THE NATURE OF HUMAN INTELLIGENCE

Edited by

Robert J. Sternberg

THE NATURE OF HUMAN INTELLIGENCE

EDITED BY

ROBERT J. STERNBERG

Cornell University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

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

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/9781316629642
DOI: 10.1017/9781316817049

First published 2018

Printed in the United States of America by Sheridan Books, Inc.

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

Library of Congress Cataloging-in-Publication Data
Names: Sternberg, Robert J., editor.
Title: The nature of human intelligence / edited by Robert J. Sternberg.
Description: New York: Cambridge University Press, 2018. |
Includes bibliographical references and index.
Identifiers: LCCN 2017044205| ISBN 9781107176577 (hardback) |
ISBN 9781316629642 (paperback)
Subjects: LCSH: Intellect.
Classification: LCC BF431.N384 2018 | DDC 153.9–dc23
LC record available at https://lccn.loc.gov/2017044205

ISBN 978-1-107-17657-7 Hardback
ISBN 978-1-316-62964-2 Paperback

Contributors

ACKERMAN, PHILLIP L. Georgia Institute of Technology

BOUCHARD, THOMAS J., JR. University of Minnesota

BUTLER, HEATHER A. California State University, Dominguez Hills

CECI, STEPHEN J. Cornell University

CHEN, JIE-QI. Erikson Institute

CONWAY, ANDREW R. A. Claremont Graduate University

DEARY, IAN J. University of Edinburgh

ENGLE, RANDALL W. Georgia Institute of Technology

ERICSSON, K. ANDERS. Florida State University

FLYNN, JAMES R. University of Otago

GARDNER, HOWARD. Harvard University

GINTHER, DONNA K. University of Kansas

GOTTFREDSON, LINDA S. University of Delaware

GRIGORENKO, ELENA L. Baylor College of Medicine, St. Petersburg University, Russia, University of Houston, Yale University

HAIER, RICHARD J. University of California, Irvine

HALPERN, DIANE F. Claremont-McKenna College and Minerva Schools at KGI

KAHN, SHULAMIT Boston University

KAUFMAN, ALAN S. Yale University

KAUFMAN, SCOTT BARRY. University of Pennsylvania

KORNHABER, MINDY. Pennsylvania State University

KOVACS, KRISTOF. Eszterhazy Karoly University

LUBINSKI, DAVID. Vanderbilt University

LYNN, RICHARD. University of Ulster

MAYER, JOHN D. University of New Hampshire

RITCHIE, STUART J. University of Edinburgh

SHIPSTEAD, ZACH. Alma College

STERNBERG, ROBERT J. Cornell University

WILLIAMS, WENDY M. Cornell University

Preface

"The field of intelligence is dead." So said one of my graduate-school mentors, Lee Cronbach, himself an expert on intelligence, in 1972. I had just started as a graduate student at Stanford and had gone to see him about possibly doing some work with him in the field of intelligence; but he displayed no optimism about the field. Perhaps I should not have been surprised. About a decade earlier, one of his colleagues at Stanford, Quinn McNemar, had written a paper published under the title "Lost: Our Intelligence? Why?" (McNemar, 1964).

The collaboration with Cronbach never happened, and it was not until my second year as a graduate student that I started working in the field of intelligence under the mentorship of my primary adviser, Gordon Bower. But the year that I started working on intelligence, unbeknownst to me, the field that McNemar suggested was lost was found again – or, to put it in terms of Cronbach's metaphor, it became undead and resurrected. Earl Hunt, to whom this volume is dedicated, and two of his colleagues had just published a book chapter that, in some respects, would bring intelligence to life (Hunt, Lunneborg, & Lewis, 1975). Hunt and his colleagues showed that a productive path to understanding intelligence would be through the cognitive analysis of intellectual functioning. Hunt and colleagues followed up two years later with a cognitive analysis specifically of verbal ability (Hunt, Lunneborg, & Lewis, 1975). Two years after that, I proposed a related although in some respects competing approach to studying intelligence (Sternberg, 1977). The rest, as they say, is history. Today, the field of intelligence research is about as active as any field could be. Indeed, its form seems to change every few years, or, arguably, every few months!

Once upon a time, recognizing that the field of intelligence was thriving, I edited a series that updated advances in the field on a regular basis. The series started in 1982 and was called *Advances in the Psychology of Human Intelligence* (Sternberg, 1982a). But that series lasted only through five volumes. A few years after my first edited volume, Douglas Detterman (1985) started a related series, *Current Topics in Human Intelligence.* But that series too is long gone. The field continued to be updated through a series of handbooks edited by myself (e.g., Sternberg, 1982b, 2000; Sternberg & Kaufman, 2011) and others (e.g., Goldstein, Princiotta, & Naglieri, 2015; Wolman, 1985), but these handbooks were intended to be comprehensive reviews rather than updates regarding current research on particular topics. Yet, the field continued to advance rapidly.

So I recently decided to edit a volume of updates on intelligence research. In the past, I had just chosen colleagues to write whose work I admired because of its impact on the field. But at the same time, I realized that my selections were always colored by my own biases about what kinds of research were worthwhile to the field. Those biases led to some kinds of work being included, but not others. This time I wanted to do things a bit differently.

When I started this volume, I recently had coedited a volume of essays by eminent psychologists who were chosen in an objective (statistically based) way (Sternberg, Fiske, & Foss, 2016), and I decided to try an analog to this approach for the current volume. I started with what I considered to be the three principal contemporary textbooks on intelligence – ones by Hunt (2011), Mackintosh (2011), and Sternberg and Kaufman (2011) – and tabulated citations in these volumes to the various authors whose work was mentioned. I then chose as my potential authors the scholars whose work was most frequently cited. Almost everyone I wrote to then agreed to write. Earl Hunt was an exception, and I later realized that the reason was that he was in the last months of his life. It therefore is fitting that this volume is dedicated to him. (I have written elsewhere about his landmark contributions to the field – Sternberg, 2017). This volume thus represents the contributions of the most-cited authors in the field of intelligence, at least as represented in three textbooks published in 2011. Because one of the textbooks, the *Cambridge Handbook of Intelligence*, is edited, I believe it fair to say that the authors have been chosen to represent those scholars who the field believes to have made the highest-impact contributions to the study of intelligence.

Regrettably, some of the most highly cited scholars in the field of human intelligence have died in recent years, not just Hunt but also John B. Carroll (e.g., Carroll, 1993), John Horn (e.g., Horn, Donaldson, & Engstrom, 1981), and Arthur Jensen (e.g., Jensen, 1998), among others. This book would have been enriched greatly had these scholars lived and been willing to contribute.

The scholars who have written for this volume represent diverse perspectives, or "metaphors of mind" (Sternberg, 1990). These perspectives include primarily biological (including behavior-genetic), cognitive, cultural, developmental, psychometric, and group-difference approaches. This book does not include all possible approaches, and there are many excellent scholars, especially ones early in their careers, who have not written for it. But this certainly will not be the last edited book of advances in the field of human intelligence, and later volumes (edited by others) doubtless will include approaches that may be underrepresented here.

Although intelligence always has been important to society, one might argue that, in some respects, it is more important now than ever before. On the one hand, intelligence as measured by IQ tests increased greatly in the 20th century (Flynn, 2009). On the other hand, we are seeing in the 21st century more stupid behavior than one might have believed possible, given these rising IQs (Sternberg, 2002). Earl Hunt (1995) asked, before the dawn of the 21st century, "Will we be smart enough?" It was a good question to ask. I hope the essays in this book provide some enlightenment as to the answer!

References

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

Detterman, D. K. (Ed.) (1985). *Current topics in human intelligence* (Vol. 1). Norwood, NJ: Ablex.

Flynn, J. R. (2009). *What is intelligence? Beyond the Flynn effect*. Cambridge: Cambridge University Press.

Goldstein, S., Princiotta, D., & Naglieri, J. (Eds.) (2015). *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts*. New York: Springer.

Horn, J. L., Donaldson, G., & Engstrom, R. (1981). Apprehension, memory, and fluid intelligence decline in adulthood. *Research on Aging*, 3(1), 33–84.

Hunt, E. B. (1995). *Will we be smart enough? Cognitive capabilities of the coming workforce*. New York: The Russell Sage Foundation.

(2011). *Human intelligence*. New York: Cambridge University Press.

Hunt, E. B., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology*, 1, 194–227.

