specific finding that decentered thinking fosters wisdom, Grossmann and colleagues have demonstrated that wisdom generally varies by situational context. In other words, most of us are sometimes very wise and sometimes very unwise, and mostly we are somewhere in between (Grossmann, 2017; Grossmann, Gerlach, & Denissen, 2016). These findings are interesting with respect to the relationship between wisdom and intelligence because they, again, emphasize the role of noncognitive aspects for wisdom: If wisdom were only a set of intellectual competencies, using them should not be highly dependent on a particular mindset or personality. But wise thinking is not always equally easy. When we are panicking or steaming with fury or when we are absolutely certain that we are right and everyone else is wrong (Sternberg, 2005), we are unlikely to question our own intuitions, weigh different aspects of a situation carefully, and consider the arguments of people who disagree with us. In other words, we are unlikely to reason wisely. Of course, people obviously differ in the extent to which they have these wisdom-related capabilities in the first place. Thus, people differ in the wisdom-related knowledge and ways of thinking that they have acquired over time and their access to these resources varies by situations. A highly wise person may have acquired more of the relevant knowledge and reasoning strategies than other people and may also more often be in a state of mind that allows them to utilize that knowledge.

Wise Reasoning and Intelligence. In one study that is particularly relevant to the topic of this chapter, Grossmann and colleagues (2013) showed that wise reasoning was related to well-being, whereas classical measures of intelligence were not. They argued that wise reasoning is learned through personal insights gained from life experiences. Therefore, wise reasoning should be related to how individuals live and evaluate their own lives, whereas abstract intelligence measures should not be related to personal experience. Grossmann and colleagues measured wise reasoning by collecting participants' oral responses to vignettes about real-life intergroup and interpersonal conflicts. Transcripts were rated for six aspects of wise reasoning. Concerning intelligence, measures of processing speed, working memory, comprehension, and vocabulary were collected, as were various measures of well-being and personality. Results showed significant correlations in the 0.20–0.30 magnitude between wise reasoning and five out of six measures of well-being. Intelligence had few and unsystematic correlations with well-being and some of those that were found might be accounted for by age. These results suggest that wise reasoning is more related to people's self-knowledge and the way they live and evaluate their lives than general intelligence is. In other words, we can easily imagine a highly intelligent and highly unhappy person but a highly wise person would seem unlikely to be very unhappy (Weststrate & Glück, 2017b; Ardelt, 2011).

Grossmann and colleagues (2013) also reported correlations between wisdom and aspects of intelligence. Wise reasoning with respect to intergroup conflicts was negatively related to processing speed ($r = -0.25$), whereas wise reasoning with respect to both types of conflict was positively related to vocabulary ($r = 0.15$ for intergroup and $r = 0.31$ for interpersonal conflicts). These findings are consistent with the idea that wise reasoning is related to crystallized intelligence – a knowledge- and experience-based way of thinking about life – that increases with age. And, indeed, Grossmann and colleagues (2010) found that wise reasoning increased with age. Thus, again, wisdom is somewhat related to intelligence but it goes beyond intelligence in its relevance for people's real lives.

An "Investment Theory" of Wisdom and Intelligence

There are some rather obvious conclusions from the literature reviewed in this chapter but there are also a lot of open questions. The first conclusion is that wisdom is not a form of intelligence, nor is intelligence a form of wisdom. Even if one thinks of intelligence in a broad, contextualized way as, for example, in Sternberg and Wagner's (1993) conception of practical intelligence, some aspects of wisdom – an ethical orientation, compassion, self-transcendence – clearly fall outside the realm of intelligence. Second, however, wisdom clearly includes aspects of intelligence. What exactly these aspects are has not yet been sufficiently investigated: relatively few studies related wisdom to intelligence and those few included rather narrow facets of intelligence. To summarize the available evidence, inductive reasoning – used as a simple measure of fluid intelligence – was positively correlated to the 3D-WS, the BWP, and the BrWP and negatively correlated to the SAWS (Glück et al., 2013; Mickler & Staudinger, 2013; Staudinger, Lopez, & Baltes, 1997). Vocabulary – as an even simpler measure of crystallized intelligence – was positively correlated to the SAWS, the BWP, the BrWP, and the WRP (Glück et al., 2013; Grossmann et al., 2013; Mickler & Staudinger, 2013; Staudinger, Lopez, & Baltes, 1997).

These findings suggest that there are (at least) two intelligence-related components of wisdom. I would argue that the first one describes a certain way of thinking. Wise individuals are, as Ardelt (2003) has argued, deeply motivated to understand life, see through illusions, and identify the deeper reasons behind complex phenomena. The willingness and ability to think deeply about complex issues manifests itself in the 3D-WS as well as in the Berlin and Bremen wisdom paradigms, which use wisdom criteria referring to relativistic, contextualized, and uncertainty-conscious thinking. It makes sense that these measures of wisdom would be correlated with inductive reasoning. (One would expect such a relationship for Grossmann's wise reasoning method as well but inductive reasoning was not included in their study.) One could argue that this "complex thinking" form of wisdom manifests itself relatively early in life and continues to influence a person's learning about life across the life span. In other words, it could be viewed as a kind of fluid intelligence that precedes wisdom. The second component could be viewed as the crystallized intelligence part of wisdom: the experience-based in-depth knowledge about life that comes from long-term "practice" with the difficult questions of human life. This form of wisdom is what Webster's SAWS focuses on and it clearly also manifests itself in the BrWP, BWP, and WRP. All these measures are correlated with vocabulary, which is a (very crude!) indicator of life knowledge. A less favorable interpretation would be that measures based on verbal responses favor participants higher in verbal abilities. This method effect may certainly explain part of the relationships but it would not account for the correlation with the SAWS.

In sum, the model of the intelligence-related components of wisdom that I have sketched here is similar to investment theories of intelligence as proposed by Ackerman (Ackerman, 2000; von Stumm & Ackerman, 2013): Early fluid intelligence as well as certain interests and personality components predict the amount of crystallized intelligence that a person develops in a certain area. Some people are highly interested in the "big questions" of the human existence. If they are also sufficiently able to think in a complex way about complex issues and have certain personality qualities such as openness, empathy, self-reflectivity, or an ethical orientation, they are likely to develop wisdom-related knowledge as they navigate their own life challenges and observe those of other people. In the course of this development, they are likely to develop not only extraordinary knowledge but also other, noncognitive facets of wisdom such as emotion regulation skills and social competencies (Glück & Bluck, 2013; Glück, Bluck, & Weststrate, 2018; Weststrate & Glück, 2017a). Thus, wisdom involves a dynamic interaction between cognitive and noncognitive components. This interaction may also explain the situational and contextual variability of wisdom: While the knowledge and the thinking-strategy repertoire stored in our memories are not affected by short-term fluctuations, our access to that knowledge may be dependent on our mindset. When we are emotionally balanced, open-minded, and interested in an issue, we may utilize our knowledge about life as well as our skills at asking the right questions to gain important information, taking different perspectives on the issue, and giving advice to someone who may not want to hear it. In other situations, we may not be able to decenter from our own viewpoint. In other words, the cognitive capacities underlying wisdom do not vary by context but the noncognitive capacities that enable us to utilize them do.

One important aspect to add is that, while the correlations between wisdom and intelligence reviewed in this chapter suggest two distinct facets of intelligence that are relevant to the development of wisdom – complex thinking and in-depth knowledge – these two may be overlapping in real life. In fact, while Baltes and colleagues refer to value relativism, life span contextualism, and recognition and management of uncertainty as criteria for wisdom-related knowledge, one could also consider them as wisdom-related ways of reasoning, as Grossmann does. They are certainly forms of knowledge in that they are the result of learning: Wise individuals have realized that different people have different values grounded in their different cultural and individual backgrounds, that people's behavior is often shaped more by their contexts than by their personalities, and that we can neither predict nor control the future. The point of this knowledge, however, is that they translate it into the way they reason about life problems – taking different perspectives, contextual factors, and the many things they do not know into account. Thus, while complex thinking is a developmental predecessor of wisdom-related knowledge, wisdom-related ways of thinking about life problems are also part of the developmental outcome.

Issues for Future Research

As this review shows, relatively little is known about the relationship between wisdom and intelligence and much of what I said in this chapter is somewhat speculative. Most previous research has treated intelligence more or less as a control variable, limiting its assessment to narrow and decontextualized indicators of fluid and crystallized intelligence. Future research should be based on theoretical considerations about which aspects of intelligence should be closely, remotely, or even negatively related to wisdom. Somewhat "old-fashioned" cognitive constructs such as cognitive complexity (e.g., Bieri, 1955; Ceci & Liker, 1986), problem-solving (e.g., Hambrick, Burgoyne, & Altmann, Chapter 23, this volume; Sternberg, 1982), dialectical thinking (e.g., Basseches, 1984; Riegel, 1973), or, more recently, intellectual humility (e.g., Zachry et al., 2018; Leary et al., 2017) might be aspects of intelligence that are closely related to wisdom. Standard measures of intelligence such as inductive reasoning or vocabulary are more distantly related, as reviewed in the section on "An 'Investment Theory' of Wisdom and Intelligence." Aspects such as processing speed might even be negatively correlated to wisdom (e.g., Grossmann et al., 2013) and it would be interesting to see whether these relationships are accounted for by age or other variables.

A methodological note: The limitations of correlations. Staudinger and colleagues (1997) argued that intelligence is a necessary but not sufficient condition for wisdom and my literature review suggests that this idea is very convincing: Wisdom is intelligence *plus* other things. However, if this is true, correlations are not a good indicator for the relationship between wisdom and intelligence. If one thing is a necessary but not sufficient precondition for another, then the scatterplot of the two things should look like a triangle: People high in wisdom should also be high in intelligence but people high in intelligence could be high in wisdom or low or anything in between. As an example, Figure 47.1 shows a scatter plot of the relationship between wisdom, measured by the BWP, and a measure of inductive reasoning. The data are from a study by Glück and Baltes (2006), in which 318 participants from three age groups participated. In that study, the BWP had an age-corrected correlation of $r = 0.29$, $p < 0.001$, with inductive reasoning. As the figure shows, the relationship may be more consistent with the idea of a necessary but not sufficient condition than with the assumption of linearity underlying correlations. For a linear relationship, one would expect to find about the same (and small) number of cases in the upper left and lower right corner of the scatter plot – that is, equally few participants high in wisdom and low in intelligence as there are high in intelligence and low in wisdom. Figure 47.1, however, shows many cases in the lower right corner and relatively few in the upper left corner – that is, there are a considerable number of participants high in inductive reasoning and low in wisdom but far fewer showing the opposite pattern. I am not aware of a statistical parameter that indicates the "triangularity" of a relationship but I assume it would be significant here. About two-thirds of the sample were inside the triangle and one-third was outside it. In other words, the correlations that have been reported in earlier research may not provide optimal information about the actual relationship between wisdom and intelligence.

One potentially interesting question came up when I showed these data in the context of a talk on wisdom and giftedness (see Reis & Renzulli, Chapter 13, this volume). The giftedness researchers in the audience pointed out that, while they could see the triangular relationship I just described, they also saw that the participants at the top end of the intelligence distribution tended to have somewhat lower wisdom scores than those slightly below the top. In other words, the most intelligent participants in my sample might be somewhat less wise, on average, than those a bit less smart. While my sample is definitely not suited for investigating this issue, it might be an interesting question for further research whether extremely high levels of intelligence make it more difficult to develop wisdom.

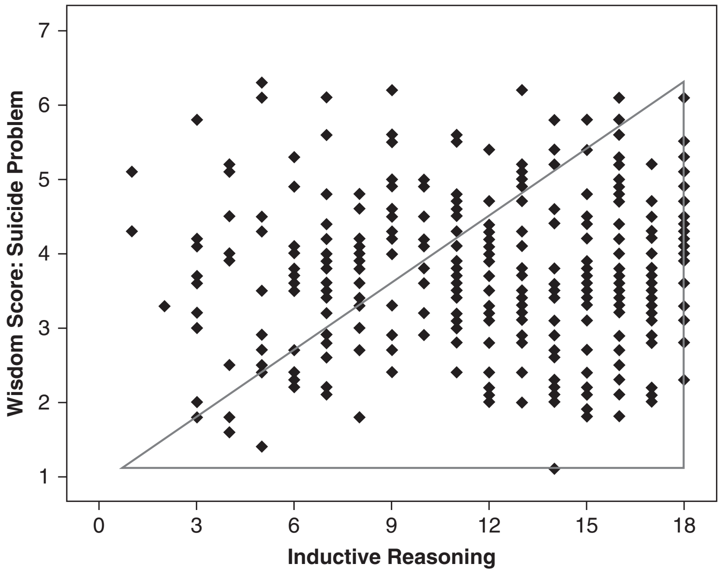

Figure 47.1 *Scatter plot of inductive reasoning with Berlin Wisdom Paradigm performance (data from Glück & Baltes, 2006).*

One important caveat concerns possible cultural differences in the relationship between wisdom and intelligence. Several studies have found that people in "non-Western" cultures associate wisdom far less with intelligence than do "Western" individuals (e.g., Takahashi & Overton, 2002; Yang, 2001). In other words, the role of intelligence for wisdom may be somewhat culture-specific. Little research has gone beyond investigating people's conceptions of wisdom by actually measuring wisdom in non-"Western" cultures and that little research mostly just used translations of "Western" measures of wisdom that may not be consistent with the respective culture's understanding of wisdom. The relationship between intelligence and wisdom has not been investigated outside "Western" cultures at all.

Conclusions

Wisdom is more than intelligence – it certainly requires intelligence but it also requires noncognitive capacities such as openness, empathy, self-reflectivity, and morality. Wisdom is also a developmental construct – no one is born with it. Wisdom develops (or does not develop) over our lifetime as we navigate the challenges and opportunities of life and try to make sense of them. Wisdom is something that today's world needs more than intelligence. The growth in intelligence over the twentieth century does not really seem to have made the world a better place (Sternberg, 2019). How can we increase the average level of wisdom in the world? This seems to be a highly important question for psychological research. Research has identified ways to reduce people's self-centeredness in psychological experiments

(Grossmann, 2017) – What do we need to also reduce it in the real world? How can we reduce the situational factors that tend to impede people's ability to access their wisdom-related knowledge? How can we teach for wisdom in schools and universities (Sternberg, 2001)? It would be wonderful if this chapter could motivate some researchers interested in intelligence to start thinking about wisdom as well.

References

Ackerman, P. L. (2000). Domain-specific knowledge as the "dark matter" of adult intelligence: Gf/Gc, personality and interest correlates. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *55*(2), P69-P84.

Aldwin, C. M., Igarashi, H., & Levenson, M. R. (2019). Wisdom as self-transcendence. In R. J. Sternberg & J. Glück (Eds.), *The Cambridge handbook of wisdom* (pp. 122–143). Cambridge: Cambridge University Press.

Ardelt, M. (2003). Development and empirical assessment of a three-dimensional wisdom scale. *Research on Aging*, *25*, 275–324.

Ardelt, M. (2004). Wisdom as expert knowledge system: A critical review of a contemporary operationalization of an ancient concept. *Human Development*, *47*, 257–285.

Ardelt, M. (2011). Wisdom, age, and well-being. In K. W. Schaie & S. L. Willis (Eds.), *Handbook of the psychology of aging* (7th ed., pp. 279–291). San Diego: Elsevier.

Baltes, P. B., & Smith, J. (1990). Toward a psychology of wisdom and its ontogenesis. In R. J. Sternberg (Ed.), *Wisdom: Its nature, origins, and development* (pp. 87–120). Cambridge: Cambridge University Press.

Baltes, P. B., & Staudinger, U. M. (2000). Wisdom: A metaheuristic (pragmatic) to orchestrate mind and virtue towards excellence. *American Psychologist*, *55*, 122–136.

Basseches, M. (1984). Dialectical thinking. Norwood, NJ: Ablex.

Bieri, J. (1955). Cognitive complexity-simplicity and predictive behavior. *The Journal of Abnormal and Social Psychology*, *51*(2), 263–268.

Ceci, S. J., & Liker, J. K. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General*, *115*(3), 255–266.

Clayton, V. P., & Birren, J. E. (1980). The development of wisdom across the life span: A reexamination of an ancient topic. In P. B. Baltes & J. O. G. Brim (Eds.), *Life-span development and behavior*, Vol. 3 (pp. 103–135). New York: Academic Press,

Curnow, T. (1999). *Wisdom, intuition and ethics*. Aldershot: Ashgate Publishing.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406

Glück, J., & Baltes, P.B. (2006). Using the concept of wisdom to enhance the expression of wisdom knowledge: Not the philosopher's dream, but differential effects of developmental preparedness. *Psychology and Aging*, *21*, 679–690.

Glück, J., & Bluck, S. (2013). MORE Wisdom: A developmental theory of personal wisdom. In M. Ferrari & N. Weststrate (Eds.), *The scientific study of personal wisdom* (pp. 75–98). New York: Springer.

Glück J., Bluck, S., & Weststrate, N. M. (2018). More on the MORE Life Experience Model: What we have learned (so far). *The Journal of Value Inquiry* (Advance online publication).

Glück, J., König, S., Naschenweng, K., Redzanowski, U., Dorner, L., Strasser, I., & Wiedermann, W. (2013). How to measure wisdom: Content, reliability, and validity of five measures. *Frontiers in Psychology*, *4*, https://doi.org/10.3389/fpsyg.2013.00405

Grossmann, I. (2017). Wisdom in context. *Perspectives on Psychological Science*, *12*(2), 233–257.

Grossmann, I., Gerlach, T. M., & Denissen, J. J. (2016). Wise reasoning in the face of everyday life challenges. *Social Psychological and Personality Science*, *7*(7), 611–622.

Grossmann, I., & Kross, E. (2014). Exploring Solomon's Paradox: Self-distancing eliminates the self-other asymmetry in wise reasoning about close relationships in younger and older adults. *Psychological science*, *25*(8), 1571–1580.

Grossmann, I., Na, J., Varnum, M. E. W., Kitayama, S., & Nisbett, R. E. (2013). A route to well-being: Intelligence versus wise reasoning. *Journal of Experimental Psychology: General*, *142*, 944–953.

Grossmann, I., Na, J., Varnum, M. E. W., Park, D. C, Kitayama, S., & Nisbett, R. E. (2010). Reasoning about social conflicts improves into old age. *PNAS Proceedings of the National Academy of Sciences*, *107*, 7246–7250.

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34–45.

Koller, I., Levenson, M. R., & Glück, J. (2017). What do you think you are measuring? A mixed-methods procedure for assessing the content validity of test items and theory-based scaling. *Frontiers in Psychology*, *8*, 126.

Kross, E., & Grossmann, I. (2012). Boosting wisdom: Distance from the self enhances wise reasoning, attitudes, and behavior. *Journal of Experimental Psychology: General*, *141*, 43–48.

Kunzmann, U. (2019). Performance-based measures of wisdom: state of the art and future directions. In R. J. Sternberg & J. Glück (Eds.), *The Cambridge handbook of wisdom* (pp. 277–296). Cambridge: Cambridge University Press.