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

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

First, intelligence is, more or less, contextually (and culturally) bounded. That is, because performance criteria (such as success in school or work) differ to some degree from one cultural environment to another, the underlying components of intelligence that are relevant to predicting success may differ from one environment to another. For example, 'intelligence' for writing a novel is not *exactly* the same as 'intelligence' for solving calculus problems. That is not to say that these two intelligences are unrelated to one another. Indeed, there are many intelligences that are highly related to each other, which ultimately gives rise to the notion of 'general intelligence' (or *g*).

Second, intelligence is a 'relative' or normative construct. One of Binet's seminal contributions to the assessment of intelligence was to introduce the idea that we can best index intelligence, especially during childhood when rapid cognitive development occurs, as the individual's performance *in comparison to* a reference group (e.g., all six-year-old children). It is almost universally accepted that one can only quantify an individual's intelligence by referring to the reference or norming group. The principal advantage to this approach is that an individual's intelligence is indexed in a way that it has the same meaning, even though norming groups may change from one decade to the next (e.g., in terms of the core knowledge and skills that are within the capabilities of the larger reference group). The principal disadvantage to this approach is that it renders comparisons *across* norming groups somewhat problematic. For example, it is arguably nonsensical to say that a large sample of today's 18-year-olds is more or less 'intelligent' than a large sample of 18-year-olds in 1930. The average 18-year-old today has very different knowledge and skills from the 18-year-old in 1930, in areas of math, science, arts and literature, and so on (see, e.g., Learned & Wood, 1938). An intelligence test designed for 18-year-olds in 1930 would be expected to yield very different performance norms if administered today, yet an IQ score for 18-year-olds in 1930 on a then-current test has the same normative meaning as an IQ score for an 18-year-old today on a current test. The IQ score only tells us the individual's standing with respect to other members of the norming sample.

Third, intelligence is dynamic. That is, although one's IQ score may be relatively constant (e.g., see Thorndike, 1940), the underlying capabilities of the individual (and the reference group) change with age. Over the course of the life span, intellectual development is quite rapid in early childhood, slows in adolescence and early adulthood, and then, for many components of intellectual ability, shows declines in middle-to-late adulthood (e.g., see Schaie, 1996).

Fourth, because prediction is the key determinant of the utility of intelligence assessments, one can make a critical distinction between intelligence *potentiality* and intelligence *actuality.* These terms are derived from Aristotle's *Metaphysics* (see Gill, 2005), but they are especially appropriate for understanding the construct of intelligence, the practicalities of intelligence assessment, and the insights that can be derived from individual intelligence scores. Moreover, as will be introduced later, this particular consideration illustrates the importance of non-ability constructs in the development and expression of intelligence.

Potentiality, in Aristotle's view, can be imagined in terms of a block of bronze (metal). It has the 'potential' to become a statue of a person or many other objects. Yet, in order to realize the goal of a statue, 'work' must be done to transform the block of bronze, by carving or hammering and so on. A completed bronze statue represents an actuality – which is the result of the work done to it by the artist. In terms of intelligence, performance scores on an IQ test are an actuality, but they are not generally of interest, in and of themselves, for many of the reasons provided previously. Consistent with Wechsler's (1975) suggestions, the goal for an intelligence assessment is an index of the individual's *potential* for intellectually demanding learning and task performance. Yet, there are three problems that prevent one from reasonably equating an IQ score with an individual's potential: (a) the test score only represents the individual's actual performance, and as such, potential can only be indirectly inferred (see Anastasi, 1983); (b) although one may be able to make effective predictions of later academic and occupational achievement from a current IQ score, it is impossible to know what future scientific and/or medical developments might be made that would fundamentally change the capability of individuals of different IQ levels to acquire new intellectual skills and knowledge (e.g., so-called brain drugs or new educational instructional techniques); and (c) like Aristotle's example, the translation from the block of bronze to a statue requires the substantial investment of work time and effort on the part of the artist. For an individual to acquire new intellectually demanding knowledge and skills, he/she must invest time and effort, which in turn, implicates non-ability constructs, such as personality and motivation. In the next sections, I will discuss how these key concepts relate to the scientific study of intellectual development and expression.

Adolescent and Adult Intellectual Development

Prior to adolescence in the developed world, nearly all children are subjected to a set of relatively common educational topics (e.g., the traditional reading, writing, and arithmetic). Once they reach early adolescence, however, educational experiences become differentiated across individuals. In addition to core courses in language, math, and sciences, most secondary schools allow students to select a subset of 'elective' or optional courses across the arts, humanities, sciences, and technology domains. These opportunities present both an opportunity and a challenge to researchers who hope to use intelligence assessments for predicting individual differences in subsequent educational and occupational success. The opportunity is represented by the fact that students can choose among courses that have greater or lesser intellectual demands, and they can choose to specialize in a particular domain or to broaden their intellectual horizons across multiple domains. Selective enrollment in these courses provides the researcher with natural experiments, where the researcher can examine differences in the acquisition of knowledge and skills of students who have varied educational experiences. Researchers can examine how such enrollments lead to changes in the depth and breadth of an individual's intellectual repertoire.

The challenge for intellectual assessment, though, is perhaps more daunting than is the opportunity for understanding of intellectual growth and diversification. That is, when students no longer have educational experiences in common, it becomes problematic to compare them using a standard intelligence test. If one student chooses to complete elective courses in Spanish throughout high school, and another student chooses instead to take courses in computer programming, then it becomes difficult to figure out how to rank-order the individuals on their respective levels of intelligence. An intelligence test that included Spanish vocabulary knowledge would put the computer science student at a disadvantage, because he/she would receive no credit for knowledge of computer science, and vice versa. On one hand, an intelligence test that excluded both Spanish and computer science would inadequately sample the knowledge of these individuals, but, on the other hand, an intelligence test that sampled all of the different domains of both in-school elective courses and out-of-school courses of study would be unreasonably long and impractical to administer. This challenge only gets more difficult as students transition from secondary school to higher education or occupations, because the content of their respective intellectual repertoires gets increasingly differentiated and specialized.

The traditional solution to such challenges has been to focus only on what knowledge and skills are common to most students (i.e., not directly sampling knowledge and skills from elective courses), and is further compromised when testing adults, who are many years beyond their high school educational experience. For example, the SAT and ACT tests, used for college/university selection, only assess mathematics knowledge and skills through algebra and geometry, because only a portion of the college-applying population advances to elective courses beyond these topics (e.g., calculus). Four years after the student completes the SAT or ACT, he/she might be considering postgraduate study. Yet, because of the lack of common core courses at the college/university level, the most widely used entrance examination for graduate study, the Graduate Records Examination (GRE) is still only testing algebra and geometry – topics that some students may have only encountered in high school, while other students may have continued with a rigorous study of advanced mathematics at university. Cattell (1957) called this testing of 'historical' crystallized intelligence (Gc), as opposed to 'current' Gc.

As individuals reach adulthood, what they can accomplish on intellectually demanding tasks becomes much more importantly determined by their prior specialized experience. Nearly every profession or expert performance depends on knowledge that has been acquired over a long period of learning and practice. Indeed, I have previously argued (e.g., Ackerman, 1996) that most of the tasks that adults perform on a day-to-day basis are much more highly associated with an adult's specialized knowledge and skills, rather than the kinds of intelligence associated with abstract reasoning and working memory. Jobs that vary broadly share this fundamental property, whether in health care (doctors, nurses), other knowledge work (e.g., accounting, law, science), and in various 'trades' (e.g., carpentry, plumbing). Ultimately, this turns out to be fortuitous for adults, because with increasing age into the middle-adult years, there is typically a decline in the 'fluid' intellectual abilities (Gf), relative to adolescents and younger adults (Cattell, 1943; Hebb, 1942). The implication of these changes is that middle-aged and older adults are less effective in performing abstract reasoning kinds of tasks, that in turn, appear to be important for the acquisition of novel task knowledge and skills. But adults who have acquired expertise in their own professions or other areas often have an advantage in acquiring new knowledge and skills within their own areas of expertise, because transfer-of-training/transfer-of-knowledge is a very powerful positive influence for acquisition of new knowledge, when it can be incorporated into existing knowledge structures (e.g., see Ferguson, 1956).

Intelligence of Young and Middle-Aged Adults

For much of the modern period of intelligence theory and assessment, it has been claimed that intelligence declines in middle-aged years, compared to adolescents and young adults (for a review, see Ackerman, 2000). The evidence for this is somewhat complex, because as noted earlier, IQ scores for different age cohorts – those born in different decades – are fundamentally incommensurable, because intelligence tests are normed for particular cohort groups. Thus cross-sectional studies, where groups of individuals of different age cohorts are given the same intelligence test, yield results where age effects are confounded with cohort differences (Schaie & Strother, 1968). Longitudinal studies, where the same individuals are given the same intelligence test repeatedly, are more informative about the effects of aging compared to cross-sectional studies, but they have other confounds that must be taken into account (such as practice effects). Nonetheless, the accumulated evidence across these studies strongly supports the notion that in adulthood, there is a normative decline in Gf abilities, but much less decline or stability in 'historical' Gc, at least into later adulthood, when there are normative declines, with stronger decline gradients for Gf, compared to Gc (Schaie, 1996). Great efforts have been expended in recent decades to determine factors that may slow or stop the decline of intellectual abilities with increasing age in adulthood, ranging from so-called brain-training games to physical exercise. A discussion of the efficacy of such programs is beyond the scope of this chapter, but see Hertzog and colleagues (2009).

Directly Assessing the Knowledge Components of Intelligence

In studies examining *current* Gc in young and middle-aged adults, we developed tests of content knowledge across a wide spectrum of domains of intellectual expertise. While one cannot reasonably hope to sample all different types of knowledge possessed by adults, we obtained a representative sampling of areas of knowledge that are found in both traditional classrooms and advanced study areas in postsecondary education (e.g., physical and social sciences, literature, art, business, and law), and also domains outside the traditional educational context (knowledge of current events, health and safety, technology, financial planning). Performance was indicated by raw scores rather than norm-based, so that direct comparisons are made between age groups, while keeping in mind that different cohort groups may have different levels of experience and score would be referenced to the sex of the examinee (Yerkes, Bridges, & Hardwick, 1915). Terman, however, decided that there was adequate justification for equality of IQ scores across the sexes, and so he constructed his IQ test to be specifically balanced by including subtests where the sex differences in the overall scale were eliminated. Subsequent IQ tests generally adopted this same approach to eliminating sex differences.

But, when it comes to individual domain knowledge tests that are content-referenced rather than norm-referenced, sex differences are clearly observed. The majority of the academic domain knowledge tests (e.g., Ackerman & Rolfhus, 1999) show advantages to males, though such differences are not typically found in current-events knowledge tests, and women have a distinct advantage in domains of health knowledge (Beier & Ackerman, 2001, 2003). When one examines sex differences in knowledge tests where the individuals self-select into particular areas of study, these differences are also seen in young adults (College Board, 2011). Ultimately, these results suggest that both individual and sex differences relate to the *direction* and *intensity* of effort devoted to the acquisition of domain-specific knowledge and skills.

Ability and Non-ability Traits and Intellectual Investment

Elementary education is largely a system for transmitting core educational content, and as such, there is great commonality among students in terms of the instruction they receive. Homework, for example, starts off relatively modest in demands for time and effort on the part of students. Once students reach secondary school, they have options toward or away from the investment of their time and effort for acquiring knowledge in intellectually demanding domains. Homework often increases in terms of time and effort, and demands consequently increase for self-regulated cognitive investments. It is during this critical period that an individual's personality and motivational traits appear to increase in influence on the direction and intensity of intellectual investments. For a conceptual discussion of investment and intellectual development, see Cattell (1971; also see Schmidt, 2014; von Stumm & Ackerman, 2013). Intellectual investments continue through decisions about postsecondary education, including whether to attend university study, selection of a major, and choice of early career paths. Together with Gf and both historical and current Gc abilities, non-ability traits also appear to be influential in determining how individuals invest their cognitive resources well into middle adulthood, in terms of seeking out or avoiding intellectual challenges, such as acquiring new knowledge and skills in and out of the workplace, and in terms of refining and improving one's performance on relatively routine tasks.

Several personality and motivational traits are, or become, associated with individual differences in intellectual abilities and domain knowledge during adolescent and adult development. Affective (personality) traits such as openness to experience and conscientiousness are positively related to individual differences in domain knowledge in many areas, while personality traits like neuroticism and extroversion tend to be negatively related to domain knowledge. Similarly, conative (will, motivation) traits such as a mastery orientation or a desire to learn are positively related to individual differences in domain knowledge, while worry and anxiety in achievement contexts are negatively related to individual differences in domain knowledge. In addition, there is a moderate association of vocational interests to differences in domain knowledge, such as investigative interests and artistic interests being positively associated with domain knowledge in the sciences and humanities, and a negative association between social and enterprising interests and a variety of academic knowledge domains. These non-ability traits are related to one another, even though they represent different aspects of individuals. This commonality has been a major factor in the development of the concept of "trait complexes" (Ackerman & Heggestad, 1997) – that is, constellations of personality, motivation, and other traits that: (a) appear more frequently in the population, and (b) are associated with orientations toward or away from intellectual development. Trait complexes of intellectual/cultural traits and science/math traits are associated with higher levels of domain knowledge in the arts, humanities, and social sciences, and in STEM (science, technology, engineering, and math) domains, respectively. Complexes of social and conventional traits are associated with lower levels of knowledge in a variety of academic and other intellectually demanding domains (Ackerman, 2000). Based on these considerations, a general framework for understanding adult intellectual development can be illustrated as shown in Figure 1.2.

In the figure, early adolescent intellectual potentiality is represented in terms of what is measured with an IQ test, that is, Gf and historical Gc. As individuals develop into adulthood, non-ability trait complexes interact with levels of intellectual potentiality to determine the investment (time and effort) the individual makes into one or more of a variety of different directions, both intellectual and non-intellectual. The result is found in an adult's breadth and depth of domain knowledge and skills, which represent the vocational and avocational (e.g., hobbies) intellectual repertoire of the individual. I propose this is the *main* source of individual differences in

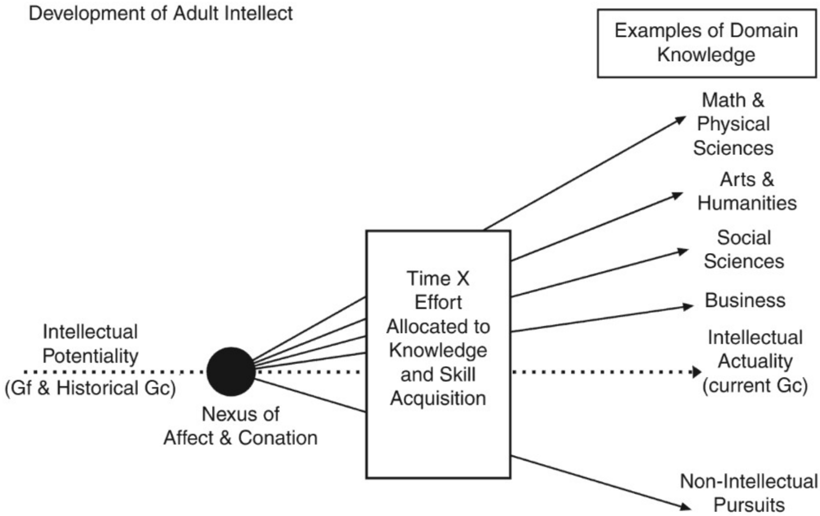

Figure 1.2 An abstracted illustration of the influences on intellectual growth from adolescent intellectual potentiality to adult intellect actuality, based on Ackerman (1996).

adult intelligence. This is *not* to say that Gf-type abilities are unimportant in adulthood (Deary et al., 2004), but it is to say that for the majority of day-to-day activities, Gf abilities are much less influential in determining an individual's performance, in comparison to the individual's current domain-specific and domain-general Gc abilities. In a nutshell, the proposition is that it is easier and more effective to *know and recall* the answer to an intellectually demanding problem than it is to figure it out from scratch using abstract reasoning (Ackerman, 1996) – which in essence, is the difference between an 'expert system' and a 'general-problem-solving' engine. One would hardly be advised to get onto an airplane piloted by an individual with high Gf but no pilot experience, over an individual with perhaps an average level of Gf but 20 years of experience piloting commercial aircraft.

Current Issues

From the theoretical foundation outlined earlier, and the body of research reviewed, two issues are important for educational and public policy perspectives. First and foremost is the finding of sex differences in domain knowledge among adolescents and adults. Large sex differences in knowledge about STEM areas among adolescents are likely a nontrivial determinant of disparities between men and women who select and persist in STEM majors in postsecondary study, and in later career choices (Ackerman, Kanfer, & Calderwood, 2013). That these differences are manifest during the high school years suggests that it may generally be too late to remedy these differences by the time an adolescent starts college. Efforts are needed to understand whether there are systematic influences external to the individual that are responsible for these differences (e.g., school policies, parental or peer influences), or whether the influences are largely internally driven, in terms of the student's interests, preferences, and personality characteristics.