Kunzmann, U., & Baltes, P. B. (2003) Wisdom-related knowledge: Affective, motivational, and interpersonal correlates. *Personality and Social Psychology Bulletin*, *29*(9), 1104–1119.

Leary, M. R., Diebels, K. J., Davisson, E. K., Jongman-Sereno, K. P., Isherwood, J. C., Raimi, K. T., Deffler, S. A., & Hoyle, R. H. (2017). Cognitive and interpersonal features of intellectual humility. *Personality and Social Psychology Bulletin*, *43*(6), 793–813.

Levenson, R., Jennings, P. A., Aldwin, C., & Shiraishi, R.W. (2005). Self-transcendence, conceptualization and measurement. *International Journal of Aging and Human Development*, *60*, 127–143.

Mickler, C., & Staudinger, U. M. (2008). Personal wisdom: Validation and age-related differences of a performance measure. *Psychology and Aging*, *23*, 787–799.

Oakes, H., Brienza, J. P., Elnakouri, A., & Grossmann, I. (2019). Wise reasoning: Converging evidence for the psychology of sound judgment. In R. J. Sternberg & J. Glück (Eds.), *The Cambridge handbook of wisdom* (pp. 202–225). Cambridge: Cambridge University Press.

Reznitskaya, A., & Sternberg, R. J. (2004). Teaching students to make wise judgments: The "teaching for wisdom" program. In P. A. Linley & S. Joseph (Eds.), *Positive psychology in practice* (pp. 181–196). New York: Wiley.

Riegel, K. F. (1973). Dialectic operations: The final period of cognitive development. *Human development*, *16*(5), 346–370.

Smith, J., & Baltes, P. B. (1990). Wisdom-related knowledge: Age/cohort differences in responses to life-planning problems. *Developmental Psychology*, *26*, 494–505.

Staudinger, U. M. (2019). General and personal wisdom. In R. J. Sternberg & J. Glück (Eds.), *The Cambridge handbook of wisdom* (pp. 182–201). Cambridge: Cambridge University Press.

Staudinger, U. M., Dörner, J., & Mickler, C. (2005). Wisdom and personality. In R. J. Sternberg & J. Jordan (Eds.), *A handbook of wisdom: Psychological perspectives* (pp. 191–219). New York: Cambridge University Press.

Staudinger, U. M., Lopez, D., & Baltes, P. B. (1997). The psychometric location of wisdom-related performance: Intelligence, personality, and more? *Personality and Social Psychology Bulletin*, *23*, 1200–1214.

Sternberg, R. J. (1982). Reasoning, problem solving, and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 225–307). Cambridge: Cambridge University Press.

Sternberg, R. J. (1985a). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, *49*, 607–627.

Sternberg, R. J. (1985b). Beyond IQ: A triarchic theory of human intelligence. New York: Cambridge University Press.

Sternberg, R. J. (1997). Thinking styles. Cambridge: Cambridge University Press.

Sternberg, R. J. (1998). A balance theory of wisdom. *Review of General Psychology*, *2*, 347–365.

Sternberg, R. J. (2001). Why schools should teach for wisdom: The balance theory of wisdom in educational settings. *Educational Psychologist*, *36*(4), 227–245.

Sternberg, R. J. (2005). Foolishness. In R. J. Sternberg & J. Jordan (Eds.), *A handbook of wisdom: Psychological perspectives*. (pp. 331–351). Cambridge: Cambridge University Press.

Sternberg, R. J. (2007). A systems model of leadership: WICS. *American Psychologist*, *62*(1), 34–42.

Sternberg, R. J. (2019). Why people often prefer wise guys to guys who are wise: An augmented balance theory of the production and reception of wisdom. In R. J. Sternberg & J. Glück (Eds.), *The Cambridge handbook of wisdom* (pp. 162–181). Cambridge: Cambridge University Press.

Sternberg, R. J., & Glück, J. (2019). Wisdom, morality, and ethics. In R. J. Sternberg & J. Glück (Eds.), *The Cambridge handbook of wisdom* (pp. 551–574). Cambridge: Cambridge University Press.

Sternberg, R. J., Reznitskaya, A., & Jarvin, L. (2007). Teaching for wisdom: What matters is not just what students know, but how they use it. *London Review of Education*, *5*(2), 143–158.

Sternberg, R. J., & Wagner, R. K. (1993). The g-ocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science*, *2*, 1–5.

Sternberg, R. J., Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: The nature and role of tacit knowledge in work and at school. In H. Reese & J. Puckett (Eds.), *Advances in lifespan development* (pp. 205–227). Hillsdale, NJ: Lawrence Erlbaum Associates.

Takahashi, M., & Overton, W. F. (2002). Wisdom: A culturally inclusive developmental perspective. *International Journal of Behavioral Development*, *26*(3), 269–277.

Von Stumm, S., & Ackerman, P. L. (2013). Investment and intellect: A review and meta-analysis. *Psychological Bulletin*, *139*(4), 841–869.

Webster, J. D. (2003). An exploratory analysis of a self-assessed wisdom scale. *Journal of Adult Development*, *10*, 13–22.

Webster, J. D. (2007). Measuring the character strength of wisdom. *International Journal of Aging & Human Development*, *65*, 163–183.

Webster, J. D. (2019). Self-report wisdom measures: Strengths, limitations, and future directions. In R. Sternberg & J. Glück (Eds.), *The Cambridge handbook of wisdom* (pp. 297–320). Cambridge: Cambridge University Press.

Weststrate, N. M., & Glück, J. (2017a). Hard-earned wisdom: Exploratory processing of difficult life experience is positively associated with wisdom. *Developmental Psychology*, 53, 800–814.

Weststrate, N. M., & Glück, J. (2017b). Wiser but not sadder, blissful but not ignorant: Exploring the co-development of wisdom and well-being over time. In M. D. Robinson & M. Eid (Eds.), *The happy mind: Cognitive contributions to well-being* (pp. 459–480). New York: Springer.

Yang, S.-Y. (2001). Conceptions of wisdom among Taiwanese Chinese. *Journal of Cross-Cultural Psychology*, *32*, 662–680.

Zachry, C. E., Phan, L. V., & Blackie, L. E. R., & Jayawickreme, E. (2018). Situation-based contingencies underlying wisdom-content manifestations: Examining intellectual humility in daily life. *The Journals of Gerontology, Series B: Psychological Sciences*, *73*(8), 1404–1415.

48 Intelligence and Expertise

Phillip L. Ackerman

Defining Terms

One traditional approach to starting a discussion of the relations between two constructs is to attempt to define one's terms. Various methods are often used for providing such a foundation for discussion but the two most common, and central to the current purposes, are the "lexical" and "stipulative" forms of definition (see Robinson, 1950). Lexical definitions are those that are essentially "dictionary" definitions. They are historically documented and based on current and prior usage. The truth-value of a lexical definition is one that can be determined in a straightforward fashion, merely by reference to original source material. Stipulative definitions are those that are proposed by the individual who chooses to use a word to mean a particular concept. As such, there is no way to determine the truth-value of a stipulative definition. The value of the stipulative definition is instead determined by other indicators, such as its consistency in a wider network of other constructs. Why provide a short discourse on definition here? The answer lies in the need to relate two different concepts that rely on different kinds of definitions. For expertise, we can rely on a lexical definition but, for intelligence, it is largely impossible to provide a coherent discussion without a stipulative definition.

Expertise

The lexical definition of expertise is both straightforward and useful for the current discussion. "Expertise" refers to having the skill of an expert. An expert, according to the *Oxford English Dictionary* (OUP, 1971), is someone who is experienced and who has been "trained by experience or practice, skilled" (p. 930). The term "expert" has been used since Chaucer's time and current usage is generally consistent with usage over the past 600 years. The foundation for expertise, then, is the notion that one has a skill or skills and that they are obtained through practice or other experiences. The one addition that should be provided here is that, in modern usage, expertise need not be limited to skills that involve a significant physical component (such as playing the violin or performing heart bypass surgery) but they may also involve "knowledge" in a more general sense. In psychology and education, three forms of knowledge have been articulated. One kind of knowledge is called "procedural knowledge" or "knowing how" (Ryle, 1949/2000). Skills that involve physical components generally fall into this category of knowledge. Such skills range in complexity from carpentry and bricklaying to neurosurgery and world-class musical performances. The second kind of knowledge is called "declarative knowledge" or "knowing that" (Ryle, 1949/2000). Declarative knowledge is essentially factual knowledge, whether it is the knowledge of a lawyer, novelist, physicist, psychologist, or a member of many other "knowledge-worker" professions. A third kind of knowledge has been called "tacit knowledge" (Polanyi, 1966/1983; Wagner & Sternberg, 1985) or "knowing with" (Broudy, 1977). This kind of knowledge is less well understood than the other two forms of knowledge. This kind of knowledge is called "tacit" because it is not usually spontaneously articulated nor is it often easily accessible to verbal reports. It is thought to develop through one's educational and cultural experiences but it is something that is not directly trained or practiced. Nonetheless, such knowledge is especially important when individuals are faced with problem-solving that is outside of their normal areas of declarative or procedural expertise.

From a practical perspective, declarative knowledge can be categorized into a variety of different topic domains and procedural knowledge can be categorized by particular skills. Tacit knowledge, as conceptualized by Polanyi and Broudy, cannot be easily categorized and thus is quite difficult to study. These categories of knowledge will be discussed in greater detail with respect to fluid and crystallized intelligence.

Intelligence

Lexical definitions of intelligence are especially problematic because there have been literally hundreds of different definitions offered for the concept over the past several hundred years. Psychologists have several times attempted to come to a consensus over how to define intelligence (e.g., *Journal of Educational Psychology*, 1921; Sternberg & Detterman, 1986), without much success. One can find a wide variation in how intelligence is defined by the different chapter authors in this book.

In order to have a coherent discussion of intelligence and expertise here, I will propose a stipulative definition – one that allows for consideration of how aspects of intelligence relate to different kinds of expertise. The definition is based on theories initially articulated by Hebb (1942) and by Cattell and Horn (Cattell, 1943, 1957, 1971; Horn, 1968, 1989; Horn & Cattell, 1966). Although their theories are more nuanced than is represented here, the fundamental property of the theories is that there are two central components of intelligence – one that is associated with "process" and the other associated with "knowledge." The component of intelligence that is associated with "process" is typically called "General Fluid Intelligence" (G*f*) and the other component is associated with "knowledge" and is typically called "General Crystallized Intelligence" (G*c*) (see Cattell, 1943). G*f* refers to abstract reasoning, short-term memory, and working memory. G*f* is most often involved in the solution of novel problems or keeping track of decontextualized information in one's head for brief periods of time (e.g., letters, numbers, or random words). G*f* also plays an important role in learning, especially for young children. Cattell proposed that G*f* is essentially innate, that is, one is born with a certain level of G*f* that determines one's success in later learning and intellectual development – a conceptualization that was consistent with Spearman's notion of general intelligence or *g* (Spearman, 1904). In contrast to G*f*, G*c* is developed through education and experience. It represents the knowledge that people acquire throughout the life span. Language, such as vocabulary and reading comprehension, reasoning, and problem-solving in context-dependent domains (math, science, arts and humanities, law, business, etc.) all make up an individual's G*c*. In practice, however, assessments of individual differences in G*c* focus on broad knowledge but almost always only at a surface level rather than a deep level. For adults, this brings us to a distinction between what Cattell (1957) referred to as "historical" G*c* and "current" G*c*.

Historical and Current Gc

Because G*c* represents the entire repertoire of knowledge and skills that an individual has, it does not directly translate to "expert" levels of performance in any single domain. Cattell suggested that, as individuals reach adolescence and adulthood, G*c* becomes more diverse and differentiated, especially as young adults acquire direct experience in occupational and avocational domains. The problem of assessment is that, to measure an adult's G*c*, one must develop tests of every possible domain of knowledge, both declarative and procedural. Without such a wide array of tests, for example, master carpenters are given no credit for their knowledge/skills at carpentry, dentists are given no credit for their skills at filling cavities, a psychologist is given no credit for knowing the current and historical theories of the field, and so on. The alternative to this impossible task of developing hundreds of tests for expert knowledge, according to Cattell, was to assess only what the individual had learned prior to receiving specialized training or practice, namely one's "historical" G*c*.

Historical Gc

The assessment of historical G*c* is in essence how G*c* is usually assessed for adolescents and adults. The quantitative sections of the SAT examination, for example, contain only algebra and geometry problems, even though students take the examination in their junior and senior years of high school, when some of the students have proceeded to trigonometry and a smaller set of students has moved even further to calculus courses. Four years later, when college/university students want to apply for graduate study, they often take the Graduate Record Examination (GRE). Although some students have majored in mathematics or related fields, the math section of the GRE General test is made up of algebra and geometry problems – even though it may have been six or more years since the student had completed a course in these areas. Such assessments are one example of testing for historical G*c* rather than current G*c*.

Current G*c*

For adults, assessments of current G*c* are frequently narrow occupational and professional tests designed to measure a particular domain of expert knowledge and skills. That is, they do not attempt to determine the individual's entire repertoire of knowledge; rather, they attempt to determine whether the individual has acquired an acceptable level of expertise to be licensed to practice in a particular profession. These assessments can be a grueling ordeal as they often require extensive education, experience, and months of study and preparation. In addition, the tests themselves can last for several days. For example, among people seeking to pursue a law career, admission to the Bar (the professional certification process for lawyers in the United States) requires an examination that typically involves two to three consecutive days of testing, with six hours a day or more of tests, depending on the state. Similarly, the US medical licensing examination (Step 3, from the Federation of State Medical Boards) requires two eight-hour days of testing. Sonographers seeking board certification for Registered Diagnostic Medical Sonographer examination must complete multiple general and specific tests (ARDMS, 2019), and each year nearly 175,000 people worldwide take one of the three, eight-hour tests conducted over a single ten-hour period to achieve the status of Chartered Financial Analyst (CFA Institute, 2008; *The Economist*, 2008). Individuals who seek state licensure to practice psychology typically take a four-hour, fifteen-minute examination (ASPPB, 2008), which is often supplemented by a state-specific examination taken on the same day. These tests are aimed at measuring individual differences in expertise but they are also measuring one aspect of current G*c*.

Summary

To this point, I have defined expertise as knowledge and skills that have been acquired through experience/practice. In addition, I have proposed a stipulative definition of intelligence as composed of two broad components: G*f* is associated with abstract reasoning and short-term memory and G*c* is associated with knowledge and skills. Within G*c*, I have distinguished between historical G*c* (knowledge/skills common to a culture) and current G*c* (both common knowledge and specialized knowledge/skills). Expertise is most highly identified with current G*c* in adults. However, I have not addressed the relationship between the components of intelligence and the acquisition of expertise – that is, answering the question of what are the roles of G*f* and G*c* in determining who develops expertise and whether an individual's level of G*f* and G*c* relates to the domain of expertise that is developed. The next section focuses on how expertise is developed and the role that intelligence plays in the development of expertise. First, however, a review of the difficulties in researching individual differences in expertise is provided.

Methods for the Study of Individual Differences in Expertise

The study of individual differences in the acquisition of expertise is fraught with difficulties. First, most scholars of expertise agree that it takes several years of study or practice (e.g., see Simonton, 1988) to develop high levels of expertise within a single domain. Although one can study early acquisition curves for knowledge and skill development in the laboratory, because of the substantial time and effort investment typically required to develop expertise, it is generally not feasible to randomly assign individuals of a wide range of intelligence levels to the task of acquiring expertise in nearly any domain in any society that allows for freedom of choice in educational or occupational development.

Most researchers rely on one of two different methods for studying individual differences in expertise. The first method employs intact groups of individuals who have already acquired a high level of expertise in a field. These individuals are compared to one another and sometimes to a group of individuals who are not expert in the field. Both kinds of comparisons have limited utility. Looking for individual differences in intelligence among a group of PhD-level physicists, for example, who have already been the subject of repeated selections (at college entry, at graduate school entry, and through exams in graduate school) is likely to reveal very little useful information, because correlations are severely attenuated (i.e., close to zero) when the range of talent is very small. By way of analogy, consider that even though one could reasonably assert that height is a critical requirement for expert performance in basketball, the correlation between player height and performance in the National Basketball Association (NBA) is attenuated because the *average* height of NBA players in the 2007–2008 season is 6 feet 7 inches and the shortest NBA player was 5 feet 9 inches (NBA, 2009).

The second method, that is, comparing a group of experts with a group of nonexperts (e.g., master bridge players vs. nonmaster bridge players who have been playing for a similar amount of time), may be informative but such a method suffers from the classic problem of unknown third variables that may also contribute to the differences between those individuals who acquire high levels of expertise and those who do not. Other variables may also differentiate the experts from the population at large, including individual differences in intelligence, but, without random assignment to practice/training, one cannot know the amount of influence these other variables might have on the development of expertise. Finding an appropriate group of nonexpert individuals for comparison purposes is a nearly impossible task. One may well ask, for example, if board-certified neurosurgeons are, as a group, more intelligent than nonboard-certified neurosurgeons, or more intelligent than doctors without surgical specialties, or college graduates, or the at-large population, and so on. Such comparisons *suggest* that there are many domains of expertise that are associated with higher levels of intelligence but they do not definitively indicate whether high levels of intelligence are necessary for the development of expertise, partly because people who are lower in intellectual abilities are less likely to be encouraged to, or allowed to, pursue these professions. For additional details, see Ackerman and Beier (2006).

Professions or hobbies that allow for the development of expertise but do not have strict educational gatekeepers, such as betting at the racetrack (Ceci & Liker, 1986a) or playing Scrabble (Halpern & Wai, 2007), bridge, and chess (Gobet & Charness, 2006), or having other skills (for a review, see Ericsson & Charness, 1994), are more amenable to expert/novice comparisons with respect to differences in intellectual abilities. However, individuals who acquire expertise in these domains are likely to have done so with vastly different experiences from those of professionals in medicine, who have gone through very structured educational/training programs (e.g., see Norman et al., 2006).

Acquiring Expertise

Closed Skills

For some kinds of expertise, the domain of knowledge or procedural skill to be acquired is relatively fixed and finite. In a narrow sense, becoming an expert typist represents a "closed" skill, as the number of keys to be used on the computer keyboard is fixed and no changes are made to their arrangement. Increasing levels of practice leads to increasing performance, though after the initial phase of practice, performance improvements show diminishing returns with additional practice. Newell and Rosenbloom (1981) called this the "power law of practice." In essence, an equal amount of improvement in speed of performance is found for the first 10 trials, the next 100 trials, the next 1,000 trials, and so on. Performance keeps improving with practice but the increments in improvement become smaller and smaller over time. The literature on closed procedural skills suggests that intellectual abilities may be influential in the first phase of skill acquisition, when learners are still figuring out strategies for completing the task. With high levels of practice, there is a reduction in both the range of differences between individuals in performance and a concomitant reduction in the influence of intellectual abilities on individual differences in performance (e.g., see Ackerman, 1987, 1988). Thus, acquiring expertise on relatively straightforward closed skills is within the capabilities of much of the population. Once learned, these tasks are often "automatic," in that it requires little or no effort on the part of the individual to perform them at a high level of expertise. This is not to say that such skills are at a world-class level. To reach that level, more extensive practice is necessary, even for text messaging, driving a car, or mental multiplication. In addition, to achieve truly exceptional performance in such domains, individuals have to focus their attention on the task while it is being performed (e.g., think of the difference in driving performance when one eliminates all distractions in comparison to when one drives and talks on their cell phone at the same time).