The second issue is related to the first. That is, what is needed is a better understanding of the malleability of the connections between non-ability traits and an individual's investment of effort toward acquisition of knowledge and skills in particular domains, and overall. While a scientific consensus exists that Gf and historical Gc are limited in malleability, at least within the variety of environments students encounter in the developed world, substantially less research has been conducted that explores the limits of guidance or instruction on either developing affective and conative traits to better focus student efforts toward acquisition of domain knowledge, or to modifying the connections between these non-ability traits and acquisition of domain knowledge. Efforts in this area might have benefits both in reducing sex differences in STEM achievements and in generally improving the educational and occupational outlooks for many students.

Future Directions

In many ways, theory and research on traditional IQ assessment for children and early adolescents has become moribund, perhaps partly because of the clear success of the Binet-inspired tests for predicting overall academic success in the elementary school system. Yet, if one considers that adult intelligence differs fundamentally from Gf and historical Gc, in that it includes the breadth and depth of current Gc knowledge and skills, understanding of adult intelligence is woefully incomplete. Assessments that give credit to adults for the wide variety of knowledge and skills that they possess have yet to be developed. A high proportion of an adult's day-to-day intellectual life is simply unaccounted for by modern IQ assessments. As a result, there is little knowledge about how current Gc develops, is maintained, or declines, especially in older-age populations. Current Gc is essentially equivalent to the 'dark matter' hypothesized by physicists.

That is, current Gc is clearly necessary to gain a complete understanding of how adults function in an intellectually demanding society, even as Gf abilities decline with each additional decade of adult life, yet it has not been adequately measured and described. Perhaps the second century of modern intelligence theory and assessments will usher in a reorientation to both the intellectual and nonintellectual determinants of the adult repertoire for intelligent task performance.

References

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, 22, 229–259.
(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., & Beier, M. E. (2006). Determinants of domain knowledge and independent study learning in an adult sample. *Journal of Educational Psychology*, 98, 366–381.
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., & Calderwood, C. (2013). High school Advanced Placement and student performance in college: STEM majors, non-STEM majors, and gender differences. *Teachers College Record*, 115 (10), 1–43.
Ackerman, P. L., & Rolfhus, E. L. (1999). The locus of adult intelligence: Knowledge, abilities, and non-ability traits. *Psychology and Aging*, 14, 314–330.
Anastasi, A. (1983). Evolving trait concepts. *American Psychologist*, 38, 175–184.
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.
(2003). Determinants of health knowledge: An investigation of age, gender, abilities, personality, and interests. *Journal of Personality and Social Psychology*, 84 (2), 439–448.
Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40, 153–193.
(1957). *Personality and motivation structure and measurement*. Yonkers-on-Hudson, NY: World Book Company.
(1971/1987). *Abilities: Their structure, growth and action*. Amsterdam: North-Holland.
College Board (2011). AP Data 2011. http://professionals.collegeboard.com/data-reports-research/ap/data
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.

It is worth emphasizing Darwin's astute comment that there is a difference between intelligence and motivation (zeal) and effort (hard work), and that the difference is important. Galton himself was well aware of the difference and argued that all three were influenced by heredity. "The triple event, of ability combined with zeal and with capacity for hard labour[,] is inherited" (Galton, 1892/1962, p. 78). Galton's speculative proposal has been nicely confirmed. We now know that virtually all traits (human and nonhuman, psychological and otherwise) are influenced by heredity (Bouchard, 2004; Lynch & Walsh, 1998, p. 175; Polderman et al., 2015).

What Is *g*?

The definition of intelligence given previously in this chapter, like almost all others, implies that many mental processes (abilities) underlie intelligence and that they are related at a deeper level. The idea that there are many independent mental abilities (faculties) is a very old one and used to be called faculty psychology. It continues to be manifest today in the brilliant work of Tooby and Cosmides (2015). These authors argue strongly that many abilities evolved to solve very specific problems – they are modular and adaptive. I largely agree with this view. What intelligence researchers like myself are concerned with is the fact that virtually all measures of mental abilities correlate positively and in many instances quite strongly. Guttman and Levey (1991) call this the first law of intelligence testing. This empirical fact can be characterized as a general factor, or, more specifically, as Spearman's *g*, in honor of the investigator who first described it mathematically (Spearman, 1904). For investigators who approach the question "What is intelligence?" from a factor point of view, the more technical question is, "What is *g*?"

A useful way to address this question is by looking at it from the point of view of its most severe critics. Steven J. Gould[2] (1981, 1996), the distinguished paleontologist, evolutionary biologist, and historian of science, argued strongly against the idea of a general factor of intelligence. He argued that belief in *g* constituted an error of reification:

> The notion that such a nebulous socially defined concept as intelligence might be identified as a "thing" with a locus in the brain and a definite degree of heritability – and that it might be measured as a single number, thus permitting a unilinear ranking of people according to the amount they possess. (Gould, 1981, p. 239)

In addition he argued that *g* was chimerical.

> Spearman's g is not an ineluctable entity; it represents one mathematical solution among many equivalent alternatives. The chimerical nature of g is the rotten core of Jensen's edifice, and the entire hereditarian school. (Gould, 1981, p. 320)

Spearman's *g* is a theoretical construct, not a "thing." Whether it has a "locus in the brain" is an empirical question (Korb, 1994). Is there a physical substructure to intelligence? Modern brain mapping suggests that there may well be (Haier, 2017). Indeed, I believe that advances in this domain, in conjunction with molecular genetics, are among the most promising future avenues of research in the domain of intelligence (Colom, 2014; Ponsoda et al., 2016).

Is *g* chimerical or ineluctable? Gould's understanding of factor analysis was, in Bartholomew's words, "half a century out of date" (Bartholomew, 2004, p. 70). Experts in the field strongly reject Gould's views (Reeve & Charles, 2008, p. 685). There is simply no doubt that a *g* factor is unavoidable (ineluctable) when correlation matrices of mental abilities are examined empirically (Reeve & Blacksmith, 2009). It is sometimes asserted that the *g* in one battery of tests is different from the *g* derived from a different battery. This is simply not true if each battery contains a reasonable number of tests and samples a broad set of abilities (Major, Johnson, & Bouchard, 2011). When such an assessment is carried out appropriately, there is just one *g* (Johnson & Bouchard, 2011; Salthouse, 2013).

Do we need a definition of intelligence? My answer is yes. The literature is rife with poor measures of *g* and claims that "this or that *g*" fails to predict key outcomes (i.e., academic achievement, etc.) better than some alternative (e.g., so-called complex problem solving) (Lotz, Sparfeldt, & Greiff, 2016). As pointed out earlier, my preferred definition requires numerous mental abilities (reasoning, planning, solving problems, thinking abstractly, comprehending complex ideas). Without such a guide, the choice of tests to include in a battery designed to measure *g* will be much too narrow. The classic case is the use of the Raven as a substitute for *g*; it is far from sufficient, as no single measure is adequate (Gignac, 2015). This is the problem of factor indeterminacy. Lee and Kuncel (2015) provide a thoughtful discussion applicable to any general factor. There is no substitute for careful measurement of the construct of interest and that requires both a meaningful definition implemented in the form of a theory and adequate quantification of the theoretical construct. To paraphrase my former colleague Paul Meehl, "theories built around poorly conceived constructs are scientifically unimpressive and technologically worthless" (Meehl, 1978, p. 806). I and others (Borsboom, 2013) believe that his arguments continue to be valid.

For those who prefer *Grit* and *Practice* over *g*, and zeal and hard work as alternative explanations of achievement, I refer them to Simonton(2016), who found the arguments of the major proponents of these constructs less than adequate. Others agree with this argument (Crede, Tynan, & Harms, 2016; Hambrick et al., 2016). I continue to prefer the terms "zeal" and "hard work." As McNemar pointed out long ago, the first cardinal principle of psychological progress is: "Give new names to old things" (1964, p. 872).

Those who believe in the threshold hypothesis – "there is little evidence that those scoring at the very top of the range in standardized tests are likely to have more successful careers in the sciences" (Muller et al., 2005) – are simply wrong (Arneson, Sackett, & Beatty, 2011). Monotonicity even applies when one looks at the top 1% of the ability distribution (Makel et al., 2016). Indeed the opposite of a threshold effect (increasing predictive power at higher levels) may be true (Coyle, 2015).

The Heritability of Intelligence

I have always had an interest in biology, and my fondest memory of high school was laboratory work in a biology course. Early exposure to biological and genetic thinking (Bouchard, 2016b) prepared me for my most influential work, a study of twins reared apart (Bouchard et al., 1990a; Segal, 2012). I would like to emphasize two important facts about this work. It is an experimental study, a fact that is widely underappreciated and often barely recognized. First, twins are an experiment of nature. In simple terms, monozygotic twins (MZ) share all their genes and dizygotic twins (DZ) share half their genes. When they are reared apart, they allow us to estimate the magnitude of genetic influence on any trait. In particular, the correlation for IQ of monozygotic twins reared apart (MZA) directly estimates the heritability (Bouchard et al., 1990b). There are additional complexities (Segal, 2017) but, as I show in what follows, it is a very good model when used in conjunction with other designs. Second, adoption is an experiment of society. Again there are complexities, but it is still a very good model. Thus, we have a combination of an experiment of nature and an experiment of nurture. Unlike laboratory experiments, the adoption experiment is a very powerful one in terms of magnitude of influence, as it involves a treatment that is applied day in and day out for many years.[3] Before turning to work on the genetics of intelligence, I will again discuss the point of view of an intense critic.

Gould Again

The idea that intelligence has a heritable component has been vigorously attacked for a very long time, sometimes viciously, as noted by Gould's use of the term "rotten core" cited earlier. Gould's attacks focused on the works of Cyril Burt with monozygotic twins reared apart and Burt's work on social mobility (1981, chap. 6). Gould drew largely on the writings of Leon Kamin (1974) and Donald Dorfman (1979). I hesitate to argue that this long-running controversy is over, but I will assert that the so-called evidence brought to bear against Burt in support of the accusation of fraud is far from conclusive. Dorfman claimed that "The eminent Briton is shown, beyond reasonable doubt, to have fabricated data on IQ and social class"[4] (p. 117). According to a panel that reexamined the "Burt Affair" (Mackintosh, 1995), the data on IQ and social class were key in deciding that Burt was guilty of fraud, as there was reasonable doubt about other charges. Tredoux (2015) has shown that critics of Burt's social mobility work did not understand his procedures and demonstrated that methods available to Burt at the time of his analysis easily explain his purportedly falsified results. Gould's larger focus, captured in his title "The Mismeasure of Man," argued that work in this domain was largely characterized by bias. There is now striking evidence to suggest that it was Gould who was biased in his analysis and interpretation of the data gathered by others (Fancher, 1987; Glenn & Ellis, 1988; Lewis et al., 2011; Zenderland, 1988). Even more interestingly, data Gould took at face value – work by Franz Boas (1912) – rather than examining for bias, and that he used as a basis of criticism of the hereditarian paradigm (Gould, 1981, p. 108), have turned out to support the hereditarian paradigm (Sparks & Jantz, 2002). For a more thorough analysis of Gould's many errors, omissions, and distortions, see Bouchard (2014) and Rushton (1997). Pinker discusses the similar role Gould played in the larger "sociobiology wars" (Pinker, 2002, chap. 6).

The Wilson Effect

A major reason why there was so much controversy regarding the magnitude of genetic influence on intelligence is the fact that there is a massive age (developmental) effect. For many years, psychologists believed that genetic influence was manifest at birth and experience altered behavioral and other traits. For example, J. P. Scott, one of the founders of behavior genetics as a systematic discipline, asserted:

> We thought that the best time to study the effects of genetics would be soon after birth, when behavior still had little opportunity to be altered by experience. On the contrary, we found that the different dog breeds were most alike as newborns; that is, genetic variation in behavior develops postnatally, in part as a result of the timing of gene action and in part from the interaction of gene action and experience, social, and otherwise. (Scott, 1990, p. vii)

This effect appears to be rather general (Bergen, Gardner, & Kendler, 2007). In the domain of intelligence, the effect is now called the Wilson Effect (Bouchard, 2013), and the results of the various relevant studies are shown in Figure 2.1.

Figure 2.1 illustrates that the results are consistent a) across multiple research designs (twins, adoptees, various combinations of kinships), b) measure of intelligence, c) Westernized industrialized countries, and d) kinds of samples, some very comprehensive and others much more restricted.

I have always thought it was amazing that while psychologists and others heavily emphasize the role of family environment, thus the emphasis on socioeconomic status (SES), in the shaping of intelligence in children, they conducted almost no studies of unrelated individuals reared together (URT). The URT design is the most powerful one to assess this source of influence. As Figure 2.1 shows, this design suggests a value near zero in adulthood for shared environment (see the asterisks in Figure 2.1), a value below that suggested by twin designs, namely, about 10%. My view is that psychologists have been plagued by confirmation bias and highly resistant to strong inference and refutation of their theories (Bouchard, 2009). The influence of genes on IQ and SES was laid out for us a great many years ago by a brilliant and highly underappreciated psychologist, namely Barbara Burks (Burks, 1938; King, Montanez-Raminez, & Wertheimer, 1996).

The Structure of Mental Abilities

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

I have revised the theory somewhat and applied it to intelligence (Bouchard, 2014), genius (Johnson & Bouchard, 2014), and personality (Bouchard, 2016a). The "theory" is admittedly weak in the sense it is difficult to refute as currently formulated and it should perhaps be called a "meta-theory" or a "heuristic" pointing investigators in a potentially fruitful direction. What it does do, however, is give a specific name to what I believe is a widely held point of view, namely, that the mind has been shaped by the environment, via evolution, and that the content of individual minds is shaped to an important extent by the content of the environment.

Behavior geneticists have long held this view. A nice example applied to genetic influence on social attitudes is given next.

> In no way does our model minimize the role of learning and social interaction in behavioral development. Rather, it sees humans as exploring organisms whose innate abilities and predispositions help them select what is relevant and adaptive from the range of opportunities and stimuli presented by the environment. The effects of mobility and learning, therefore, augment rather than eradicate the effects of the genotype on behavior. (Martin et al., 1986, p. 4368)

EPD theory needs to be more rigorously formulated, but, if correct, it has the virtue of answering the "how" question (Anastasi, 1958). The answer is "nature via nurture"; that is, "the genome impresses itself on the psyche largely by influencing the character, selection, and impact of experience during development" (Bouchard et al., 1990a. p. 228).

Notes

1 The book was originally published in 1869. In the 1892 edition Galton admitted that the title was misleading, that it had little to do with genius, and that it should have been titled *Hereditary Ability* (Galton, 1892/1962, p. 26). As Darwin noted in the quote that follows, the idea of "intellect," a fixed characteristic or a trait in which individuals did not differ, has a very long history.

2 Gould was one of the most widely read scientists of the 20th century and was highly influential among both academics and the literate public (Shermer, 2002).

3 It is not widely recognized that any given experimental manipulation is simply one of many possible implementations of a causal mechanism and is not an infallible procedure (Johnson & Bouchard, 2014, footnote 1). The "fadeout effect" is a dramatic example in the domain of intelligence research. Interventions appear to influence *g*, but the effect fades with time (Protzko, 2016).

4 The reason the IQ and social class issue enters the discussion is because it relates to the hereditarian argument that higher-IQ individuals migrate to higher social classes via the influence of IQ on education and occupational success.

Ipso facto higher-social-status individuals are genetically superior, at least with respect to IQ. This has been a taboo topic (Bouchard, 1995). The classic adoption studies on IQ and social class are, in my view, dispositive (Scarr & Weinberg, 1978), and molecular genetic techniques have begun to confirm that conclusion (Kong et al., 2017; Selzam et al., 2016), although the actual effect sizes remain quite small.

5 I had a special interest in spatial and mental rotation abilities (Bouchard & McGee, 1977; Lubinski, 2010), and, as a result, the MISTRA test batteries more adequately represent this domain relative to most other batteries. Work on the VPR model was spearheaded by my colleague Wendy Johnson.

References

Anastasi, A. (1958). Heredity, environment, and the question "how?" *Psychological Review*, 65, 197–208. doi:10.1037/h0044895.

Arneson, J. J., Sackett, P. R., & Beatty, A. S. (2011). Ability–performance relationships in education and employment settings: Critical tests of the more-is-better and the good-enough hypotheses. *Psychological Science*, 22(10), 1336–1342. doi:10.1177/0956797611417004.

Bartholomew, D. J. (2004). *Measuring intelligence: Facts and fallacies*. Cambridge: Cambridge University Press.

Bergen, S. E., Gardner, C. O., & Kendler, K. S. (2007). Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: A meta-analysis. *Twin Research and Human Genetics*, 10(3), 423–433. doi:10.1375/twin.10.3.423.

Boas, F. (1912). Changes in the bodily form of descendants of immigrants. *American Anthropologist*, 14, 530–562.