Open Skills

Most domains of expertise that depend on declarative knowledge rather than procedural knowledge are open, in the sense that more knowledge brings about improved performance and in the sense that, once one component of the skill is acquired, another, more complex component of the skill is yet to be learned. Becoming an expert at mathematics has this characteristic: Once learners acquire arithmetic skills at addition and subtraction, they are presented multiplication/division, then algebra, geometry, trigonometry, derivative and integral calculus, and so on. Although each separate component of the skill may be "closed" – with a fixed set of rules, facts, and procedures to be learned – to become an expert requires that one acquire knowledge and skill at each of the more complex components of the task. Acquiring expertise in such domains is a lifelong task and one that depends on intellectual abilities because these abilities are integral in acquiring expertise when faced with each increasingly complex component of the skill to be learned. At some level, the individual may choose to "specialize," in which case the challenges to acquire more complex task components *might* be diminished, depending on how rapidly the domain changes. Any time there is a change in the field of expertise, such as the introduction of new technology (whether it be, for example, new equipment for surgical procedures or diagnostic tools in medicine, or new computer systems for the solution of technical problems or design), the challenge to stay up-to-date is one that will make demands on the individual's intellectual abilities. (The decline in G*f* associated with increasing age in middle-aged and older adults tends to make such new learning more difficult than it is for younger adults; see Kubeck et al., 1996.)

Expert Short-Term/Working Memory

There have been a few notable studies that have attempted to develop expertise in short-term and working memory capabilities. The general framework proposed by Miller (1956) is that humans have a capacity of keeping about 7 +/–2 items active in short-term memory at any one time. Individuals differ in their short-term memory capacity and such differences are considered to be an integral part of fluid intellectual abilities. Strategies for memorizing new information in a more efficient and effective manner have been common since the time of the Greeks (e.g., the Method of Loci; see Yates, 1966). These strategies, along with "chunking" – that is, combining new information into larger units – are not aimed at creating expertise in memory per se but rather they are aimed at more effective use of one's limited attentional resources. In the normal day-to-day environment, having expert memory appears to be a matter of G*f* abilities, the allocation of effort to memorize information, and the use of effective strategies. Remembering phone numbers or names of people at a party, for example, is dependent mainly on those three factors.

However, in one research program (Chase & Ericsson, 1981), the authors were able to train *some* individuals to keep track of much more information. In one noteworthy case, with 250 hours of practice over the course of two years, one learner was able to develop the skill of keeping track of more than eighty random digits read aloud at a rate of one per second (where the typical individual can keep track of no more than seven digits). This individual was able to use his extensive long-term memory of running speeds (e.g., world records for various distance races) as reference tags for chunks of numbers and then to retrieve the numbers on a recall test. Attempts to train other individuals without such deep knowledge of numerically based information, or to train the same individual to recall long sequences of letters instead of numbers, were largely unsuccessful. Being able to recall more than eighty random digits when presented only once at a rapid rate is clearly an example of expertise but whether it represents skilled short-term or working memory (which would make it a G*f* ability) or a unique use of a highly organized long-term memory (which would make it more of a G*c* ability) is debatable.

Is G*f* a Limiting Factor?

One of the most contentious issues in the study of individual differences in expertise is the question of whether G*f* is a limiting factor. There are, in essence, two related questions. The first question is whether there is a *threshold* of G*f* needed for developing expertise. The second question is whether higher levels of G*f* lead to higher levels of expertise or a faster development of expertise, ceteris paribus (that is, all other things being equal). Each of these issues is treated in turn.

G*f* as Threshold

One conceptualization of the acquisition of expertise is that there is a threshold level of G*f* or general intelligence, below which an individual is unlikely to develop expertise in a particular domain (e.g., Gibson & Light, 1992). In the limit, this is surely true. For example, moderately or profoundly retarded individuals are highly unlikely to develop expertise in nuclear physics, compared to individuals who have high levels of Gf. However, there is no "fixed" threshold for the development of expertise in most areas. Indeed, early studies of the relationship between intelligence and occupational status (e.g., Stewart, 1947) showed that there is a wide range of intelligence levels for nearly all occupations, even though mean levels of intelligence for the occupations of doctor, lawyer, and scientist are well above average. There are two likely explanations for these findings. First, standard group measures of intellectual abilities (both G*f* and G*c*) – the kind most frequently administered to large groups of job and school applicants – are not comprehensive, in that they may miss some important components of intelligence that are relevant to educational and occupational success (such as spatial abilities; see, e.g., Webb, Lubinski, & Benbow, 2007). Second, because the acquisition of expertise depends on the investment of practice and study over an extended period of time, individuals with relatively lower levels of intellectual abilities may sometimes compensate for their abilities by working harder and longer to acquire the knowledge/skills necessary to develop expertise. In practice, however, the overwhelming majority of regressions between ability and job performance is found to be linear (Coward & Sackett, 1990), suggesting that the intelligence threshold conceptualization is not particularly viable and that higher levels of intelligence lead generally to higher levels of occupational performance.

Impact of Higher G*f* – Declarative Knowledge

Even though there may not be a fixed threshold for G*f* in determining the acquisition of expertise, extant data suggest that, ceteris paribus, higher levels of G*f* will result in a higher likelihood of developing expertise in a variety of academic and other declarative knowledge-dependent domains. Studies of individuals who have extremely high levels of intellectual abilities indicate a much higher representation of experts in such fields (e.g., see Lubinski & Benbow, 2006). At some point in the acquisition of expertise, however, the role of G*f* appears to be diminished in favor of an increasing influence from G*c*, in the form of transfer.

Impact of Higher G*f* – Procedural Knowledge

G*f* is, however, not as important in the development of many procedural skills. For expertise that depends on procedural skills, especially when initial performance on such tasks is within the capabilities of most individuals (even if slow and error prone), G*f* has a much-diminished association with acquisition of expertise. For this to happen, though, the procedural skill to be learned needs to be "closed" rather than "open" (for discussions, see Ceci & Liker, 1986b; Ericsson & Lehmann, 1996). If the skill has increasingly complex procedures that must be learned, then it can be expected that intellectual abilities will have an increased effect on individual differences in performance at each higher level of complexity required by the skill to be acquired.

G*c* and Transfer

In the section on "Intelligence," when a stipulative definition of intelligence was provided, the two main components of intelligence were denoted G*f* and G*c*. If current G*c* represents acquired knowledge and skills, then domain-specific expertise represents a subset of an individual's intellectual repertoire. By definition, then, expertise is closely related to intelligence. But this assertion does not address the role of G*c* in the acquisition of expertise. Gf has been shown to be instrumental in reasoning and problem-solving in the absence of prior context, a critical component when one attempts to acquire knowledge and skill in a novel domain. But, as people begin to learn about a particular domain, new knowledge and skill are developed partly on the foundation of earlier learning and skills. Ferguson (1956) offered a strong thesis along these lines. He suggested that learning of only a newborn child occurs in the absence of transfer – that is, building new knowledge on existing knowledge. In that sense, individual differences in existing knowledge are the most important determinant of acquiring new knowledge in the same domain. As learners attempt to acquire expertise, what they already "know" is the main limiting factor for new learning.

If Ferguson's assertion is true, then current G*c*, in that it represents the individual's repertoire of knowledge and skill, should be more highly related to an individual's current level of expertise than is G*f*, and G*c* should be more highly related to the acquisition of new knowledge in the same general domain. Scientifically evaluating this assertion is difficult, for some of the same reasons that comparisons of individual differences among experts or contrasts between experts and novices is problematic. One can evaluate the individual's current domain-specific knowledge with tests that allow for assessment of deep domain knowledge but people cannot be randomly assigned to control and experimental groups for domain-learning situations that require years of experience to develop high levels of expertise. Although there have not been extensive studies that have related historical G*c* to domain-specific expertise, assessments of adult knowledge in the physical sciences, technology, social sciences, humanities, business/law, health and nutrition, and current events illustrate a consistent pattern of correlations (e.g., Ackerman, 2000; Ackerman & Rolfhus, 1999; Beier & Ackerman, 2001, 2003). For all of these knowledge domains, measures of G*c* show substantial correlations with individual differences in the depth of knowledge (correlations in the range of $r = 0.48$ to 0.80). Correlations between Gf and domain knowledge are usually much smaller (in the range of $r = 0.33$ to 0.49) for most domains, with the exception of physical sciences and technology, where both G*c* and G*f* abilities are highly correlated with domain knowledge. These studies do not necessarily point to direct transfer of knowledge from historical G*c* to domain-specific expertise, especially because a third variable could account for both high G*c* and high levels of domain-specific expertise. However, they are consistent with Ferguson's conceptualization that transfer is a key ingredient to intellectual development and to the development of expertise.

Expertise Transfer and Intelligence

Just as individual differences in intellectual abilities and skills can be expected to transfer to the development of expertise, the development of expertise can be expected to transfer to intellectual abilities. The problem in assessing the degree of transfer from domain knowledge and skills, or even memory skills to intellectual abilities, lies in determining how best to assess the transfer. On the one hand, because standardized intelligence tests, as discussed in the section on "Historical G*c*," tend to sample broadly but at a surface level, developing expertise in, say, medicine, might have a small beneficial effect on a vocabulary subscale but little effect on digit-span or reading comprehension. Developing expert memory skills, on the other hand, might have much larger effects on standard intelligence measures, especially those that depend on short-term and working memory. Other researchers have suggested that the challenges of complex jobs through adulthood lead to better maintenance of intellectual abilities (Kohn & Schooler, 1978; Schooler, 2001; Willis & Tosti-Vasey, 1990; for a review, see Krampe & Charness, 2018). Another issue to be considered is whether intelligence, per se, represents "developing expertise" (e.g., see Sternberg, 1999) or is a form of expertise. Certainly one general aim of education is the development of knowledge and skills that make up a significant portion of what is considered to be intellectual, especially in the basic skills in literacy and foundations of science, math, and other areas (e.g., see Alexander & Murphy, 1999; Snow, 1996; Stanovich & West, 1989). These are important aspects of the development and expression of intelligence but they relate more to a view of "expertise" that is much more general than I have discussed to this point and they probably fall into the tacit/knowing with kinds of knowledge proposed by Polanyi (1966/1983) and discussed by Broudy (1977).

The "Deliberate Practice" View

A view proposed by Ericsson and his colleagues (e.g., Ericsson, 2006; Ericsson, Krampe, & Tesch-Römer, 1993) and popularized by lay observers, such as Gladwell (2008), is that intellectual abilities are largely irrelevant to the development of expertise and that individual differences in the depth and extent of "deliberate practice" are the main determinants of expertise. The sources of expertise studied by these researchers include tasks such as playing chess, typing, performing music, and playing sports. Comparison groups are typically those individuals who have practiced a task for similar amounts of time as the expert group but have not achieved high levels of expert performance. The lack of a substantial difference between these groups on standard ability tests is taken as evidence that intellectual abilities are not relevant for distinguishing between experts and nonexperts. The professional basketball player height analogy mentioned in the section on "Methods for the Study of Individual Differences in Expertise" applies to these comparisons. That is, when one is dealing with a group of individuals who are severely limited in range-of-talent (because even those individuals deemed "nonexperts" perform many levels higher than the at-large population), one expects that even if an individual-differences variable is related to success in a random sample of people, it will not be revealed in a group that has a severe restriction in range-of-talent. Such studies do not inform one about the role that intellectual abilities play at the various stages of entry to the domain, the speed with which one develops expertise, or the level of expertise ultimately attained.

A definitive study of expertise to evaluate the deliberate practice has yet to be conducted. Such a study would involve taking a large number of otherwise unselected individuals, testing them on intellectual abilities at the outset of the study, and subjecting them to extensive (e.g., 10,000 hours) of deliberate practice. After the several years of practice, correlations between the intellectual abilities and acquired levels of expert performance would need to be computed. Ericsson's notions would be confirmed if, and only if, the correlations were essentially zero. In the absence of such a study, numerous sources of indirect evidence have been presented to show the likely folly of the deliberate practice viewpoint (e.g., see the review by Hambrick, Macnamara, et al., 2016; Hambrick, Oswald, et al., 2014; see also Ackerman, 2014a, 2014b; and Hambrick, Campitelli, & Macnamara, 2018). The lack of low-IQ individuals who are chess grandmasters, successful hedge-fund traders, neurosurgeons, PhD-level chemists, physicists, or even psychologists suggests that there is indeed an important role of intellectual abilities for the acquisition of expertise in cognitively demanding domains; and that the relationship between intellectual abilities and occupational performance is largely linear.

Maintenance of Expertise

One of the interesting aspects of expertise that provides an additional basis for aligning it closer to G*c* than to G*f* is the pattern of growth and decline of expertise that occurs during middle age and older adulthood. Both theory and extant data indicate that G*f* reaches a peak for most people in early adulthood, generally between the age of eighteen and the mid-twenties (Horn, 1989; Salthouse, 1996). In contrast, both historical and current G*c* is maintained well into middle age and some studies have suggested that current G*c* also shows growth into middle age (e.g., see Ackerman, 2000; Horn, 1989; Horn & Masunaga, 2006). Domain-specific expertise in many areas is also well preserved into middle age and beyond, in reviews that have been conducted on this issue (e.g., Simonton, 1988). In one of the first longitudinal studies of intelligence of adults, Owens (1953) found that, on an information test first administered thirty-one years earlier when participants were nineteen, a group of adults performed much better on a test of general information. The average score was nearly one standard deviation higher than at initial testing. In a more extensive longitudinal study, Schaie (1996, 2005) found that general verbal knowledge grows and is maintained up to about age sixty and then shows declines as people reach their seventies and eighties.

For narrower areas of domain knowledge, other studies have indicated that knowledge and skills are well preserved, *if it has been well learned to begin with*, even if the individuals do not actively use the knowledge in the intervening years. Studies of Spanish-language knowledge by Bahrick (1984), and algebra and geometry knowledge (Bahrick & Hall, 1991) acquired first in high school and college, found high levels of recall over periods of up to fifty years, though A-grade students performed much better at recall than did C-grade students. Procedural knowledge and skills, once acquired, have also been shown to be well preserved over long periods of time. The old adage about retaining skill in riding a bicycle, even after many years of nonuse, is consistent with the extant data. For juggling, when an individual was trained for forty-two daily sessions at initial acquisition, performance assessed six years later was nearly as good as the last performance during initial acquisition (Swift, 1910). In a remarkable study of typewriting skill retention, Hill, a novice typist, acquired expertise at typing over five months of daily practice (Hill, Rejall, & Thorndike, 1913). In two follow-up assessments, he assessed his retained typing skill, first, after a twenty-five-year period, during which he did not use the typewriter (Hill, 1934), and then, after a total of fifty years after the initial training, again without using the typewriter (Hill, 1957), when he was eighty years old. After twenty-five years of nonuse, he performed at a level that he had only achieved after twenty-seven days of initial practice. After fifty years, even though his perceptual/motor abilities had surely declined with age, he was able to achieve the same level of performance after only eight days of retraining.

It should be emphasized that the important finding is that, when the procedural skills are well developed to begin with, the retention period can be very long indeed, even when the skill is not regularly exercised. Of course, continued use of the skill can be expected to lead to even better maintenance or improvement, up to the limits of a person's perceptual and motor abilities, as was exemplified in the skills of Michael DeBakey, the pioneering heart surgeon. By the time he finally retired from practice at age ninety, he had performed more than 60,000 cardiovascular procedures and was still considered one of the best surgeons in the field (see Nuland, 2007). Similarly, several world-class classical musicians have performed well into their seventies and eighties (e.g., Isaac Stern, Arthur Rubinstein, Vladimir Horowitz). At advanced ages, these musicians were more likely to perform the standard repertoire pieces, yet their skills were exceptionally well maintained.

Tacit Knowledge Expertise

Determining the relationship between intelligence and tacit knowledge expertise is even more of a challenge than it is for declarative and procedural knowledge. Where declarative knowledge can be reasonably well measured, with tests designed to assess knowledge that can be verbally reported, and procedural knowledge can be measured by asking the individual to perform the skill in question, tacit knowledge is, by definition, not spontaneously articulated nor is it often easily accessible to verbal reports. A few studies have been conducted to assess tacit knowledge by providing scenarios in the domain to examinees (e.g., in-basket management problems; see Wagner, 2000) and then evaluating the quality of the responses. Under these circumstances, good or excellent performance is determined not by evaluating the difference between optimal strategies and the individual's response but rather by determining the similarity of the individual's response and a consensus response by experts (e.g., see Wagner & Sternberg, 1987). To date, studies in this area have suggested relatively low correlations between tacit knowledge and standard tests of intellectual abilities, although the comparisons between experts and novices made in these studies are subject to the same limitations noted in the discussion of "deliberate practice" about evaluating individual differences in samples where there is a restriction in range-of-talent (e.g., see Cianciolo et al., 2006).

Summary and Conclusions

The study of intelligence and expertise is a much more recent focus for researchers than is the study of, say, intelligence and academic performance. Nonetheless, based on research from experimental psychology that has focused on understanding the development and expression of expertise, and a small number of studies that have examined individual differences in expertise, a relatively consistent pattern of results has been found. Individual differences in expertise are not directly measured by historical G*c* assessments. Most current G*c* measures do not involve the kind of depth in assessment necessary to probe an individual's expertise, in contrast to measures of professional competency or professional certification tests. Experts in domains that are highly dependent on declarative knowledge, most often acquired through extensive education and experience (e.g., in law, medicine, science), will have higher levels of intellectual abilities (both G*f* and G*c*) than the lay public. Whether higher intellectual abilities are *necessary* for acquisition of such levels of expertise is not directly known, because gatekeepers to entry for these occupations depend on intellectual ability tests for selection into the educational or occupational programs. However, intellectual abilities are not *sufficient* for the development of expertise; other factors, such as motivation and effort for learning and task practice over long periods of time, play an important role in determining who becomes an expert.

Studies of experts in domains that are more highly dependent on procedural knowledge show mixed results in the correlations with intellectual abilities. In several studies, researchers have claimed that there is essentially a zero correlation between expertise in these domains and intellectual abilities, though such inferences are dependent on the interpretation of data from individuals who are already restricted in range-of-talent or nonexpert comparison groups that may or not be equivalent to the expert groups. For experts in the domain of tacit knowledge, it is, as yet, difficult to draw conclusions regarding the role of intellectual abilities. Improved measurement techniques for assessing tacit knowledge may ultimately help address these issues. In addition, a better understanding of how tacit knowledge is acquired can be expected to provide additional insights into the relationships between G*f*, G*c*, and tacit knowledge.