Borsboom, D. (2013). Theoretical amnesia. *Centre for Open Science.* Retrieved from: http://centerforopenscience.github.io/osc/2013/11/20/theoretical-amnesia/.

Bouchard, T. J., Jr. (1995). Breaking the last taboo. Review of "The Bell Curve." *Contemporary Psychology*, 40, 415–421.

(1996). Behavior genetic studies of intelligence, yesterday and today: The long journey from plausibility to proof – the Galton lecture. *Journal of Biosocial Science*, 28, 527–555. doi:10.1017/S0021932000022574.

(2004). Genetic influence on human psychological traits: A survey. *Current Directions in Psychological Science*, 13, 149–151.

(2009). Strong inference: A strategy for advancing psychological science. In K. McCartney & R. Weinberg (Eds.), *Experience and development: A festschrift in honor of Sandra Wood Scarr* (pp. 39–59). London: Taylor and Francis.

(2013). The Wilson effect: The increase in heritability of IQ with age. *Twin Research and Human Genetics*, 16, 923–930. doi:10.1017/thg.2013.54.

(2014). Genes, evolution and intelligence. *Behavior Genetics*, 44, 549–577. doi:10.1007/s10519-014-9646-x.

(2016a). Experience producing drive theory: Personality "writ large." *Personality and Individual Differences*, 90, 302–314. doi:10.1016/j.paid.2015.11.007.

(2011). Intelligence and social inequality: Why the biological link? In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences*. Malden, MA: Wiley-Blackwell.
Gould, S. J. (1981). *The mismeasure of man*. New York: W. W. Norton.
(1996). *The mismeasure of man* (2nd edn.). New York: Norton.
Gustafsson, J. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 179–203.
Guttman, L., & Levey, S. (1991). Two structural laws for intelligence tests. *Intelligence*, 15, 79–103.
Haier, R. J. (2017). *The neuroscience of intelligence*. New York: Cambridge University Press.
Hambrick, D. Z., Macnamara, B. N., Campitelli, G., Ullén, F., & Mosing, M. A. (2016). Beyond born versus made: A new look at expertise. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 64, pp. 1–55). Elsevier Academic Press.
Hayes, K. J. (1962). Genes, drives, and intellect. *Psychological Reports*, 10, 299–342. doi:10.2466/pr0.1962.10.2.299.
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.
(2005b). The structure of human intelligence: It's Verbal, Perceptual, and Image Rotation (VPR), not Fluid Crystallized. *Intelligence*, 33, 393–416. doi:10.1016/j.intell.2004.12.002.
(2011). The MISTRA data: Forty-two mental ability tests in three batteries. *Intelligence*, 39(2–3), 82–88. doi:10.1016/j.intell.2011.02.010.
(2014). Genetics of intellectual and personality traits associated with creative genius: Could geniuses be Cosmobian Dragon Kings?. In D. K. Simonton (Ed.), *The Wiley handbook of genius* (pp. 269–296). Oxford: Wiley.
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. doi:10.1016/j.intell.2006.05.002.
Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2), 183–202. doi:10.1007/bf02289343.
Kamin, L. J. (1974). *The science and politics of IQ*. Potomac, MD: Erlbaum.
King, D. B., Montanez-Raminez, L. M., & Wertheimer, M. (1996). Barbara Stoddard Burks: Pioneer behavioral geneticist and humanitarian. In G. A. Kimble, C. A. Boneay, & M. Wertheimer (Eds.), *Portraits of pioneers in psychology; Volume II* (pp. 213–225). Washington, DC: American Psychological Association.
Kong, A., Frigge, M. L., Thorleifsson, G., Stefansson, H., Young, A. I., Zink, F., … Stefansson, K. (2017). Selection against variants in the genome associated with educational attainment. *Proceedings of the National Academy of Science USA*, 114(5), E727–E732. doi:10.1073/pnas.1612113114.
Korb, K. B. (1994). Stephen Jay Gould on intelligence. *Cognition*, 52(2), 111–123.

Lee, J. J., & Kuncel, N. R. (2015). The determinacy and predictive power of common factors. *Industrial and Organizational Psychology*, 8(03), 467–472. doi:10.1017/iop.2015.64.
Lewis, J. E., Degusta, D., Meyer, M. R., Monge, J. M., Mann, A. E., & Holloway, R. L. (2011). The mismeasure of science: Stephen Jay Gould versus Samuel George Morton on skulls and bias. *PLoS Biology*, 9(6), e1001071. doi:10.1371/journal.pbio.1001071.
Lotz, C., Sparfeldt, J. R., & Greiff, S. (2016). Complex problem solving in educational contexts – Still something beyond a "good g"? *Intelligence*. doi:10.1016/j.intell.2016.09.001.
Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences*, 49, 344–351. doi:10.1016/j.paid.2010.03.022.
Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sunderland, MA: Sinauer.
Mackintosh, N. J. (Ed.) (1995). *Cyril Burt: Fraud or framed*. Oxford: Oxford University Press.
Major, J. T., Johnson, W., & Bouchard, T. J., Jr. (2011). The dependability of the general factor of intelligence: Why small, single-factor models do not adequately represent *g*. *Intelligence*, 39, 418–433. doi:10.1016/j.intell.2011.07.002.
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. doi:10.1016/j.intell.2012.07.006.
Makel, M. C., Kell, H. J., Lubinski, D., Putallaz, M., & Benbow, C. P. (2016). When lightning strikes twice: Profoundly gifted, profoundly accomplished. *Psychological Science*, 27(7), 1004–1018. doi:10.1177/0956797616644735
Martin, N. G., Eaves, L. J., Heath, A. C., Jardine, R., Feingold, L. M., & Eysenck, H. J. (1986). Transmission of social attitudes. *Proceedings of the National Academy of Science USA*, 83(12), 4364–4368.
McNemar, Q. (1964). Lost: Our intelligence? Why? *American Psychologist*, 19, 871–882.
Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46, 806–834.
Muller, C. B., Ride, S. M., Fouke, J., Whitney, T., Denton, D. D., Cantor, N., … Robinson, S. (2005). Gender differences and performance in science. *Science*, 307(5712), 1043. doi:10.1126/science.307.5712.1043b.
Pearson. (1914–1930). *The life, letters and labours of Francis Galton, Volume 1*. Cambridge: Cambridge University Press.
Pinker, S. (2002). *The blank slate: The modern denial of human nature*. New York: Viking.
Platt, J. R. (1964). Strong inference. *Science*, 146, 347–353.
Polderman, T. J. C., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, 47, 702–709. doi:10.1038/ng.3285.

Ponsoda, V., Martinez, K., Pineda-Pardo, J. A., Abad, F. J., Olea, J., Roman, F. J., … Colom, R. (2016). Structural brain connectivity and cognitive ability differences: A multivariate distance matrix regression analysis. *Human Brain Mapping*. doi:10.1002/hbm.23419.

Protzko, J. (2016). Does the raising IQ–raising g distinction explain the fadeout effect? *Intelligence*, *56*, 65–71. doi:10.1016/j.intell.2016.02.008.

Reeve, C. L., & Blacksmith, N. (2009). Identifying *g*: A review of current factor analytic practices in the science of mental ability. *Intelligence*, 37, 487–494. doi:10.1016/j.intell.2009.06.002.

Reeve, C. L., & Charles, J. E. (2008). Survey of opinions on the primacy of g and social consequences of ability testing: A comparison of expert and non-expert views. *Intelligence*, 36(6), 681–688. doi:10.1016/j.intell.2008.03.007.

Rushton, J. P. (1997). Race, intelligence, and the brain: The errors and omissions of the "revised" edition of S. J. Gould's *The mismeasure of man* (1996). *Personality and Individual Differences*, 23, 169–180.

Salthouse, T. A. (2013). Evaluating the correspondence of different cognitive batteries. *Assessment*. doi:10.1177/1073191113486690.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype -> environment effects. *Child Development*, 54, 424–435. doi:10.2307/1129703.

Scarr, S., & Weinberg, R. A. (1978). The influence of "family background" on intellectual attainment. *American Sociological Review*, 43, 674–692.

Scott, J. P. (1990). Foreword. In M. E. Hahn, J. K. Hewitt, N. D. Henderson, & R. H. Benno (Eds.), *Developmental behavior genetics: Neural, biometrical, and evolutionary approaches*. New York: Oxford.

Segal, N. L. (2012). *Born together–reared apart: The landmark Minnesota twin study*. Cambridge, MA: Harvard University Press.

(2017). *Twin mythconceptions: False beliefs, fables, and facts about twins*. New York: Academic Press.

Selzam, S., Krapohl, E., Von Stumm, S., O'Reilly, P. F., Rimfeld, K., Kovas, Y., … Plomin, R. (2016). Predicting educational achievement from DNA. *Molecular Psychiatry*, 22(2), 267–272. doi:10.1038/mp.2016.107.

Shermer, M. B. (2002). Stephen Jay Gould as historian of science and scientific historian, popular scientist and scientific popularizer. *Social Studies of Science*, 32, 489–525.

Simonton, D. K. (2016). Intelligence, inheritance, motivation, and expertise. Review of *Grit: The power of passion and perseverance*, A. Duckworth, and *Peak: Secrets from the new science of expertise*, A. Ericsson & R. Pool. *Intelligence*, 58, 80–81. doi:10.1016/j.intell.2016.05.005.

Sparks, C. S., & Jantz, R. L. (2002). A reassessment of human cranial plasticity: Boas revisited. *Proceedings of the National Academy of Sciences*, 99(23), 14636–14639. doi:10.1073/pnas.222389599.

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

CHAPTER 3

Culture, Sex, and Intelligence

Stephen J. Ceci, Donna K. Ginther, Shulamit Kahn, & Wendy M. Williams

In this chapter we focus on findings from our research on sex differences in academic achievement and what they say about the role of culture in shaping mathematical and spatial cognition. Our research focuses on the policy and educational implications of spatial and mathematical ability that are correlated with psychometric data (e.g., SAT, GRE, NAEP) and raises questions about the nature and development of these differences and what role policy has in ameliorating them.

Sex Differences in Quantitative Fields

Women are underrepresented in all math-intensive fields in the academy. According to the NSF's 2010 Survey of Doctorate Recipients (SDR), women in **G**eoscience, **E**ngineering, **E**conomics, **M**ath/Computer science, and the **P**hysical sciences (**GEEMP**) in 2010 comprised only 25%–44% of tenure-track assistant professors and only 7%–16% of full professors. There is debate over why women are so conspicuously absent in these fields compared with the **L**ife sciences, **P**sychology, and **S**ocial sciences (**LPS**), where the comparable figures show women at 66% of tenure-track assistant professorships in psychology, 45% in social sciences (excluding economics), and 38% in life sciences; for full professors, the figures are 35%, 23%, and 24%, respectively. So, compared with their presence in LPS fields, women's presence in mathematically intensive (GEEMP) fields is much lower. Why is this? What does it say about spatial and quantitative aptitude? And what, if anything, ought to be done to narrow this gap between the sexes? To answer these questions, we have taken a developmental perspective, starting early in life and tracking cohorts through adulthood. Before examining early sex differences, however, we review sex differences at the college, graduate school, and professional levels.

Females comprise 57% of college graduates and 57% of STEM (Science, Technology, Engineering, and Mathematics) majors. Figure 3.1 shows the

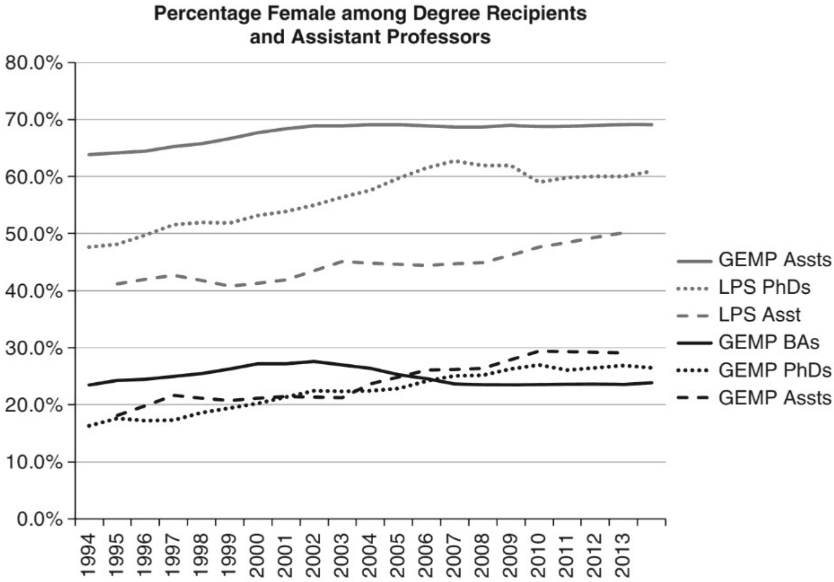

Figure 3.1 Percent of bachelor and doctorate degrees awarded to females and percent female of assistant professors, by STEM category. Data drawn from the NSF's WebCASPAR database. (ncsesdata.nsf.gov/webcaspar/) and the NSF's 1973–2010 Survey of Doctorate Recipients.

percentages of females among college graduates, PhD recipients, and assistant professors. Since 2000, 69% of majors in the life sciences, psychology, and social sciences (LPS) are women. In contrast, women comprise only a quarter of undergraduates in GEEMP fields. Because of these disparities across fields, combining across all STEM majors misses important field-specific sex differences.

As seen in this figure, women have made significant gains in both categories of STEM at each of these levels over the past 40 years. By 2011, there was little difference in women's and men's advancement from baccalaureate to PhD and then to tenure-track assistant professorships – *in GEEMP fields only*. Thus, although far fewer women begin in GEEMP fields, of those who do, their progress resembles male GEEMP majors and in fact slightly exceeds males in transitioning from baccalaureate to PhD (Ceci et al., 2014, fig. 11). Recently, research showed that women with undergraduate engineering degrees persisted in the workforce 7–8 years post-baccalaureate degree nearly identically to males (Kahn & Ginther, 2015).

In contrast, in 2011, the probability of advancing from an LPS baccalaureate to a PhD was not as high for women as for men, nor as high as

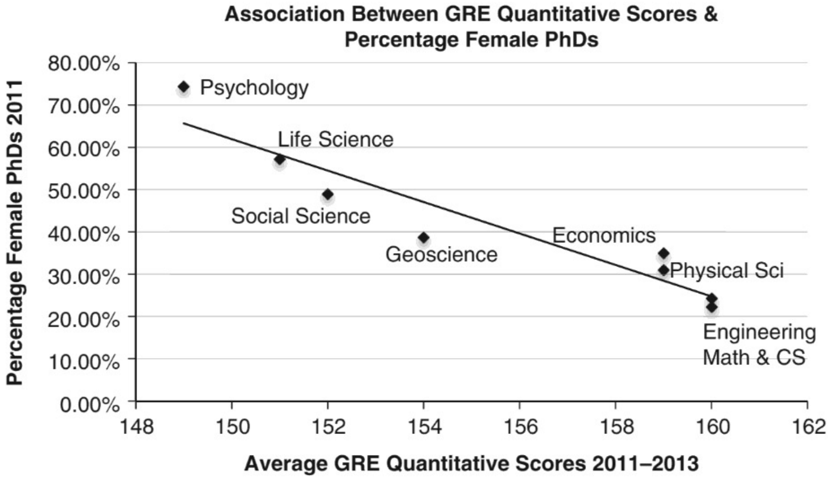

Figure 3.2 Association between average GRE-Quantitative scores and % PhDs awarded to females (CS refers to computer science).
Source: Ceci et al., 2014, data from ETS and Webcaspar.

For example, sex differences favoring men in spatial 3-D tasks do not translate directly to superiority in geometry, but depend on whether the dependent variable is grades in geometry or scores on standardized tests that cover materials not directly taught (Else-Quest, Hyde, & Linn, 2010; Lindberg et al., 2010). For that matter, sex differences in mathematics scores do not translate into grades in math classes, including complex math classes in college (Ceci et al., 2009): women obtain slightly higher grades in college classes and 40–48% of baccalaureates in mathematics for two decades (see Ceci et al., 2014, table A1A). None of this means that biological sex differences play no role in the shortage of women in GEEMP fields. But it does mean that care must be taken in touting them as *the* primary causal factor.

At the midpoint of the quantitative distribution, there are no systematic sex differences through middle school (see Ceci et al., 2014). Hyde and her colleagues have analyzed the sex gap in average mathematics ability, using large-scale national probability samples (Hyde, Fennema & Lamon, 1990; Hyde et al., 2008). They showed that mean scores highly overlap (*d*'s between 0.05–0.26 favoring males): Hyde and colleagues' 1990 meta-analysis of 100 studies found no significant sex differences for children at any age and for any type of mathematical problem – the only exception was a small male advantage, d = 0.29, for complex math problems for high school–aged students. Hyde and colleagues (2008) even found small female advantages for most years through ninth grade. However, they also found significant male advantages in grades 10 and 11.