In the final analysis, higher levels of intellectual abilities appear to give the learner a head start or an overall advantage in the acquisition of expertise over learners with lower levels of intelligence. For closed tasks, especially those that are mostly dependent on procedural skills, the influence of intellectual abilities diminishes with increasing practice, as motivation, effort, and persistence increase in influence. For open tasks, especially those that are mostly dependent on declarative knowledge, intellectual abilities, and especially G*c*, appear to be important determinants of higher levels of expertise.

References

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin*, *102*, 3–27.

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, *117*, 288–318.

Ackerman, P. L. (2000). Domain-specific knowledge as the "dark matter" of adult intelligence: gf/gc, personality and interest correlates. *Journal of Gerontology: Psychological Sciences*, *55B*(2), P69–P84.

Ackerman, P. L. (2014a). Nonsense, common sense, and science of expert performance: Talent and individual differences. *Intelligence*, *45*, 6–17.

Ackerman, P. L. (2014b). Facts are stubborn things. *Intelligence*, *45*, 104–106.

Ackerman, P. L., & Beier, M. E. (2006). Methods for studying the structure of expertise: Psychometric approaches. In A. Ericsson, P. Feltovich, N. Charness, &

R. R. Hoffman (Eds.), *Cambridge handbook on expertise and expert performance* (pp. 147–166). New York: Cambridge University Press.

Ackerman, P. L., & Rolfhus, E. L. (1999). The locus of adult intelligence: Knowledge, abilities, and non-ability traits. *Psychology and Aging*, *14*, 314–330.

Alexander, P. A., & Murphy, P. K. (1999). Learner profiles: Valuing individual differences within classroom communities. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 413–436). Washington, DC: American Psychological Association.

ARDMS (American Registry for Diagnostic Medical Sonography). (2019). Get certified. Registered Diagnostic Medical Sonograhper® (RDMS®). www.ardms.org/get-certified/rdms/

ASPPB (Association of State and Provincial Psychology Boards). (2019). *EPPP candidate handbook*. https://cdn.ymaws.com/www.asppb.net/resource/resmgr/eppp_/eppp_cand-handbook-1_16_2019.pdf

Bahrick, H. P. (1984). Fifty years of second language attrition: Implications for programmatic research. *Modern Language Journal*, *68*(2), 105–118.

Bahrick, H. P., & Hall, L. K. (1991). Lifetime maintenance of high school mathematics content. *Journal of Experimental Psychology: General*, *120*(1), 20–33.

Beier, M. E., & Ackerman, P. L. (2001). Current events knowledge in adults: An investigation of age, intelligence and non-ability determinants. *Psychology and Aging*, *16*, 615–628.

Beier, M. E., & Ackerman, P. L. (2003). Determinants of health knowledge: An investigation of age, gender, abilities, personality, and interests. *Journal of Personality and Social Psychology*, *84*(2), 439–448.

Broudy, H. S. (1977). Types of knowledge and purposes of education. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 1–17). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*, 153–193.

Cattell, R. B. (1957). *Personality and motivation structure and measurement.* Yonkers-on-Hudson NY: World Book.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. New York: Houghton Mifflin.

Ceci, S. J., & Liker, J. K. (1986a). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General*, *115*, 255–266.

Ceci, S. J., & Liker, J. (1986b). Academic and nonacademic intelligence: An experimental separation. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 119–142). New York: Cambridge University Press.

CFA (Chartered Financial Analyst) Institute. (2008). CFA® exam information. www.cfainstitute.org/en/programs/cfa/exam

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 141–189). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cianciolo, A. T., Matthew, C., Sternberg, R. J., & Wagner, R. K. (2006). Tacit knowledge, practical intelligence, and expertise. In A. Ericsson, P. Feltovich, N. Charness, & R. R. Hoffman (Eds.), *Cambridge handbook on expertise and expert performance* (pp. 613–632). New York: Cambridge University Press.

Coward, W. M., & Sackett, P. R. (1990). Linearity of ability-performance relationships: A reconfirmation. *Journal of Applied Psychology*, *75*, 297–300.

The Economist. (2008). Charter School. *The Economist*, June 5. www.economist.com/finance-and-economics/2008/06/05/charter-school
Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In A. Ericsson, P. Feltovich, N. Charness, & R. R. Hoffman (Eds.), *Cambridge handbook on expertise and expert performance* (pp. 683–703). New York: Cambridge University Press.
Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725–747.
Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.
Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305.
Ferguson, G. A. (1956). On transfer and the abilities of man. *Canadian Journal of Psychology*, *10*, 121–131.
Gibson, J., & Light, P. (1992). Intelligence among university scientists. In R. S. Albert (Ed.), *Genius and eminence* (2nd ed., pp. 109–111). International series in experimental social psychology, 22. Elmsford, NY: Pergamon Press.
Gladwell, M. (2008). *Outliers: The story of success*. New York: Little, Brown and Company.
Gobet, F., & Charness, N. (2006). Expertise in chess. In A. Ericsson, P. Feltovich, N. Charness, & R. R. Hoffman (Eds.), *Cambridge handbook on expertise and expert performance* (pp. 523–538). New York: Cambridge University Press.
Halpern, D. F., & Wai, J. (2007). The world of competitive Scrabble: Novice and expert differences in visuospatial and verbal abilities. *Journal of Experimental Psychology: Applied*, *13*, 79–94.
Hambrick, D. Z., Campitelli, G., & Macnamara, B. N. (Eds.) (2018). *The science of expertise: Behavioral, neural, and genetic approaches to complex skill*. New York: Routledge.
Hambrick, D. Z., Macnamara, B. N., Campitelli, G., Ullén, F., & Mosing, M. A. (2016). Beyond born versus made: A new look at expertise. *Psychology of Learning and Motivation*, *64*, 1–55.
Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34–45.
Hebb, D. O. (1942). The effect of early and late brain injury upon test scores, and the nature of normal adult intelligence. *Proceedings of the American Philosophical Society*, *85*(3), 275–292.
Hill, L. B. (1934). A quarter century of delayed recall. *Journal of Genetic Psychology*, *44*, 231–238.
Hill, L. B. (1957). A second quarter century of delayed recall, or relearning at eighty. *Journal of Educational Psychology*, *48*, 65–69.
Hill, L. B., Rejall, A. E., & Thorndike, E. L. (1913). Practice in the case of typewriting. *Pedagogical Seminary*, *20*, 516–529.
Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, *75*, 242–259.
Horn, J. L. (1989). Cognitive diversity: A framework of learning. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 61–116). New York: W. H. Freeman.
Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*, 253–270.

Horn, J., & Masunaga, H. (2006). A merging theory of expertise and intelligence. In A. Ericsson, P. Feltovich, N. Charness, & R. R. Hoffman (Eds.), *Cambridge handbook on expertise and expert performance* (pp. 147–166). New York: Cambridge University Press.

Journal of Educational Psychology. (1921). Intelligence and its measurement: A symposium. *Journal of Educational Psychology*, *12*, 123–275.

Kohn, M. L., & Schooler, C. (1978). The reciprocal effects of the substantive complexity of work and intellectual flexibility: A longitudinal assessment. *American Journal of Sociology*, *84*, 24–52.

Krampe, R. T., & Charness, N. (2018). Aging and expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. Mark Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed., pp. 836–856). New York: Cambridge University Press.

Kubeck, J. E., Delp, N. D., Haslett, T. K., & McDaniel, M. A. (1996). Does job-related training performance decline with age? *Psychology and Aging*, *11*(1), 92–107.

Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science*, *14*, 316–345.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

(NBA) National Basketball Association. (2009). 2007–08 player survey: Height. NBA. www.nba.com/news/survey_height_2007.html

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Lawrence Erlbaum Associates.

Norman, G., Eva, K., Brooks, L., & Hamstra, S. (2006). Expertise in medicine and surgery. In A. Ericsson, P. Feltovich, N. Charness, & R. R. Hoffman (Eds.), *Cambridge handbook on expertise and expert performance* (pp. 339–353). New York: Cambridge University Press.

Nuland, S. B. (2007). *The art of aging: A doctor's prescription for well-being*. New York: Random House.

Owens, W. A., Jr. (1953). Age and mental abilities: A longitudinal study. *Genetic Psychology Monograph*, *48*, 3–54.

OUP (Oxford University Press). (1971). *The compact edition of the Oxford English Dictionary*. New York: OUP.

Polanyi, M. (1966/1983). *The tacit dimension*. Gloucester, MA: Peter Smith.

Robinson, R. (1950). *Definition*. London: Oxford University Press.

Ryle, G. (1949/2000). *The concept of mind*. Chicago: University of Chicago Press.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428.

Schaie, K. W. (1996). *Intellectual development in adulthood: The Seattle longitudinal study*. New York: Cambridge University Press.

Schaie, K. W. (2005). *Developmental influences on adult intelligence: The Seattle Longitudinal Study*. New York: Oxford University Press.

Schooler, C. (2001). The intellectual effects of the demands of the work environment. In R. J. Sternberg & E. L. Gigorenko (Eds.), *Environmental effects on cognitive abilities* (pp. 363–380). Mahwah, NJ: Lawrence Erlbaum Associates.

Simonton, D. K. (1988). *Scientific genius: A psychology of science*. New York: Cambridge University Press.

Snow, R. E. (1996). Aptitude development and education. *Psychology, Public Policy, and Law*, *2*, 536–560.

Spearman, C. (1904). “General intelligence,” objectively determined and measured. *American Journal of Psychology*, 15, 201–293.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*, 403–433.

Sternberg, R. J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology*, *24*, 359–375.

Sternberg, R. J., & Detterman, D. K. (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex.

Stewart, N. (1947). A.G.C.T. scores of army personnel grouped by occupation. *Occupations*, *26*, 5–41.

Swift, E. J. (1910). Relearning a skillful act: An experimental study in neuro-muscular memory. *Psychological Bulletin*, *7*, 17–19.

Wagner, R. K. (2000). Practical intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 380–395). New York: Cambridge University Press.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology*, *49*, 436–458.

Wagner, R. K., & Sternberg, R. J. (1987). Tacit knowledge in managerial success. *Journal of Business and Psychology*, *1*, 301–312.

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology*, *99*, 397–420.

Willis, S., & Tosti-Vasey, J. L. (1990). How adult development, intelligence, and motivation affect competence. In S. L. Willis & S. S. Dubin (Eds.), *Maintaining professional competence: Approaches to career enhancement, vitality, and success throughout a work life* (pp. 64–84). San Francisco: Jossey-Bass.

Yates, F. (1966). *The art of memory*. London: Routledge and Kegan Paul.

PART IX

Folk Conceptions of Intelligence

49 Self- and Other-Estimates of Intelligence

Aljoscha C. Neubauer and Gabriela Hofer

Today, intelligence can possibly be regarded as one of the best-researched constructs in the whole field of psychology. The structure of intelligence is clearly defined, as proposed by Carroll's (1993) widely accepted structure of abilities and its further development in the form of the Cattell-Horn-Carrol (CHC) model, which essentially is a fusion of the Carroll model with that of Cattell and Horn (McGrew, 2009). Although structural models of intelligence differ in the number and naming of factors, most intelligence researchers agree that the subfactors are positively correlated (i.e., the so-called positive manifold), which allows one to compute a total index of a general cognitive ability (GCA), or mostly simply termed *g* (Deary, 2012). *G*, and partially also the subfactors, has been well researched from many different perspectives: There are findings on its neural correlates (e.g., Haier, 2017) and the elementary cognitive processes that are involved (e.g., Sheppard & Vernon, 2008). Furthermore, a considerable body of research supports its validity with respect to predicting educational (e.g., Roth et al., 2015) and professional (e.g., Schmidt & Hunter, 2004) as well as health and longevity outcomes (for an overview, see Deary, Weiss, & Batty, 2010). Moreover, *g* is considered a partially heritable, polygenic, fitness-related trait that has evolved to solve adaptive problems like survival and reproduction (Chiappe & MacDonald, 2005; Cosmides & Tooby, 2002; Davies et al., 2011).

All that said, the pervasive importance of intelligence, as measured by psychometric tests for educational, professional, and even "life" success, has also been challenged on several accounts. As an example, it has been argued that performance on intelligence tests could depend strongly on the test setting or situation, which might be anxiety-inducing. This could happen, for example, through the so-called stereotype threat, which is a phenomenon that describes the influence that implicitly activated stereotypes can have on performance (Steele & Aronson, 1995). According to Steele (1997), a negative stereotype about a group one belongs to leads to fear and self-doubt, thereby impairing working memory and hampering cognitive performance. Other concepts assume that an individual's performance depends not only on the individual's ability but also on the individual's belief in it. Relevant to these claims are the research lines into "self-variables," such as self-concept (Shavelson, Hubner, & Stanton, 1976) and self-efficacy (Bandura, 1977). While not the topic of this chapter, they share an assumption with self-estimates of abilities, which are the focus here: A person's level of ability is not the same as the person's belief in how good they are. While the research on self-efficacy and especially on self-concept is mostly about self-belief in rather specific domains such as, for example, math or language performance in school, this chapter focuses on more broadly defined self-estimates of abilities, namely those of commonly assessed factors in intelligence tests.

Everyone has experienced situations in which they have had to assess their own abilities or performance. We constantly make such assessments, with some of them having only small consequences for our lives and others having an important influence on our future. In the days preceding an important exam, for example, students have to ask themselves, "Has my preparation been sufficient?" Before deciding on a degree or career to pursue, individuals may ask themselves, "Do I have the necessary skills to be successful in this domain?" Thus, self-estimates shape important life choices (Ackerman & Wolman, 2007; Ehrlinger & Dunning, 2003). Furthermore, despite the existence of a vast number of different objective measures of abilities, self-estimates are widely used in practical fields, such as career counseling (Freund & Kasten, 2012). Self-report measures are less time-consuming, easier to administer, and, overall, more economical than psychometric tests (Herreen & Zajac, 2018). Such measures might also provide additionally useful information (Ackerman & Wolman, 2007; Freund & Kasten, 2012). The accuracy of self-estimates is therefore interesting from both a theoretical and a practical standpoint. In the second section of our chapter, we attempt to integrate the existing literature on the accuracy of self-estimates. As the reader will see, there is a large body of research on these questions, well documented by several meta-analyses. Before we review these meta-analyses and other relevant literature, however, a second main aim of this chapter shall be introduced.

In the same way as we can ask about the accuracy of self-perceptions of abilities, we can ask whether others might have the same or maybe even a better perception of an individual's abilities than the individual themselves. These *others* might either be acquaintances, like peers, friends, teachers, or parents, or else even individuals who do not personally know the "target person." The last of these is known in the literature as "zero acquaintance." Compared to self-estimates, there is much less research on the accuracy of other-estimates of abilities. Nevertheless, we consider it promising to look at this body of evidence, especially when comparing it to the self-perspective. This approach has so far focused mostly on the Big Five personality traits (e.g., Vazire, 2010) but practically no research on abilities is available, which is why we finally propose a new research agenda that brings self- and other-perspectives on human intelligence factors together in one model.

To sum up, this chapter seeks to provide (1) an up-to-date outline of the current state of research on self-estimates of intelligence; (2) a shorter, solely narrative review of the "other-estimates of intelligence" literature; and (3) an attempt (the first) to bring both lines of research together.

The Relationship Between Self-Estimated and Objectively Measured Intelligence

The first studies on the accuracy of self-estimates were conducted more than 100 years ago (Cogan, Conklin, & Hollingworth, 1915; for a historical account, see Ackerman & Wolman, 2007). Since then, a vast amount of research on this topic has accumulated, leading to the publication of several meta-analyses (e.g., Freund & Kasten, 2012; Mabe & West, 1982; Ross, 1998) and even one metasynthesis, which is a combination of several meta-analyses (Zell & Krizan, 2014). Given the topic of this handbook, we focus our discourse on research conducted on the topic of accuracy of self-estimates of intelligence. However, because a lot of the existing body of literature generalizes to abilities as a whole, we will include relevant research from this broader area as well. Most research on the accuracy of self-estimates relies on correlations with criterion measures such as psychometric tests, grades, or performance estimates given by others (Dunning & Helzer, 2014). Thus, we will begin by reviewing existing correlative findings, before moving on to other measures of accuracy.

In an early meta-analysis on the question of the accuracy of self-estimates of ability, Mabe and West (1982) included findings of fifty-five studies that investigated diverse abilities ranging from intelligence to sewing. They found small to medium associations between self-estimated and performance measures of abilities (r_{mean} = 0.29), even though results varied strongly between studies (SD = 0.25). Of greater importance to the topic of our chapter is a more recent meta-analysis by Freund and Kasten (2012), which focused entirely on the accuracy of self-assessed intelligence. The authors combined the findings of forty-one studies conducted between 1915 and 2009. They included studies that contained self-estimates and psychometric measures of verbal, numerical, and spatial abilities, as well as of overall intelligence. Results revealed a positive, albeit only moderate, association between self-estimates and psychometric measures (r_{mean} = 0.326, 95 percent CI [0.284, 0.368]). They also investigated whether assessing self-estimated numerical, spatial, or verbal ability directly resulted in higher levels of accuracy compared with self-estimates of overall intelligence. However, only self-estimates of numerical abilities, but not of spatial or verbal abilities, resulted in comparably more accurate self-estimates. Given the low to at best medium validities of self-estimates of cognitive abilities, the authors proposed applying them to gain additional information but warned against using them as a substitute for standardized psychometric tests.

Returning to the broader field of self-estimates of abilities in general, we cannot fail to mention work by Zell and Krizan (2014), who confirmed the low to moderate relationship between self-estimates and objective measures. They used a novel approach called "metasynthesis" and assembled results from twenty-two meta-analyses published between 1982 and 2012 that had focused on academic ability, intellectual ability (the aforementioned meta-analysis by Freund and Kasten), language competence, medical skills, memory ability, nonverbal skills, perceived knowledge, sports ability, and vocational skills. The overall relationship between self-estimates and criterion measures (in this case psychometric tests, grades, or evaluations by superiors) was also moderate (r_{mean} = 0.29) and largely depended on the type of ability that was investigated, with correlations ranging from 0.09 (interpersonal abilities) to 0.63 (language competence). The high validity of self-estimated language competence is interesting, especially when compared to Freund and Kasten's results on verbal abilities. The diverging results might have arisen because the meta-analysis on language competence (Ross, 1998), which Zell and Krizan referred to, did not focus on overall verbal abilities but on second-language skills specifically, a domain in which individuals usually receive repeated feedback while learning.

Thus far, we have seen that, although self-estimates of intelligence and other abilities are related to performance measures, correlations are only moderate (see also Sternberg, 1985; Sternberg et al., 1981). It is important to note that all three meta-analytic/meta-synthetic studies included a wide range of moderating variables. We will come back to these moderation analyses in the following section, where we summarize potential explanations for the low to at best moderate association between self-estimated and objectively measured abilities.

Recently, Dunning and Helzer (2014) have criticized relying solely on the correlation coefficient to investigate the accuracy of self-estimates. They argued that even though correlation coefficients give us important information about the (lack of) accuracy of self-estimates, other measures can also provide interesting insights. Additionally, it has to be kept in mind that correlation coefficients alone contain no information about the direction of misestimation. Even though both over- and underestimation of one's abilities could lead to negative outcomes, they probably do so in different ways (Ackerman & Wolman, 2007). When someone underestimates their own abilities, they might fail to attain goals by not even trying, despite the fact that the goals would actually be achievable. Overestimation of one's intellectual capacities might also lead to nonattainment of goals, for example if it is related to too little effort. Even though most research in this area is correlational, we would provide an incomplete picture if we failed to mention the other methodological approaches that have been applied.

Looking at the difference between self-estimated and objectively measured abilities and traits, researchers have established that there is an overall tendency for individuals to overestimate themselves, that is, that average self-estimated abilities are significantly higher than the average actual performance (Dunning, Meyerowitz, & Holzberg, 1989). This effect is called the above-average or better-than-average effect (Alicke & Govorun, 2005) and has been found across a wide range of domains ranging from humor (Kruger & Dunning, 1999) to driving ability (Horrey et al., 2015). Similar findings have been shown for intellectual performance (e.g., Ehrlinger, Mitchum, & Dweck, 2016; Kruger & Dunning, 1999). As an example, university students were found to rate their intelligence on all eight domains of Gardner's multiple intelligences as above-average compared to other students at their institution (Visser, Ashton, & Vernon, 2008). A more detailed discussion of this effect and potential explanations can be found in Krizan and Suls (2008).

Another phenomenon that is often discussed in relation to the above-average effect (e.g., Dunning, 2011) is overclaiming alleged knowledge of concepts that do not exist (Paulhus et al., 2003). To measure this effect, Paulhus and colleagues (2003) have asked individuals about their familiarity with topics belonging to diverse domains within academic or everyday knowledge (e.g., nuclear fusion or asteroids) and interspersed them with fake topics (e.g., "cholarine" or "ultra-lipid"). Remarkably, participants claimed to be familiar with about 25 percent of the fake topics, compared to around 44 percent of the actually existing topics. In another study, the accuracy of claimed knowledge, that is, the proportion of correctly claimed knowledge to incorrectly claimed knowledge, was found to be positively associated with self- and peer-rated as well as psychometric intelligence, whereas a bias toward overclaiming, that is, an overall tendency to claim knowledge of both existent and nonexistent items, correlated positively with subclinical narcissism (Paulhus & Harms, 2004).

The picture painted by research on the above-average effect might appear rather dark, especially when one considers the potentially negative consequences of overestimation. However, some of these and other studies have suggested that overconfidence in one's abilities is not as universal as initially believed. In this research, and similar to correlational studies on the accuracy of self-estimates, several moderators have received theoretical and empirical attention. It has even been suggested that some factors are actually related to an underestimation of one's own abilities. We will discuss these variables in more detail in the next section.

Moderators of the Accuracy of Self-Estimates of Intelligence

Two main sources of moderating variables of the accuracy of self-estimates have been discussed: measurement conditions related to the self-assessment process, on the one hand, and characteristics of the person assessing themselves, on the other hand. As the meta-analyses summarized in the previous section mainly focused on the first group of moderators, we will begin our discussion with them. In the second subsection, we will review evidence on moderators within the person.

Measurement conditions

In their early meta-analysis, Mabe and West (1982) identified what they called "favorable measurement conditions" that are positively related to the validity of self-estimates of abilities. Accuracy of self-estimates was especially high when (1) self-estimate measures used relative scales that included a comparison group, (2) participants expected that their self-estimates would be compared with an objective criterion, (3) participants were assured that they would remain anonymous, and (4) participants had the opportunity to gain prior experience with self-estimates within the same study. Two other predictors were associated with comparatively high validity: self-estimate measures that evaluated specific kinds of performance compared with overall ability; and positioning the self-estimate measures after, instead of before, the criterion measures. Thirty years later, Freund and Kasten (2012) replicated the results regarding the use of relative instead of absolute scales and found an additionally beneficial effect of including an explicit reference group. In contrast to Mabe and West, they found no significant effect of whether self-reports preceded or followed objective measures. This lack of an order effect was confirmed by Zell and Krizan (2014), who have observed additional moderators while using a descriptive rather than statistical approach toward moderation. The accuracy of self-estimates of abilities appeared to be comparatively high if (1) self-estimates were specific compared with global, (2) objective instead of subjective performance measures were applied, (3) participants were familiar with the task, and (4) the tasks, for which participants had to evaluate their performance, were of low complexity.

In conclusion, it can be said that there are several measurement conditions that impact the validity of self-estimates of abilities. As Mabe and West (1982) had already stated more than thirty years ago, favorable measurement conditions are easy to apply and small changes in designs can lead to improvements in validity. Some researchers have already applied these suggestions. As an example, Ackerman and Wolman (2007) wanted to investigate the accuracy of self-estimated verbal, numerical, and spatial ability under optimal conditions, while also testing changes of accuracy over time. Their results indicated medium to large correlations between self-estimates and objective measures, with slightly higher validity of self-estimated numerical than spatial and particularly verbal ability. Additionally, they observed a significant increase of accuracy from before to after the intelligence test for verbal and spatial but not numerical ability. This – again − raises the question of the importance of the positioning of self-estimates within studies but could also be seen as confirmation of the beneficial effect of prior experience with self-estimates. If making accurate self-estimates in at least some domains is indeed something individuals can learn from a little practice, this result could have essential implications for practical fields like counseling and should also be considered when conducting research in this area.

Moderators Within Individuals

On overconfidence and incompetence: the Dunning-Kruger effect. Several characteristics within the person who has to judge their abilities may influence the accuracy of self-estimates. Among them, one has received an extraordinary amount of attention: the level of ability itself. More specifically, it has been argued that individuals with low abilities are particularly bad at assessing themselves. This effect is called the Dunning-Kruger effect, named after the two scientists who established it (Kruger & Dunning, 1999). The authors proposed that participants with low abilities suffer a dual burden: "Not only do they reach mistaken conclusions and make regrettable errors, but their incompetence robs them of the ability to realize it" (Kruger & Dunning, 1999, p. 1132). The problem of these individuals lies with the fact that one often needs the same set of skills to judge one's abilities as one would need to give correct responses in a test of the same abilities. In order to estimate one's performance in a mathematical test accurately, one must possess a certain level of mathematical skills. Thus, these individuals suffer from a lack of relevant metacognition.

In their early studies, Kruger and Dunning (1999) asked individuals to rate their ability to recognize humor, their logical reasoning skills, and their grammatical ability relative to their peers. They then compared these self-estimates with objective ability measures. Results indicated that individuals whose measured abilities were in the lowest 25 percent of those tested gave the largest overestimations. They overestimated their abilities relative to others by around 40 to 50 percent points. The higher the measured abilities of a person were, the less the person overestimated their abilities.

Individuals whose test performance placed them among the top 25 percent within the sample actually underestimated themselves. Interestingly, the lowest-quartile performers were not only unable to estimate their own abilities accurately but showed the same problems in assessing the performance of others. They further were unable to make more accurate self-estimations after having been confronted with the performance of others. In line with the authors' suspicion that a lack of ability leads to misjudgments of the exact same ability, the quality of self-estimates of logical reasoning did improve after individuals had received some training in logical reasoning. Since the discovery of the Dunning-Kruger effect, several studies have shown similar effects in diverse domains (for summaries of the literature, see Dunning, 2005, 2011).

Despite the broad body of literature supporting the Dunning-Kruger effect, it has to be noted that it has also received some criticism and several alternative explanations have been proposed. Some of these alternative explanations refer to statistical artifacts such as regression to the mean combined with measurement inaccuracy (Krueger & Mueller, 2002) or problems related to floor effects and an asymmetric distribution in Kruger and Dunning's original sample (Krajč & Ortmann, 2008). However, two studies found support for the existence of the Dunning-Kruger effect while correcting for measurement error (Ehrlinger et al., 2008; Feld, Sauermann, & de Grip, 2017) and another (Schlösser et al., 2013) disputed the alternative explanation proposed by Krajč and Ortmann (2008). Burson, Larrick, and Klaymann (2006) have argued that low task difficulty was actually responsible for the Dunning-Kruger effect: For easy tasks, nearly all individuals would report performing better than their peers, leading to overestimation in individuals with low performance. Ehrlinger and colleagues (2008) found consistent effects of overestimation in low performers over five studies, in which individuals faced challenging real-life tasks and argued that the alternative explanation regarding task difficulty did not hold true. One might also argue that poor performers are not sufficiently motivated to admit their shortcomings due to a lack of incentives (for a discussion, see Dunning, 2011). Yet Ehrlinger and colleagues (2008) showed that neither monetary ($5 or even $100) nor social incentives improved the accuracy of assessments. Lastly, Miller and Geraci (2011) argued against the notion that low performers were completely "unaware" by showing that they report lower confidence in their (faulty) assessments than did high performers.

In summary, there is a large body of literature supporting the Dunning-Kruger effect. Even though some authors provide compelling alternative explanations, which should be kept in mind, other researchers have repeatedly disputed them. Additionally, the effect is in line with studies that showed that high levels of intelligence are associated with more accurate self-judgment (for a review, see Mabe & West, 1982). As regards the dire picture that studies on the accuracy of self-estimates have painted, research on the Dunning-Kruger effect offers one promising avenue for enhanced accuracy: Individuals improved in their self-assessment skills after having improved the skills that were to be assessed (Kruger & Dunning, 1999).

Personality. Several personality characteristics are related to self-estimates of intelligence. In the five-factor model, extraversion, conscientiousness, and openness to experience correlate positively with self-estimated intelligence (e.g., Visser et al., 2008). Neuroticism and agreeableness are negatively associated with self-estimates of intelligence (e.g., Furnham, Moutafi, & Chamorro-Premuzic, 2005). Unfortunately, there is only limited literature that has directly investigated moderating effects of the five-factor traits on the accuracy of self-estimates of intelligence. Jacobs, Szer, and Roodenburg (2012) found low extraversion and high agreeableness to be associated with more accurate self-estimates. Soh and Jacobs (2013) investigated this matter for men and women separately and found high conscientiousness and openness to be positive predictors of accuracy in men and low extraversion to be associated with higher accuracy in women.

Narcissism has also repeatedly been discussed in relation to self-estimates of intelligence. Narcissism correlates with self-estimates of abilities or performance but not with actual performance measures (Farwell & Wohlwend-Lloyd, 1998; Gabriel, Critelli, & Ee, 1994; Goncalo, Flynn, & Kim, 2010). In a recent meta-analysis, Grijalva and Zhang (2016) concluded that narcissism is associated with overly positive self-views on a wide range of domains, with intelligence as one of those domains. In other words, individuals high in narcissism might believe themselves to be better performers than others but fail to prove so in objective measures.

Also worth mentioning in this context are (implicit) theories of the modifiability of intelligence. These refer to two opposed views individuals can hold on intelligence (Dweck & Leggett, 1988): In the *entity view*, individuals believe intelligence is fixed and nonchangeable, whereas individuals holding an *incremental view* see intelligence as more flexible and changeable over time. Individuals holding the entity view overestimate their performance to a higher degree than those holding the incremental view (Ehrlinger et al., 2016).

Viewing the existing body of literature in this area, numerous studies have investigated how personality traits could be related to self-estimates of intelligence (for trait complexes, i.e., compounds of traits, see Ackerman & Wolman, 2007) but there are lamentably few that have directly investigated associations between personality and the accuracy of self-judgments. This is surprising, since nearly forty years ago, researchers were already asking themselves whether the capacity to make accurate self-judgments of abilities might be a trait in itself (Mabe & West, 1982). Future research investigating the (lack) of accuracy of self-estimates of intelligence might therefore benefit from including personality traits. Identifying personality traits related to more and less accurate intelligence self-estimates might then also benefit individuals who are in career counseling.

Gender. Very similar to personality traits, a considerable body of research has investigated gender differences in self-estimates of intelligence. Syzmanowicz and Furnham (2011) brought together the sometimes inhomogeneous results in four meta-analyses encompassing overall, mathematical, spatial, and verbal intelligence. Their results indicated higher self-estimates in men compared with women for overall ($d = 0.37$), mathematical ($d = 0.44$), and spatial ($d = 0.43$) intelligence but not for verbal intelligence ($d = 0.07$). The authors pointed out that, for all but spatial abilities, gender differences in self-estimates exceed those known from the literature on gender differences in psychometric intelligence (for a discussion of gender differences in psychometric intelligence, see Halpern and Wai, Chapter 14 in this volume). Thus, these results largely point toward more accurate self-estimates or underestimation in women and overestimation in men.

The account suggesting underestimation in women and overestimation in men was recently confirmed for fluid and general intelligence in older adults (Herreen & Zajac, 2018). Also in line with this proposition, male adolescents give higher self-estimates of figural, numerical, and reasoning abilities than do female adolescents and do so even when actual abilities are controlled for (Steinmayr & Spinath, 2009). Freund and Kasten (2012) included gender as a moderator in their meta-analysis and found no effect on the accuracy of self-estimates of intelligence. However, they only investigated whether accuracy depended on whether the participants of a study consisted of women, men, or both genders.

In short, many studies have investigated the influence of gender on self-estimates of intelligence but hardly any research has been done on its impact on accuracy. The importance of this question was demonstrated in a study conducted by Ehrlinger and Dunning (2003), who found that women's lower self-estimates of the ability to "reason about science" were related to lower self-estimated performance in a scientific reasoning test, even though men and women performed equally well. Consequently, women were also less willing to take part in a science competition than men.

Other-Estimates of Intelligence

Compared with self-estimates, the literature on other-estimates of intelligence appears to be rather sparse (see also Denissen et al., 2011). In this section, we will first discuss findings on the accuracy of estimates made by known others, before then summarizing literature on those made in zero-acquaintance contexts.

Probably the largest number of studies that have been conducted on the validity of intelligence estimates by known others have focused on the accuracy of teacher- and parent-estimates of children's intellectual abilities. The interest in this topic is of little surprise, since both teachers and parents can shape a child's educational future, for example by proposing specific schools to attend. Steinmayr and Spinath (2009) found that parents of adolescents were fairly accurate in estimating their child's performance in intelligence tests. Mostly medium accuracy levels were found for verbal ($r_{\text{boys}} = 0.26$, $r_{\text{girls}} = 0.35$), numeric ($r_{\text{boys}} = 0.34$, $r_{\text{girls}} = 0.36$), figural ($r_{\text{boys}} = 0.28$, $r_{\text{girls}} = 0.34$), and reasoning abilities ($r_{\text{boys}} = 0.26$, $r_{\text{girls}} = 0.34$). Recently, Zippert and Ramani (2017) showed that parents' estimates of their preschoolers' numerical abilities were already reasonably accurate. Sommer, Fink, and Neubauer (2008) found that both teachers and parents are relatively good at judging overall intelligence ($r_{\text{teachers}} = 0.56$, $r_{\text{parents}} = 0.50$) and moderately good at judging creativity ($r_{\text{teachers}} = 0.34$, $r_{\text{parents}} = 0.24$), while neither are successful at estimating the pupils' social competence.

Less research in this area has focused on adult populations. Denissen and colleagues (2011) investigated the development of accuracy of peer-estimates of intelligence in university students. After knowing their peers for only one week, individuals could already estimate the peers' intelligence with low to medium accuracy (r = 0.25). Interestingly, accuracy rates were similar in size after four (r = 0.27) and eight months (r = 0.22) of acquaintance. Peer- but not self-ratings directly predicted academic success, that is, grades and staying in school. Borkenau and Liebler (1993) asked cohabiting pairs − most of them romantic couples − to estimate their own and each other's intelligence and found medium accuracy for both types of estimates (r_{other} = 0.29, r_{self} = 0.32). The researchers also filmed the individuals during the seemingly mundane tasks of entering a room and reading a weather forecast aloud. They later presented these videotapes with or without sound to strangers and asked them to rate the participants' intelligence. Strikingly, stranger-ratings were even slightly more accurate (r = 0.43) than those by participants themselves and their close others. Strangers failed to make accurate judgments only in the silent-video condition. These findings raise the question of sources of accuracy of other-estimates: How is it possible that an individual who has never met me can estimate my intelligence with a similar degree of accuracy to someone living with me or even myself?

Two years after their initial study, Borkenau and Liebler (1995) reanalyzed their data to shed more light on sources of accuracy in zero-acquaintance contexts and referred to Brunswik's lens-model as a theoretical framework (Brunswik, 1956, as cited in Borkenau & Liebler, 1995). This model considers two aspects when it comes to accuracy: cue utilization and cue validity. While cue utilization refers to the degree to which a signal is used by individuals to estimate another person's intelligence or other characteristics, cue validity refers to the degree to which a signal is actually associated with the estimated characteristic. For judgments to be made with high accuracy, the utilized cues must also be valid. For their reanalyses, Borkenau and Liebler scanned their video material for potential cues to intelligence and correlated the expression of each of these cues in the targets with stranger-ratings of intelligence (cue utilization) and measured intelligence (cue validity). Several valid cues for intelligence were identified, with most of them speech-related, like pleasantness of speech (r = 0.29) or understandability (r = 0.44). Strangers used a wide range of cues when making intelligence judgments, with speech-related cues but also visual cues like attractiveness being utilized.

In a similar study, Reynolds and Gifford (2001) filmed high school students while they answered intellectually challenging questions like, "What do you see as the future of the world environment?" and showed these videos to strangers, who judged the students' intelligence afterwards. Accuracy of stranger-estimates varied, depending on whether the videos were presented audiovisually (r = 0.22), only visually (r = 0.30), or only auditorily (r = 0.38), and only reached statistical significance in the auditory-only condition. Speech rate (r = 0.40) and number of words used (r = 0.50) were valid cues of intelligence. Utilized cues were also mainly speech-related, such as the use of standard speech.

These studies point out the importance of speech-related cues for intelligence judgments. Notwithstanding, there are also valid and utilized visual cues of intelligence. As an example, Zebrowitz and colleagues (2002) investigated zero-acquaintance estimates of intelligence based on pictures of individuals taken over the course of their lives. There were small correlations between estimates and measured intelligence in childhood (r = 0.14), puberty (r = 0.18), and middle adulthood (r = 0.15), which were rendered nonsignificant when attractiveness was controlled for. Recently, Lee and colleagues (2017) found that a combination of facial characteristics extracted from pictures significantly explained 2 percent of variance in measured intelligence. Strangers inferred intelligence from these pictures with low accuracy (r = 0.15) and the effect remained significant when controlling for attractiveness (cf. Zebrowitz et al., 2002). Facial characteristics mediated this relationship, indicating that they might be used to make accurate intelligence estimates.

In conclusion, unknown others seem to know us astonishingly well, attaining similar accuracy levels as our close others or even ourselves. They have some sense of our intelligence after seeing pictures or short videos of us (see also Borkenau et al., 2004). Speech-related cues seem to be more valid indices of intelligence than solely physical ones. Accordingly, videos are related to higher accuracy rates than static pictures. Findings on the role of target attractiveness seem to remain inconclusive. Altogether, apart from studies on cue utilization and validity, there seems to be little research on mediating and moderating variables that affect the accuracy of intelligence estimates by others. One other variable has received repeated attention in the relevant literature: gender. Some studies have indicated that both known (Steinmayr & Spinath, 2009; but cf. Sommer et al., 2008) and unknown others (Borkenau et al., 2004) give higher intelligence estimates for men than for women. There is also some indication that controlling for gender (and age) is related to lower accuracy of other-ratings (Borkenau et al., 2004; but cf. Borkenau & Liebler, 1993). However, presently, the literature on this topic seems to be inconclusive and further research is needed.

Discussion

Summing up the reviewed literature on intelligence and its most common subfactors, it can be concluded that self-estimates can be surprisingly inaccurate and often do not surpass other-estimates regarding accuracy. Indeed, they can even be more inaccurate than estimates by others, even if those are made in zero-acquaintance contexts. In the following, we will first discuss self-estimates and then other-estimates. For self-estimates, we will discuss (1) potential causes/sources of inaccuracy, specifically those of overestimation, and (2) its consequences on (a) the use of self-estimates for diagnostic purposes and (b) the development of pertinent ability. We will then turn to other-estimates and discuss their accuracy in light of their source. Other-estimates could have different implications, depending on whether they come from acquaintances or are zero-acquaintance estimates. Acquaintances' perspectives might affect the development of the individual's respective ability and their self-concept, among other things. Findings from zero-acquaintance judgments, on the contrary, might be more relevant for human-resources contexts, such as interviews, assessment centers, and so on. Finally, we will try to bring both perspectives together, an approach that hitherto has rarely been followed.

Explaining the Low Accuracy of Self-Estimates

Dunning and Kruger explained the imprecise self-estimations that particularly the less able are susceptible to (i.e., the Dunning-Kruger effect) through two factors: lack of meta-skills and lack of knowledge. In their experiments (e.g., Kruger & Dunning, 1999), manipulations to increase meta-skill and to improve knowledge on how the task-in-question can be solved successfully led to more realistic self-perceptions. To quote Kruger and Dunning (1999): "one way to make people recognize their incompetence is to make them more competent" (p. 1131). However, it can be questioned whether these interventions could apply to the domain of intelligence, since it still seems unclear whether intelligence can be enhanced, for example through working memory training (Shipstead, Redick, & Engle, 2012) or brain-stimulation (Neubauer et al., 2017).

Other explanations for the persistence of overconfidence advanced by Kruger and Dunning (1999) refer to the lack of (negative) feedback most of us receive when we do not perform well, especially when it comes to skills and abilities. And, even when we get negative feedback, we all too frequently attribute it to external factors, such as a lack of luck, having had a "bad day," or having had "not enough sleep," and so on. We often tend to overgeneralize positive feedback and minimize or trivialize negative feedback (Mezulis et al., 2004).

Consequences of Inaccurate Self-Estimates

Implications for the assessment of intelligence. As previously mentioned, self-estimates are widely used in practical fields such as career counseling (Freund & Kasten, 2012). As self-estimates are usually collected via classical questionnaires, that is, self-report measures, they have the advantage of being overall more economical than psychometric ability tests (Herreen & Zajac, 2018). From that viewpoint, it could be desirable to use them as "proxies" for "true" ability instead of ability tests. The surprisingly low correlations reported in the above-mentioned literature clearly advise against that course of action. Although several authors have proposed measures to improve accuracy, the resulting correlations still cannot be considered high enough to allow for a replacement of ability tests. Freund and Kasten (2012) concluded that self-estimates – at best – could be used to assess how realistically a person views themselves. The low mean correspondence of 0.33 in their meta-analysis points to only around 11 percent of shared variance.

Moreover, several studies (see "Moderators Within Individuals" section) have shown that self-report measures of abilities exhibit considerable associations with personality traits. Herreen and Zajac (2018) showed that self-estimates are only highly valid for crystallized intelligence (around 0.5), whereas *g*, fluid intelligence (G*f*), and visual-processing ability (G*v*) showed the typically low correlations of 0.2–0.3 reported in previous meta-analyses. In fact, in some cases, self-estimates correlated even higher with some of the Big Five traits (up to 0.53) than with performance measures. In multiple regressions on self-estimates as the criterion, the Big Five traits were similarly strong or in some cases even stronger predictors than abilities. The authors concluded “the regression models reported herein indicate that self-estimated Gf and Gv more strongly reflect one’s personality than their actual cognitive performance ability” (Herreen & Zajac, 2018, p. 12). Similar findings have been reported by Williams and colleagues (2017) for a self-report scale of attentional control, only they found even lower associations with actual performance measures (practically zero), while correlations with the Big Five were again up to 0.5.

Ackerman and Wolman (2007) have pointed out that self-estimates correlate most highly with self-efficacy and self-esteem, another finding that suggests that self-estimates do not reflect ability per se but rather the belief in possessing the ability to a certain degree. These findings support the notion of some authors that self-estimates might provide additionally useful information compared with performance measures (Ackerman & Wolman, 2007; Freund & Kasten, 2012), an assumption that has also received some empirical support (Furnham, Zhang, & Chamorro-Premuzic, 2005).

The findings on associations with personality traits give rise to another interesting aspect: Although self-estimate measures of abilities clearly ask for maximum behavior, they exhibit high correlations with typical performance measures. It is a common finding that maximum versus typical performance measures often correlate poorly (Beus & Whitman, 2012). It has even been shown that situational judgment tests of emotional competencies show divergent validity patterns, depending on whether maximum versus typical performance instructions are given, with maximum performance correlating with intelligence and typical performance correlating with the Big Five (Freudenthaler & Neubauer, 2007). It can be questioned whether people responding to self-estimate measures really have their maximum performance in mind or rather report their typical, daily behavior. Furthermore, one could ask whether typical behavior might not be more informative regarding the prediction of real-life behaviors. However, especially for the field of intelligence, it seems rather unlikely that self-estimate measures could supersede the high validities of objective tests, which, at 0.5–0.6, are among the highest in psychology (e.g., Roth et al., 2015; Schmidt & Hunter, 2004).

In brief, self-estimates are an inadequate substitute for psychometric intelligence tests. The more promising question is whether they could add incremental information to the prediction of real-life criteria. Even if they do, it has to be questioned whether this increment is really provided by self-estimates per se and not the “personality aspects” that are obviously a part of them. As at least some of the Big Five contribute to school and professional success (Barrick, Mount, & Judge, 2001; Poropat, 2009), it could simply be their contributions that enhance the prediction.

Effects of over- versus underestimations on the development of pertinent ability. In addition to the psychometric implications, we can also ask whether misestimations might have long-term effects on the development of an ability, specifically in terms of the costs and benefits of over- versus underestimations. While one could assume that slight overestimations might be “encouraging” and increase (intrinsic) motivation, large overestimations might, in the long term, lead to faults and failures in professional contexts or in education. Conversely, underestimations might be detrimental to achievement motivation: The person might be less motivated to tackle higher and more ambitious educational or professional goals.

Until very recently, these effects had not been studied. While this appears to remain true for underestimation, an especially informative article on the impact of overestimation has been published: In a series of six (partly experimental) studies, Sanchez and Dunning (2018) analyzed the development of knowledge and contrasted it to individuals' (over-)confidence. While beginners were initially not overconfident, they quickly developed the so-called beginner's bubble, that is, they strongly overestimated their performance. With more experience, individuals not only improved their performance but also developed more realistic self-views. In addition, the authors discussed problems of self-selection that might arise from overestimation: What are the implications of overconfidence on the process of choosing careers, education, or specific jobs? People tend to volunteer for tasks in which they consider themselves competent (Camerer & Lovallo, 1999; Koellinger, Minniti, & Schade, 2007) and might show more overconfidence in these tasks, which in the long run might shape further learning. However, this has not been studied empirically so far (for a discussion of perspectives, see Sanchez & Dunning, 2018). If a bit of learning strongly increases confidence, learners face the dilemma that learning is necessary to acquire abilities but also leads people to overestimate these abilities. Sanchez and Dunning advise being mindful of English philosopher R. G. Collingwood's conclusion that "people cease to be beginners in any craft or science, and become instead masters, at the moment they realize they are going to be beginners for the rest of their lives" (Sanchez & Dunning, 2018, p. 26). While this commentary might apply to the development of knowledge or expertise in some domains, it might not apply to the trait of intelligence, which seems – at least from the beginning of adulthood – not really to be highly malleable. Therefore, the following question remains open: How can individuals arrive at a valid estimate of their inherent abilities in order to avoid misguided educational or job decisions? Should they rely more on the estimates of other persons, such as their parents, peers, or teachers? We will come back to this issue after a short discussion of another perspective on the potential consequences of overly positive self-views.

Intra- and interpersonal effects of self-enhancement. Overly positive self-views have also been studied thoroughly in social psychology (Dufner et al., 2012, 2018; Sedikides & Alicke, 2012, in press), in which research has focused on phenomena such as self-enhancement, that is, the motive to see oneself more positively than external criteria would warrant (Sedikides & Alicke, in press). This research line focuses less on performance aspects of self-enhancement and more on personal and interpersonal adaptation, that is, effects on personal well-being and on developing and maintaining personal relationships. Interestingly, here self-enhancement turns out as a rather positive – at worst neutral – psychological trait. The most recent meta-analysis (Dufner et al., 2018) on this topic analyzed studies that related self-enhancement to indicators of personal (life satisfaction, positive affect, negative affect, depression) and interpersonal (informant reports of social valuation, agency, communion) adjustment. While results showed small (mostly $r < 0.2$) but generally positive relationships of self-enhancement with personal adjustment, relationships with interpersonal adjustment were even smaller and depended on the level of acquaintance.

The implications for research on self-estimates of intelligence, however, are currently unclear. So far, self-enhancement research has mostly used self-reports alone and rarely compared self-estimates with objective criteria. Furthermore, it has only roughly distinguished between domains. Findings basically show that more optimistic people with more positive and less negative emotions also have higher tendencies toward self-enhancement, based mostly on the correlation of self-reports with self-reports. It is unclear, however, how this finding translates into "real performance." Therefore, we are currently left with a somewhat puzzling picture of (mostly) performance-based research that points to the dangers of overestimation versus (mostly) self-report–based research that sees positive to at worst neutral effects of self-enhancement.

Other-Estimates of Intelligence

As shown above, other-estimates of intelligence have been less studied than self-estimates. Surprisingly, other-estimates seem to result in sometimes higher, sometimes similar, but rarely lower correlations with psychometric intelligence tests than self-estimates. Additionally, other-estimates can provide even better predictions of educational outcomes (grades and dropout). In particular, teachers and parents can provide quite valid estimates. For adults, there is much less research and existing studies show moderate correlations when acquaintances are asked.

The most surprising outcome of research into other-estimates of intelligence comes from zero-acquaintance studies, which show moderate to high correlations when unacquainted individuals judge the intelligence of a person who is only briefly presented in a video clip employing a standardized situation. As mentioned, this finding clearly provokes the question of how it can be possible that a stranger can estimate my intelligence with a similar or even higher degree of accuracy than someone knowing me very well or even than myself, a question only partially answered by existing studies on cues of intelligence. What seems particularly important is that the other-perspective on intelligence can give important insights and therefore – potentially – feedback in domains in which one has inaccurate self-views. These findings also raise the question of whether self- and other-perspectives give similar insights for each subdimension of intelligence or whether there are some abilities that might be judged more accurately by peers, parents, teachers, or even strangers, while others might be more validly estimated by the self. In our view, the state of research summarized within this chapter requires a combination of both perspectives, an enterprise that has rarely been undertaken so far and that possibly constitutes a promising future line of research.

Future Perspectives: A Combination of Self- and Other-Estimates

When people try to develop a successful career, their abilities (especially their intelligence) as well as their interests and personality traits certainly play an important role. But how about self- and other-perceptions of their abilities? The notion that self- and other-perceptions might provide distinct insights is far from new. Already, in 1955, Joseph Luft and Harrington Ingham introduced the Johari window, a model that places information known about a person (e.g., traits) into one of four quadrants. According to this model, such information can either be known (1) by oneself and others (located in the open area), (2) only by oneself (located in the hidden area), (3) only by others (located in the blind spot), or, possibly, even (4) neither by oneself nor by others (located in the unknown area). More recently, Simine Vazire (2010) introduced the self-other knowledge asymmetry model (SOKA), in which she built on the Johari window by proposing two characteristics of traits that influence the asymmetry in accuracy of self- and other- perceptions: observability and evaluativeness (i.e., the [un-]desirability of a trait). Highly evaluative traits should be associated with lower accuracy of self-perceptions, leading to peers having an advantage in judging us on such traits. Internal traits, that is, traits that are low in observability, on the other hand, should result in comparatively less accurate other-perceptions (for a summary of the Johari window and the SOKA model, see Figure 49.1).

From Vazire's research (2010), we get first impressions of the location of the Big Five within the Johari window. However, apart from general intelligence and creativity, which were estimated more accurately by friends than by individuals themselves, we do not know about the location of specific intelligence factors. What we know from the above-reported studies is that individuals can judge numerical ability validly but that accuracy might be somewhat lower for spatial ability and considerably lower for verbal ability. Steinmayr and Spinath's (2009) results showed

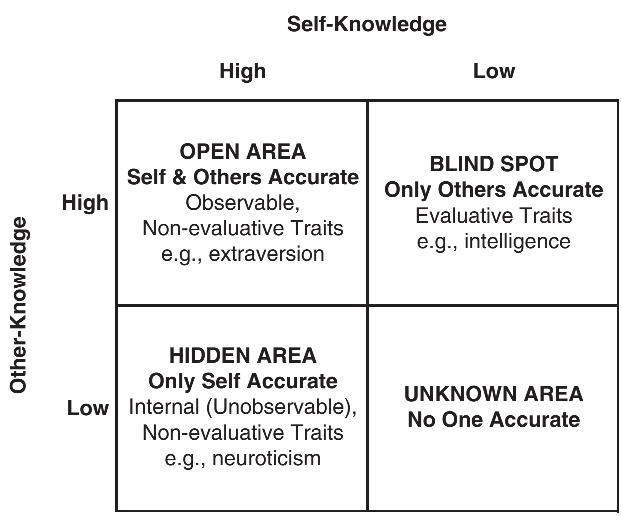

Figure 49.1 *Location of exemplary traits within the Johari window according to the self-other knowledge asymmetry (SOKA) model (after Bollich, Johannet, & Vazire, 2011; Luft & Ingham,1955).*

that others might be similarly good at judging numerical, spatial, and verbal abilities. However, a direct comparison of both perspectives might provide additional insight (see Neubauer et al., 2018). An additionally intriguing question is whether the dimensions of observability and evaluativeness that Vazire proposed in her SOKA model can also be applied to assign different abilities to the quadrants of the Johari window.

This multiperspective seems important to be addressed in future research, as individuals, when choosing a vocational career, be it apprenticeships, be it university majors, and so on, probably rely partially on their self-perception of their abilities, traits, and interests but might also use advice from peers, parents, teachers, or others. These other perspectives have rarely been addressed, let alone compared, something that should be changed in due course in order to arrive at a comprehensive view of the different perspectives and how they might inform vocational decisions.

Moreover, there are still some methodological problems regarding a psychometrically sound self-assessment of abilities that have not been completely solved. There is also the above-mentioned question of whether self-estimates (and possibly other-estimates) might enhance the quality of decisions (i.e., lead to better psychological counseling regarding educational/vocational choices). We think that following this line of research is a worthwhile future enterprise that could provide the field of intelligence research with interesting and important new insights.

References

Ackerman, P. L., & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *Journal of Experimental Psychology: Applied*, *13*, 57–78. https://doi.org/10.1037/1076-898X.13.2.57

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. Dunning, & J. I. Krueger (Eds.), *The self in social judgment* (pp. 85–106). New York: Psychology Press.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*, 9–30. https://doi.org/10.1111/1468–2389.00160

Beus, J. M., & Whitman, D. S. (2012). The relationship between typical and maximum performance: A meta-analytic examination. *Human Performance*, *25*, 355–376. https://doi.org/10.1080/08959285.2012.721831

Bollich, K. L., Johannet, P. M., & Vazire, S. (2011). In search of our true selves: Feedback as a path to self-knowledge. *Frontiers in Psychology*, *2*, 1–6. https://doi.org/10.3389/fpsyg.2011.00312

Borkenau, P., & Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology*, *65*, 546–553. https://doi.org/10.1037/0022–3514.65.3.546

Borkenau, P., & Liebler, A. (1995). Observable attributes as manifestations and cues of personality and intelligence. *Journal of Personality*, *63*, 1–25. https://doi.org/10.1111/1467–6494.ep9504261446

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, *86*, 599–614. https://doi.org/10.1037/0022–3514.86.4.599

Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology*, *90*, 60–77. https://doi.org/10.1037/0022–3514.90.1.60

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, *89*, 306–318. https://doi.org/10.1257/aer.89.1.306

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Chiappe, D., & MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *The Journal of General Psychology*, *132*, 5–40. https://doi.org/10.3200/GENP.132.1.5–40

Cogan, L., Conklin, A., & Hollingworth, H. (1915). An experimental study of self-analysis, estimates of associates, and the results of tests. *School and Society*, *2*, 171–179.

Cosmides, L., & Tooby, J. (2002). Unraveling the enigma of human intelligence: Evolutionary psychology and the multimodular mind. In R. J. Sternberg & J. C. Kaufman (Eds.), *The evolution of intelligence* (pp. 145–198). Mahwah, NJ: Lawrence Erlbaum Associates.

Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S. E., Liewald, D., . . . & Deary, I. J. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry*, *16*, 996–1005. https://doi.org/10.1038/mp.2011.85

Deary, I. J. (2012). Intelligence. *Annual Review of Psychology*, *63*, 453–482. https://doi.org/10.1146/annurev-psych-120710–100353

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest*, *11*, 53–79. https://doi.org/10.1177/1529100610387081

Denissen, J. J. A., Schönbrodt, F. D., van Zalk, M., Meeus, W. H. J., & van Aken, M. A. G. (2011). Antecedents and consequences of peer-rated intelligence. *European Journal of Personality*, *25*, 108–119. https://doi.org/10.1002/per.799

Dufner, M., Denissen, J. J. A., van Zalk, M., Matthes, B., Meeus, W. H. J., van Aken, M. A. G., & Sedikides, C. (2012). Positive intelligence illusions: On the relation between intellectual self-enhancement and psychological adjustment. *Journal of Personality*, *80*, 537–572. https://doi.org/10.1111/j.1467–6494.2011.00742.x

Dufner, M., Gebauer, J. E., Sedikides, C., & Denissen, J. J. A. (2018). Self-enhancement and psychological adjustment: A meta-analytic review. *Personality and Social Psychology Review*, *23*, 48–72. https://doi.org/10.1177/1088868318756467

Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychology Press.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In J. M. Olson & M. P. Zanna (Ed.) *Advances in experimental social psychology* (Vol. 44, pp. 247–296). San Diego, CA: Elsevier. https://doi.org/10.1016/B978-0-12-385522-0.00005-6

Dunning, D., & Helzer, E. G. (2014). Beyond the correlation coefficient in studies of self-assessment accuracy: Commentary on Zell & Krizan (2014). *Perspectives on Psychological Science*, *9*, 126–130. https://doi.org/10.1177/1745691614521244

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*, 1082–1090. https://doi.org/10.1037/0022–3514.57.6.1082

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273. https://doi.org/10.1037/0033-295X.95.2.256

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of Personality and Social Psychology*, *84*, 5–17. https://doi.org/10.1037//0022-3514.84.1.5

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, *105*, 98–121. https://doi.org/10.1016/j.obhdp.2007.05.002

Ehrlinger, J., Mitchum, A. L., & Dweck, C. S. (2016). Understanding overconfidence: Theories of intelligence, preferential attention, and distorted self-assessment. *Journal of Experimental Social Psychology*, *63*, 94–100. https://doi.org/10.1016/J.JESP.2015.11.001

Farwell, L., & Wohlwend-Lloyd, R. (1998). Narcissistic processes: Optimistic expectations, favorable self-evaluations, and self-enhancing attributions. *Journal of Personality*, *66*, 65–83. https://doi.org/10.1111/1467-6494.00003

Feld, J., Sauermann, J., & de Grip, A. (2017). Estimating the relationship between skill and overconfidence. *Journal of Behavioral and Experimental Economics*, *68*, 18–24. https://doi.org/10.1016/J.SOCEC.2017.03.002

Freudenthaler, H. H., & Neubauer, A. C. (2007). Measuring emotional management abilities: Further evidence of the importance to distinguish between typical and maximum performance. *Personality and Individual Differences*, *42*, 1561–1572. https://doi.org/10.1016/j.paid.2006.10.031

Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin*, *138*, 296–321. https://doi.org/10.1037/a0026556

Furnham, A., Moutafi, J., & Chamorro-Premuzic, T. (2005). Personality and intelligence: Gender, the big five, self-estimated and psychometric intelligence. *International Journal of Selection and Assessment*, *13*, 11–24. https://doi.org/10.1111/j.0965-075X.2005.00296.x

Furnham, A., Zhang, J., & Chamorro-Premuzic, T. (2005). The relationship between psychometric and self-estimated intelligence, creativity, personality and academic achievement. *Imagination, Cognition and Personality*, *25*, 119–145. https://doi.org/10.2190/530V-3M9U-7UQ8-FMBG

Gabriel, M. T., Critelli, J. W., & Ee, J. S. (1994). Narcissistic illusions in self-evaluations of intelligence and attractiveness. *Journal of Personality*, *62*, 143–155. https://doi.org/10.1111/j.1467–6494.1994.tb00798.x

Goncalo, J. A., Flynn, F. J., & Kim, S. H. (2010). Are two narcissists better than one? The link between narcissism, perceived creativity, and creative performance. *Personality and Social Psychology Bulletin*, *36*, 1484–1495. https://doi.org/10.1177/0146167210385109

Grijalva, E., & Zhang, L. (2016). Narcissism and self-insight: A review and meta-analysis of narcissists' self-enhancement tendencies. *Personality and Social Psychology Bulletin*, *42*, 3–24. https://doi.org/10.1177/0146167215611636

Haier, R. J. (2017). *The neuroscience of intelligence*. New York: Cambridge University Press.

Herreen, D., & Zajac, I. (2018). The reliability and validity of a self-report measure of cognitive abilities in older adults: More personality than cognitive function. *Journal of Intelligence*, *6*, 1–15. https://doi.org/10.3390/jintelligence6010001

Horrey, W. J., Lesch, M. F., Mitsopoulos-Rubens, E., & Lee, J. D. (2015). Calibration of skill and judgment in driving: Development of a conceptual framework and the implications for road safety. *Accident Analysis and Prevention*, *76*, 25–33. https://doi.org/10.1016/J.AAP.2014.12.017

Jacobs, K. E., Szer, D., & Roodenburg, J. (2012). The moderating effect of personality on the accuracy of self-estimates of intelligence. *Personality and Individual Differences*, *52*, 744–749. https://doi.org/10.1016/j.paid.2011.12.040

Koellinger, P., Minniti, M., & Schade, C. (2007). "I think I can, I think I can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology*, *28*, 502–527. https://doi.org/10.1016/j.joep.2006.11.002

Krajč, M., & Ortmann, A. (2008). Are the unskilled really that unaware? An alternative explanation. *Journal of Economic Psychology*, *29*, 724–738. https://doi.org/10.1016/j.joep.2007.12.006

Krizan, Z., & Suls, J. (2008). Losing sight of oneself in the above-average effect: When egocentrism, focalism, and group diffuseness collide. *Journal of Experimental Social Psychology*, *44*, 929–942. https://doi.org/10.1016/J.JESP.2008.01.006

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology*, *82*, 180–188. https://doi.org/10.1037/0022–3514.82.2.180

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*, 1121–1134. https://doi.org/10.1037/0022–3514.77.6.1121

Lee, A. J., Hibbs, C., Wright, M. J., Martin, N. G., Keller, M. C., & Zietsch, B. P. (2017). Assessing the accuracy of perceptions of intelligence based on heritable facial features. *Intelligence*, *64*, 1–8. https://doi.org/10.1016/j.intell.2017.06.002

Luft, J., & Ingham, H. (1955). The Johari window: A graphic model of interpersonal awareness. In *Proceedings of the Western Training Laboratory in Group Development*. Los Angeles: UCLA.

Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, *67*, 280–296. https://doi.org/10.1037/0021–9010.67.3.280

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*, 711–747. https://doi.org/10.1037/0033–2909.130.5.711

Miller, T. M., & Geraci, L. (2011). Unskilled but aware: Reinterpreting overconfidence in low-performing students. *Journal of Experimental Psychology: Learning Memory and Cognition*, *37*, 502–506. https://doi.org/10.1037/a0021802

Neubauer, A. C., Pribil, A., Wallner, A., & Hofer, G. (2018). The self–other knowledge asymmetry in cognitive intelligence, emotional intelligence, and creativity. *Heliyon*, 4, e01061. https://doi.org/10.1016/j.heliyon.2018.e01061

Neubauer, A. C., Wammerl, M., Benedek, M., Jauk, E., & Jaušovec, N. (2017). The influence of transcranial alternating current stimulation (tACS) on fluid intelligence: An fMRI study. *Personality and Individual Differences*, *118*, 50–55. https://doi.org/10.1016/j.paid.2017.04.016

Paulhus, D. L., & Harms, P. D. (2004). Measuring cognitive ability with the overclaiming technique. *Intelligence*, *32*, 297–314. https://doi.org/10.1016/j.intell.2004.02.001

Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology*, *84*, 890–904. https://doi.org/10.1037/0022-3514.84.4.890

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*, 322–338. https://doi.org/10.1037/a0014996

Reynolds, D. J., & Gifford, R. (2001). The sounds and sights of intelligence: A lens model channel analysis. *Personality and Social Psychology Bulletin*, *27*, 187–200. https://doi.org/10.1177/0146167201272005

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, *15*, 1–20. https://doi.org/10.1177/026553229801500101

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118–137. https://doi.org/10.1016/j.intell.2015.09.002

Sanchez, C., & Dunning, D. (2018). Overconfidence among beginners: Is a little learning a dangerous thing? *Journal of Personality and Social Psychology*, *114*, 10–28. https://doi.org/10.1037/pspa0000102

Schlösser, T., Dunning, D., Johnson, K. L., & Kruger, J. (2013). How unaware are the unskilled? Empirical tests of the "signal extraction" counterexplanation for the Dunning–Kruger effect in self-evaluation of performance. *Journal of Economic Psychology*, *39*, 85–100. https://doi.org/10.1016/J.JOEP.2013.07.004

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*, 162–173. https://doi.org/10.1037/0022-3514.86.1.162

Sedikides, C., & Alicke, M. D. (2012). Self-enhancement and self-protection motives. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 303–322). New York: Oxford University Press.

Sedikides, C., & Alicke, M. D. (in press). The five pillars of self-enhancement and self-protection. In R. M. Ryan (Ed.), *The Oxford handbook of motivation*. New York: Oxford University Press.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*, 407–441. https://doi.org/10.3102/00346543046003407

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*, *44*, 535–551. https://doi.org/10.1016/j.paid.2007.09.015

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628–654. https://doi.org/10.1037/a0027473

Soh, L., & Jacobs, K. E. (2013). The biasing effect of personality on self-estimates of cognitive abilities in males and females. *Personality and Individual Differences*, *55*, 141–146. https://doi.org/10.1016/j.paid.2013.02.013

Sommer, U., Fink, A., & Neubauer, A. C. (2008). Detection of high ability children by teachers and parents: Psychometric quality of new rating checklists for the assessment of intellectual, creative and social ability. *Psychology Science Quarterly*, *50*, 189–205.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613–629. https://doi.org/10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Steinmayr, R., & Spinath, B. (2009). What explains boys' stronger confidence in their intelligence? *Sex Roles*, *61*, 736–749. https://doi.org/10.1007/s11199-009-9675-8

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, *49*, 607–627. https://doi.org/10.1037/0022-3514.49.3.607

Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, *41*, 37–55. https://doi.org/10.1037/0022-3514.41.1.37

Syzmanowicz, A., & Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta analyses. *Learning and Individual Differences*, *21*, 493–504. https://doi.org/10.1016/j.lindif.2011.07.001

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*, 281–300. https://doi.org/10.1037/a0017908

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2008). What makes you think you're so smart? Measured abilities, personality, and sex differences in relation to self-estimates of multiple intelligences. *Journal of Individual Differences*, *29*, 35–44. https://doi.org/10.1027/1614-0001.29.1.35

Williams, P. G., Rau, H. K., Suchy, Y., Thorgusen, S. R., & Smith, T. W. (2017). On the validity of self-report assessment of cognitive abilities: Attentional control scale associations with cognitive performance, emotional adjustment, and personality. *Psychological Assessment*, *29*, 519–530. https://doi.org/10.1037/pas0000361

Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*, *28*, 238–249. https://doi.org/10.1177/0146167202282009

Zell, E., & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science*, *9*, 111–125. https://doi.org/10.1177/1745691613518075

Zippert, E. L., & Ramani, G. B. (2017). Parents' estimations of preschoolers' number skills relate to at-home number-related activity engagement. *Infant and Child Development*, *26*, e1968. https://doi.org/10.1002/icd.1968

PART X

Conclusion

50 Speculations on the Future of Intelligence Research

Robert J. Sternberg

In this final chapter, I briefly summarize some of the major challenges for intelligence research going forward. These issues are among those that may be highlighted when a next edition of the *Cambridge Handbook of Intelligence* is published, perhaps in a decade or so. Obviously, different readers will have different impressions of what the major challenges are. Thus, the challenges presented here represent just one scholar's view of what lies in store for the intelligence researchers of the future. Here are fifteen questions I think researchers will want to address.

Are Hierarchical Models an Endstate?

Hierarchical models of intelligence, such as that of Carroll (1993), are very much in favor today among many if not most intelligence researchers (see Chapter 1 by Sternberg; Chapter 4 by DeBoeck, Gore, Gonzalez, & San Martin; Chapter 5 by Walrath, Willis, Dumont, & Kaufman). In such models, general intelligence, or *g*, usually appears at the top, with successively more specific abilities located at lower levels of the hierarchy. Such models have been around for a while (Burt, 1940; Vernon, 1961). Some psychometric theorists have proposed other structures of human abilities, such as a radex (Guttman, 1965); and other types of theorists have proposed different kinds of models, such as the theory of multiple intelligences and the theory of successful intelligence (Gardner, 2011; Sternberg, 2003; see also Chapter 27 by Kornhaber and Chapter 28 by Sternberg).

It may be that a variation of a type of model first proposed rather early in the twentieth century will be around forever. But, in science, relatively few findings last forever. Rather, findings serve as heuristic devices to encourage the next generation of findings. The *g* construct (Spearman, 1904) may prove to be the exception. What other ideas in psychology from the turn of the twentieth century are still so prominent? Or it may be that researchers will decide, as they often do, that theories based on *g* and a hierarchy of abilities below it are a start but not an endstate for understanding intelligence? It is worth recalling that one scholar believed that the "end of history" had arrived, only to discover that what he thought was the end was nothing more than an inflection point – a turning point from one phase of history to the next (Fukuyama, 2006). At the very least, one would hope that researchers would acquire a better understanding of just what *g* is (and of what the abilities lying under it in the hierarchy are). That certainly is an exciting prospect for the future. And discovering that hierarchical models are only the beginning, not the end, might possibly be even more exciting.

What Is the Role of the Brain in Intelligence and of Intelligence-Enhancing Drugs on the Brain?

One of the hottest areas of research today in the field of intelligence is that of linking genetic, psychometric, and cognitive models of intelligence to brain functioning. How do the elements of these models link to the processing of information in the brain (see Chapter 6 by Grigorenko & Burenkova, Chapter 17 by Zentall, Chapter 18 by Bates & Byrne; Chapter 19 by Haier)? When one compares contemporary models of brain functioning, such as P-FIT (Jung & Haier, 2007), to earlier models (e.g., Hebb, 1949), it is amazing how much progress has been made in understanding biological bases of intelligence.

But important questions remain. Of course, we will need more understanding of the distribution of intelligence in the brain. Gardner's (2011) theory suggests that aspects of intelligence (or, as Gardner puts it, multiple intelligences) are modularly distributed in the brain. Jung and Haier's (2007) theory suggests that intelligence is more widely distributed throughout the brain. We also need a better understanding of the role of the brain in intelligence. Whereas Haier and some other biological researchers believe intelligence is entirely biological (e.g., Haier, 2017), other researchers believe intelligence resides only partly in biology (see Sternberg, 1990). A still further question might be that if we understood the brain-based origins of every cognitive process of intelligence, would that constitute fully understanding intelligence – would it tell us, for example, who will be able to adapt well to their environment and who would not; or who knows when to push people's buttons to get what they want from them and who knows when not to?

A further question that will need to be addressed is that of the current and future role of intelligence-enhancing drugs. Are there really drugs that enhance intelligence (starting with something as simple as caffeine) or are such enhancements merely state rather than trait variables (see Haier, 2018; Sternberg, 2014)? That is, perhaps such drugs temporarily increase intellectual performance but they do not actually increase intelligence, at least over the long term.

Does Culture Affect What Intelligence Is or Just What It Is Conceived of as Being?

Some investigators believe intelligence is at least partially culturally based (e.g., Laboratory of Comparative Human Cognition, 1982; Serpell, 2000; Sternberg, 2004; see Chapter 34 by Ang, Ng, & Rockstuhl). Is there really any cultural basis to intelligence or is intelligence and its underlying biology the same across cultures, merely manifesting itself in different ways in different cultural settings? The question is an important one. Cultural researchers would argue that the skills needed to adapt to the environment in different cultures are simply different. For example, the skills necessary for hunting or gathering might be crucial to survival in one culture but of little or no importance in another. The skills needed for reading might be important in one culture but of little or no use in a culture in which reading is not done. How can intelligence be maintained in later adulthood?

There is good evidence that some aspects of intelligence, especially fluid intelligence, tend to decline with age. One may be intelligent in childhood and even early adulthood but then become susceptible to decline once one reaches the age of sixty and above (see Chapter 8 by Gelman & DeJesus; Chapter 9 by Hertzog). Knowledge may be intact but the functioning of basic processes may be far from what it once was (see Chapter 20 by Nettelbeck, Zwalf, & Stough). There have been many efforts, discussed in Chapter 9, to maintain or even improve fluid intelligence in older age but they have met with, at best, mixed success. Nevertheless, the population of many countries in the world is aging and so there is a growing and, in some cases, desperate market for any kind of intervention – behavioral, pharmaceutical, or otherwise – to maintain or enhance intelligence. Such interventions are particularly crucial in the cases of brain disease, such as Alzheimer's or other forms of dementia.

The point is that cultures vary over time and place; and intelligence researchers may hope better to understand what makes a person smart in one culture does not necessarily cross over well to a different culture. I suspect that research in this area will grow exponentially, as societies come fully to realize the costs of having large aging populations with decreasing levels of independent adaptive skills.

What Are the Genetic Bases of Intelligence?

The genetics of intelligence is another area in which there is now an incredible amount of research (see Chapter 6 by Grigorenko & Burenkova). Not so long ago, the focus of research was on finding one or a few intelligence genes. As Chapter 6 by Grigorenko and Burenkova on genetics points out, those attempts have been, for the most part, abandoned. It is now recognized that, whatever the genes may be for intelligence, they are widely distributed across the genome. Taking this genome-wide approach has greatly increased the percentage of variance that can be accounted for by genetic factors.

One of the greatest challenges for genetic research is what, exactly, to do with the information. Back in 1997, a movie, *Gattaca*, provided a look into a dystopian society where a person's future is determined by analyses of the person's genes. The basic theme of the movie is that, the heavy use of genetic information by the futuristic society notwithstanding, genes were a highly imperfect predictor of a person's achievements and even motivation to achieve. Some unmotivated people were advanced to the metaphorical head of the line, whereas other highly motivated ones were placed at the back of the line. Truly, how could a society ever account for *all* of the diverse factors that lead to success in everyday life? Already, people fear that genetic information may be used against them and, in some cases, most notably in the capture of alleged criminals, it already is. But where does the use of this information stop? Will the issuance or cost of health insurance be next? And where does the use of such information cross the line between capturing criminals and capturing people who are thought to be possible future dangers to society? These are some of the questions that this area of research will face and will have to address, because there is no stopping this train now that it is out of the station.

What Are the Causes of the Flynn Effect and the Reverse Flynn Effect?

The Flynn effect (see Chapter 39 by Flynn) refers to secular increases in IQ over the course of the twentieth century. These increases were worldwide and large, amounting to as much as thirty points overall. Oddly, perhaps, although the Flynn effect was first proposed more than thirty years ago (Flynn, 1984), we still do not fully understand what causes it, although there have been many explanations (e.g., Dickens & Flynn, 2001; Neisser, 1998). One reason it is so important to understand the Flynn effect is that, in some places, it has started to reverse (see, e.g., Dutton, van der Linden, & Lynn, 2016). Thus, societies may not have purposely created increases in IQ over time but they perhaps could purposely try to stem decreases over time. If we well understood the bases of the Flynn effect, we might be able to answer the next question in a fully meaningful way.

Can Intelligence of Individuals Be Meaningfully Increased by Environmental Interventions?

There have been many attempts to increase intelligence and some of them have had at least modest if often temporary success (see Chapter 10 by Nickerson; Chapter 43 by Mayer). The main attempts have been cognitive training programs. When these programs succeed, they generally have small short-term effects. The effects do not last any more than effects last for diets. One can lose weight on almost any diet – the problem is keeping the weight off. Similarly, the problem is keeping the gains once one returns to one's normal environment, which, in many cases, may not be conducive to intellectual gains.

The most effective intervention for increasing intelligence is and probably always has been effective quality schooling (Ceci, 1996; see also Chapter 40 by Barnett, Rindermann, Williams, & Ceci; Chapter 43 by Mayer). But older people often do not have a lot of time or see a need for schooling. And younger people may not be able to continue with schooling if they have to work or if their schooling experience has been distinctly nonideal. So there is a need for interventions that are effective over the long term. Some of these may end up being pharmaceutical (see Chapter 19 by Haier) but such interventions are largely untested. The chemical interventions that raise intellectual performance, such as caffeine, do so only temporarily. And other kinds of stimulants, such as Ritalin, may have side effects and long-term consequences that make them unsuitable for long-term use. Perhaps no other question more needs to be addressed regarding intelligence but perhaps no other leaves researchers with such a lacuna of information regarding what works.

Is the Nature of Intelligence Evolving, or Just Its Manifestations?

Greenfield (Chapter 38) has suggested that the nature of intelligence is evolving in a way that prioritizes the needs of the society over the needs of the small community or even neighborhood. Thus, she has suggested that what is intelligent in a culture may change over time. According to Greenfield, the skills needed for adaptive success in more modern cultures are qualitatively different from those needed in less modern cultures. But will the skills that are important in today's modern culture continue to be important in the cultures of the future? For example, some intelligence tests (such as the Thurstone Primary Mental Abilities tests) once featured (in the mid-1900s) arithmetic computation as a device for measuring number ability. Would any test of intelligence include such items today? Not likely. And what will be needed to be smart in the world of tomorrow? Are we teaching children to be smart in ways that will not apply in the future? Already, some might worry that the skills older adults learned as children are not the skills children need for adaptation in a digital world. Will those same skills be the ones for the world of tomorrow or will the adults of the future have the same worries some of us have?

Is High Intelligence Having Negative as well as Positive Effects on the World, and What Is to Be Done About It?

Some investigators have worried that high intelligence is not all that it is cracked up to be – indeed, that it may have almost as many negative effects on society as it has positive ones (Sternberg, 2018c, 2018d). People can and often do use their intelligence selfishly – to benefit themselves – rather than wisely – to benefit others as well as themselves. The result is that world problems – poverty, income disparity, climate change, wars and other forms of violence, pollution, and hunger – persist (see Chapter 28 by Sternberg). A serious problem, therefore, is how we can use intelligence in order to achieve better outcomes in the world. Several scholars have speculated on this topic (Sternberg, 2018a) but it is by no means clear yet how we can harness intelligence to be used in a positive rather than neutral or negative way. The world has many serious problems, among which is the problem of nuclear weapons that could obliterate much if not all of existing humanity. We need to figure out how the increases in IQ charted by the Flynn effect can be used for a greater good rather than to cause harm and possibly undermine civilization as we know it. These words may sound apocalyptic but that makes them no less realistic with respect to the problems the world faces today.

How Should Gifted People Be Identified?

Intellectually gifted children are an important part of the future of our country and our world. If they are not going to be the ones who help build a future, who will be? But today, as in the past, the problem of how to figure out exactly who they are is a challenge (see Chapter 13 by Reis & Renzulli; also see Chapter 12 by Feldman & Morelock). Historically, such children were identified largely by IQ tests. Terman (1925–1959), one of the earliest scholars of gifted children, used an IQ test (the Stanford-Binet) to identify gifted children whose lives he would follow over the course of the years. Today, as Reis and Renzulli point out, there is a wide variety of assessments available for identifying gifted children. It is pretty clear, however, that such assessments, to be maximally effective, will need to look at more than intelligence or, at least, more than intelligence narrowly defined. Reis and Renzulli emphasize creativity and task commitment as well as above-average intellectual abilities. Sternberg (see Chapter 28) recommends assessments of creativity, practical intelligence, and wisdom in addition to assessments of analytical abilities (as measured by tests of IQ). He also has suggested that schools focus on identifying children who are likely to be the concerned and active citizens and ethical leaders of the future (Sternberg, 2017). Gardner (2011; see also Chapter 27 by Kornhaber) has suggested assessments of his various intelligences: linguistic, logical-mathematical, bodily-kinesthetic, interpersonal, intrapersonal, musical, naturalist, and spatial. Whatever system is used, it almost certainly needs to go beyond IQ (Sternberg, 1985). Exactly how will be a matter for the future to determine. Perhaps measures will include grit (Duckworth, 2016) or an incremental view of intelligence (Dweck, 2007). Whatever future measures look like, I would hope that they would be broader than the measures of today and also more open to recognizing the wide range of intellectual skills brought to the world by members of different societies and cultures – and not just Western ones.

Current ways of identifying gifted children and adults may be less than fair to members of diverse groups (see Chapter 11 by Fidler, Schworer, Swanson, & Hepburn; Chapter 14 by Halpern & Wai; Chapter 15 by Suzuki, Larson-Konar, Short, & Lee; Chapter 37 by Niu). Can any society afford to waste human resources – children who are gifted who may be passed over because they are not gifted in the narrow ways in which many schools identify gifted children? Given the problems the world is facing today, I suspect not.

Will Artificial Intelligence Ultimately Be Dangerous to the World?

For most of us, artificial intelligence (AI) is a boon. It not only can speak to us from our cell phone through various digital assistants; it also can help us navigate how to get from one place to another when we have no clue as to where we are or it even can be used to drive a car or a truck. What is not to like? But some experts on AI are worried (see Chapter 25 by Goel). Is there a moment of “singularity” at which progress on AI essentially explodes and perhaps takes off on its own? Will AI become like the computer HAL gone berserk in the novel *2001: A Space Odyssey*? Will humans lose control of their own inventions?

To some, such a possibility might seem remote. But singularity events of any kind are only slightly, if at all, predictable. And keep in mind that supercomputers already are beating humans in domains where it was never thought to be possible, such as the games of chess and even of Go (which is quite a bit more complicated than chess). There were not lots of people predicting the stock market crashes of 1929 or of 2008. And, whereas one can recover from a stock market crash, it is not clear how one would recover from a situation in which AI "decides" not to work for humans but rather to have humans work for it. At the very least, controls need to be built in. Isaac Asimov recognized this danger in his robot series (e.g., Asimov, 2004). In the Asimov series, robots were built with "positronic" brains that ensured they would serve humans rather than vice versa. But contemporary AI has no such safeguards. Maybe it is time that it should. But then, how would the safeguards be created and, if robots became smarter than humans, how could humans guarantee that the robots would not outsmart the safeguards?

What Effects Are Social Media and the Internet More Generally Having on Both Individual and Collective Intelligence?

Of course, the world changes with time. In most respects, it probably is better than it was during the large majority of times during human history (Pinker, 2018). But, in at least one way, things do not seem to be improving (Carr, 2011; Twenge, 2014). This way is with regard to the effects of the Internet and social media in particular on our thought processes. Interacting with media of all kinds can change the way we think, for better or worse (see Chapter 26 by Quiroga & Colom). One would think that, before major potential interventions with human thought, societies might want to do some research. After all, when new drugs are introduced, many societies require very laborious and expensive trials to ensure their safety with regard to our physical health. Why would a society put so many constraints on interventions that affect physical health and well-being but no constraints at all on interventions that affect mental health and well-being? But there are no such constraints and video games, social media, and easy Internet access to an astonishing array of resources seem to be embedded integrally in many if not most societies today. Quiroga and Colom, in Chapter 26, cite research suggesting that students who use social media every day score as much as 4 percent lower than average with respect to reading, mathematics, and science performance. Over large numbers of students, that is an astonishing drop. Moreover, the Internet may impair our critical thinking (many people just believe what they read), our access to a broad range of information (people more and more read only sources that agree with their preconceptions), our depth in reading (people read more superficially and have trouble finishing what they read), and even our kindness to each other (many people lose whatever courtesy they have when on the Internet) (Bauerlein, 2008, Manjoo, 2013; Maurer, 2015). Violent video games may make us more violent (Anderson & Bushman, 2001).

The Internet, social media, and video games (at least, the violent ones) are changing, at minimum, the way we use our intelligence and, possibly, even our intelligence itself. If people are having more trouble concentrating for long periods of time and critically appraising what they do manage to concentrate on that would seem to be a cause for concern. Researchers on intelligence need to study these and related phenomena and they need to provide society with results from a massive social-engineering experiment with no institutional review board (IRB) approval and no obvious control group.

Why Are People Who Are So Intelligent at Times So Lacking in Rational and Wise Thinking?

Related to the issues discussed in this chapter, research suggests that the relation of intelligence, at least as measured by IQ, and rational thinking is surprisingly modest (see Chapter 23 by Hambrick, Burgoyne, & Altman; Chapter 24 by Gigerenzer; Chapter 46 by Stanovich, Toplak, & West). Although reasoning is related to intelligence, of course, only certain kinds of reasoning show strong relations, induction, for example, more than deduction (see Chapter 21 by Conway & Kovacs; Chapter 22 by Lakin & Kell). Intelligence also appears to be necessary but far from sufficient for wise thinking (see Chapter 47 by Glück). These findings suggest that, for all of the observed increases in IQ during the twentieth century, people's processing of information may not be a lot better and may even be worse than it once was. Stanovich has referred to the lack of rational thinking, even in intelligent people, as "dysrationalia" (Stanovich, 1993, 1994). Dysrationalia may have increased in recent years as a result of extensive and perhaps sometimes excessive use of the Internet, as discussed in the previous section. Anyone who teaches (including myself) observes in student papers a tendency to accept anything they read on the Internet as true. And politicians know that merely repeating the same lie, again and again, begins to give the lie credibility. People remember what they hear but forget the source (Johnson, 2016).

We live in an age in which post–World War II hopes of the spread of liberal democracy are fading fast (Levitsky & Ziblatt, 2018; McCarty, Poole, & Rosenthal, 2008; Mounck, 2018). Dictators and would-be dictators are showing the world just how very little increases in IQ mean when it comes to people falling for the same sucker gambits that have worked for thousands of year: Others are out to get us; we need to band together to expel or even obliterate these nefarious others; our lives and livelihoods depend on it (see Sternberg, 2018b; Sternberg & Sternberg, 2008). It is hard to believe that, in 2018, people accept these fabrications much as they did in ancient times. Their own overestimation of their intelligence may lead them to believe they are immune to foolishness, when in fact they are not (see Chapter 49 by Neubauer & Hofer). So much for the usefulness of IQ when people's emotions are deliberately manipulated by despots and despots-in-waiting. And this leads us to our next issue.

Are There Noncognitive (e.g., Emotional, Attitudinal, Motivational) Aspects of Intelligence, and, If So, What Are They?

In the real world, people's thinking is constantly affected by powerful and sometimes almost overwhelming emotions and motivations (see Chapter 29 by Rivers, Handley-Miner, Mayer, & Caruso; see Chapter 44 by Carr & Dweck). All the IQ points in the world probably have little to do with how people argue about issues such as immigration, disarmament, abortion, gay marriage, taxation, and poverty. People just have too much of a personal stake and their intuitive (Type I) thinking overwhelms their rational and reflective thinking (Kahneman, 2013).

Some might argue that intelligence is not about emotion or motivation; it is about *g* and the abilities that lie under it in a hierarchical model. But, at least in my view, this is a problem not with intelligence but rather with our way of understanding and studying it. What good is a high IQ if a person is unable or unwilling to use it effectively in solving life problems? Certainly intelligence is not merely about adaptation to a highly structured test given in a sterile setting with a strict time limit; it must be also about adaptation as it occurs in the real world. We need better to understand how emotions and motivations modulate our use of intelligence when we most need it, not just when intelligence is being measured in a highly controlled and largely unrealistic environment.

What Kinds of Environmental Factors Inadvertently Decrease Intelligence?

Societies around the world have not been careful as to the environmental toxins they have allowed into the world that can adversely affect intelligence (see Chapter 41 by Bellinger). Lead, which is a toxin with clearly adverse effects on intelligence, is found far and wide and once was used in gasoline, house paint, water pipes, and jewelry. In some places, it still is. Cadmium is another brain toxin found widely distributed across our world. And who knows what the effects of untested food additives, chemicals in cosmetics, and chemicals used in agriculture are? We need better to understand whether some of the chemicals we are using are impairing not only our physical well-being but our intellectual well-being as well.

Do Societies Today Place Too Much (or Too Little) Emphasis on Intelligence?

There can be no doubt that intelligence is one of the best predictors of many and diverse kinds of behaviors in everyday life. Intelligence as measured by tests of IQ predicts academic performance, job placement, job success, marital success, health, longevity, and many other things (Sackett, Shewach, & Dahlke, in press; see Chapter 40 by Barnett, Rindermann, Williams, & Ceci). But, as noted, other factors are certainly important too, such as conscientiousness (Sackett et al., in press), creativity (see Chapter 45 by Plucker, Karwowski, & Kaufman), motivation (see Chapter 44 by Carr & Dweck), wisdom (see Chapter 47 by Glück), judgment in interpersonal relationships (see Chapter 35 by Geher, Kaufman, Planke, & Di Santo), leadership (see Chapter 33 by Boyatzis), common sense (see Chapter 30 by Hedlund; Chapter 36 by Sujan & Sujan), the ability to get along with and interact effectively with others (see Chapter 31 by Kihlstrom & Cantor; Chapter 32 by Malone & Woolley), personality and especially openness to experience (see Chapter 42 by DeYoung), expertise (see Chapter 48 by Ackerman), and the ability to understand people from diverse cultures (see Chapter 34 by Ang, Ng, & Rockstuhl), among many other things. Skills in some of these areas can be viewed as constituting special kinds of intelligence or they can be seen as useful adjuncts to intelligence. Either way, a narrow view of intelligence will always be lacking in terms of powerfully predicting in any kind of comprehensive and meaningful way broad aspects of success in life, at least as defined by adaptive behavior. And, in the end, if intelligence does not fully account for adaptive behavior, then what, exactly, do we mean by intelligence and how can we expand our conceptualizations so that adaptive behavior and its origins are more comprehensively understood?

References

Anderson, C., A., & Bushman, B. J. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature. *Psychological Science*, https://doi.org/10.1111/1467-9280.00366

Asimov, I. (2004). *I, robot*. New York: Random House.

Bauerlein, M. (2008). *The dumbest generation*. New York: Tarcher.

Burt, C. (1940). *The factors of the mind*. London: University of London Press.

Carr, N. (2011). *The shallows: What the Internet is doing to our brains*. New York: W. W. Norton.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Ceci, S. J. (1996). *On intelligence*. Cambridge, MA: Harvard University Press.

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, *108*, 346–369.

Duckworth, A. (2016). *Grit: The power of passion and perseverance*. New York: Scribner.

Dutton, E., Van der Linden, D., & Lynn, R. (2016). The negative Flynn effect: A systematic literature review. *Intelligence*, *59*, 163–169.

Dweck, C. S. (2007). *Mindset: The new psychology of success*. New York: Ballantine.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29–51.

Fukuyama, F. (2006). *The end of history and the last man*. New York: Free Press.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Guttman, L. (1965). A faceted definition of intelligence. In R. R. Eiferman (Ed.), *Scripta Hierosolymitana* (Vol. 14, pp. 166–181). Jerusalem: Magnes Press.

Haier, R. J. (2017). *The neuroscience of intelligence*. New York: Cambridge University Press.

Haier, R. J. (2018). A view from the brain. In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 167–182). New York: Cambridge University Press.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Johnson, M. K. (2016). Falling down the duck/rabbit hole. In R. J. Sternberg, S. T. Fiske, & D. J. Foss (Eds.), *Scientists making a difference* (pp. 132–135). New York: Cambridge University Press.

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(2),135–154. https://doi.org/10.1017/S0140525X07001185

Laboratory of Comparative Human Cognition. (1982). Culture and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 642–719). New York: Cambridge University Press.

Levitsky, S., & Ziblatt, D. (2018). *How democracies die*. New York: Crown.

McCarty, N., Poole, K. T., & Rosenthal, H. (2008). *Polarized America: The dance of ideology and unequal riches*. Cambridge, MA: MIT Press.

Manjoo, F. (2013). You won't finish this article: Why people online don't read to the end. Slate, June 6. www.slate.com/articles/technology/technology/2013/06/how_people_read_online_why_you_won_t_finish_this_article.html

Maurer, H. (2015). Does the Internet make us stupid? *Communications of the ACM*, *58*(1), 48–51. https://doi.org/10.1145/2629544

Mounck, Y. (2018). *The people vs. democracy: Why our freedom is in danger and how to save it*. Cambridge, MA: Harvard University Press.

Neisser, U. (Ed.) (1998). *The rising curve: Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.

Pinker, S. (2018). *Enlightenment now: The case for reason, science, humanism, and progress*. New York: Viking.

Sackett, P. R., Shewach, O. R., & Dahlke, J. A. (in press). The predictive value of general intelligence. In R. J. Sternberg (Ed.), *Human intelligence: An introduction*. New York: Cambridge University Press.

Serpell, R. (2000). Intelligence and culture. In R. J. Sternberg (Ed.), *Handbook of intelligence* (2nd ed., pp. 549–577). New York: Cambridge University Press.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Stanovich, K. E. (1993). Dysrationalia: A new specific learning disability. *Journal of Learning Disabilities*, *26*(8), 501–515.

Stanovich, K. E. (1994). An exchange: Reconceptualizing intelligence: Dysrationalia as an intuition pump. *Educational Researcher*, *23*(4),11–22.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. J. (1990). *Metaphors of mind*. New York: Cambridge University Press.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. New York: Cambridge University Press.

Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist*, *59*(5), 325–338.

Sternberg, R. J. (2014). Intelligence as trait–and state? *Journal of Intelligence*, *2*, 4–5.

Sternberg, R. J. (2017). ACCEL: A new model for identifying the gifted. *Roeper Review*, *39* (3), 139–152. www.tandfonline.com/eprint/kSvRMFf9R8tAJPDRfXrJ/full

Sternberg, R. J. (Ed.) (2018a). *The emerging role of intelligence in the world of the future* (e-book). Basel, Switzerland: MDPI.

Sternberg, R. J. (2018b). FLOTSAM: A new model for the development and transmission of hate. *Journal of Theoretical Social Psychology*. https://doi.org/10.1002/jts5.25

Sternberg, R. J. (2018c). Speculations on the role of successful intelligence in solving contemporary world problems. *Journal of Intelligence*, *6*(1), 4. https://doi.org/10.3390/jintelligence6010004

Sternberg, R. J. (2018d). Why real-world problems go unresolved and what we can do about it: Inferences from a limited-resource model of successful intelligence. *Journal of Intelligence*, *6*(3), 44. https://doi.org/10.3390/jintelligence6030044

Sternberg, R. J., & Sternberg, K. (2008). *The nature of hate*. New York: Cambridge University Press.

Terman, L. M. (1925–1959). *Genetic studies of genius*, 5 vols. Stanford: Stanford University Press.

Twenge, J. (2014). *Generation me: Why today's young Americans are confident, assertive, entitled – and more miserable than ever before*. New York: Atria Books.

Vernon, P. E. (1961). *The structure of human abilities* (2nd ed.). London: Methuen.

Author Index

Entries underlined indicate author of chapter in this book, or that the author is the main subject on these pages

Subject Index

Subject Index (For researchers in the field *see* Author Index)
Locators in **bold** refer to tables; those in *italic* to figures