By the early 2000s, average U.S. sex differences were small even on the most complex items, leading Hyde and Mertz (2009) to conclude: "effect sizes were found to average d = 0.07, a trivial difference. These findings provide further evidence that the average U.S. girl has now reached parity with the average boy, even in high school, and even for measures requiring complex problem solving" (p. 8802).

Some have found small differences on math tests earlier than middle school, but not at entrance to kindergarten. For instance, Fryer and Leavitt (2010) and Penner and Paret (2008), studying the same 1998–1999 kindergarten cohort, found no differences entering kindergarten but differences starting at tiny levels by the end of kindergarten that rose to 0.15–0.20 SDs by fifth grade. Cross-national studies found countries differ. For instance, Mullis and colleagues (2000b) reported no U.S. average sex difference on the fourth-grade TIMSS, but a male advantage for Korea and Japan.

In sum, there is agreement that in the United States, there are either nonexistent male advantages in average math scores, or very small ones relative to the overlap of the distributions, on the order of less than 0.001, or less than 0.1% SD.

Sex differences at the right tail. Perhaps the shortage of women in GEEMP fields is the result of sex differences in high math ability. As seen in Figure 3.3, most graduate students in GEEMP fields at one of our universities have GRE-Q scores in the top 18% (750), which is equivalent to the updated scale used earlier.

What is known about sex differences at the right tail? A male advantage in math ability is unreliable until early adolescence (Ceci et al., 2009). Lohman and Lakin (2009) analyzed more than 300,000 American 9- to 17-year-olds and found a higher proportion of boys in in the top 4% of the math distribution, which was stable across national samples from 1984 to 2000 (Figure 3.4). Strand and colleagues (2006) reported a similar male overrepresentation at the right tail of this test for more than 300,000 11-year-olds from the UK, with boys significantly more likely to score in the top group (+1.75 *SD*s above the mean); boys are also more likely to score in the bottom 4% in quantitative ability. Hedges and Nowell's (1995) analyses of six national data sets also showed consistency in the sex ratios at the top tail over a 32-year period.

Wai and colleagues (2010); Wai and Putallaz (2011), and Hyde and colleagues (2008) also reported substantial sex differences at the right tail

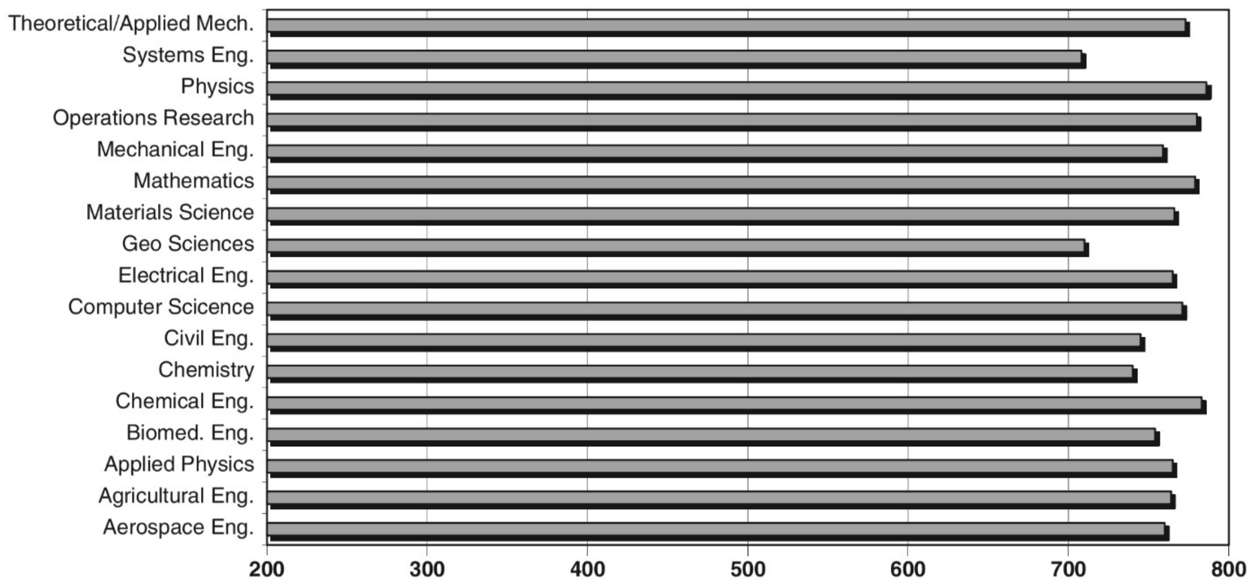

Figure 3.3 Mean GRE-Q scores of Cornell University GEEMP doctoral students, 2008–2009.

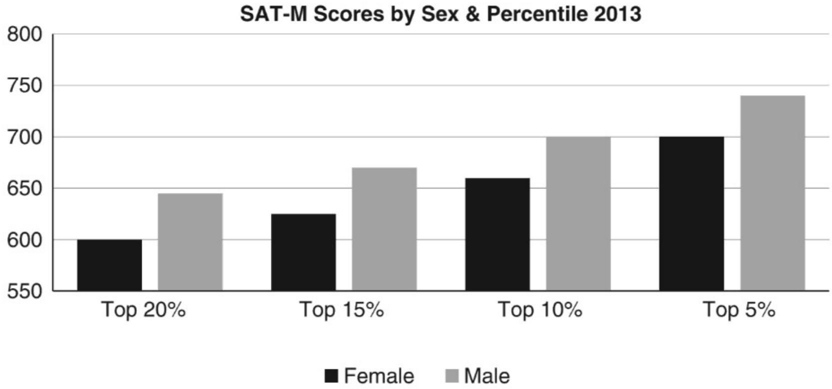

Figure 3.5 SAT-Math scores by sex and percentile, 2013.
Source: Ceci et al., 2014, data from College Board.

TIMSS at grade 8, but consistent male superiority by grade 12, particularly among the highest quartile of mathematics scorers. Relatedly, Stoet and Geary (2013) analyzed PISA data for the 33 countries that participated in all waves from 2000–2009. They, too, found slowly narrowing sex differences of 15-year-olds at the right tail, ending at 1.7:1 favoring males at the top 5% and 2.3:1 at the top 1%. Thus, large-scale analyses converge on the conclusion that there is a sizable male advantage at the right tail of the math distribution.

Sex differences favoring males on the SAT-Mathematics are similar: the same score that gets a girl into the top 5% of the female mathematics distribution gets a boy into only the top 10% of the male distribution; the same score that gets a girl into the top 10% gets a boy into only the top 20% (Figure 3.5).

There is inconsistency in sex differences at the right tail, even when comparing large national samples or meta-analyses (for review, see Ceci et al., 2009). For instance, among certain age children in Iceland, Singapore, and Indonesia, more girls scored in the top 1% than boys (Hyde & Mertz, 2009). Further, the male advantage at the right tail has been decreasing, more in some countries than in others, and the greater male variance in math scores is not always the case. In Lohman and Lakin's (2009) data, females narrowed the right tail gap on the Cognitive Abilities Test Nonverbal Battery: ninth stanine female-to-male ratios changed from 0.72 in 1984, to 0.83 in 1992, to 0.87 in 2000. Relatedly, the male-to-female ratio at the top 4% is larger in the United States (two to one) than it is in the UK (roughly three to two), further illustrating the influence of cultural factors.

Researchers have reported variations across ethnic groups in the United States. Hyde and Mertz (2009) found large differences favoring white males at the extreme right tail, but the opposite for Asian Americans, with more females at the right tail, and these differences varied by cohort. Miller and Halpern (2014), note that "sex differences in high mathematics test performance are reversed (female advantage) among Latino kindergarteners, indicating the early emerging effects of family and culture" (p. 39).

Finally, there are cross-state variations in the United States in the male/female ratio at the 95th percentile, with sex differences in some states less than half the size in others (Pope & Sydnor, 2010): the males-to-females ratio among the top 5% scorers in math/science is approximately 1.8 in the Eastern South Central states, but only 1.4 in the New England states. However, the ratio of *females-to-males* NAEP 95th percentile scores in reading is approximately 2.1 in the New England states and 2.6 in the East South Central states. States with more gender-equal math and science scores also have more gender-equal reading scores at the right tail (which otherwise has more girls), suggesting gender norms strongly influence mathematic and verbal achievement at the top tail. Along these lines, Ellison and Swanson (2010) found that local school culture was highly influential in determining how many girls competed at the highest level of mathematics in national competitions.

Boys are overrepresented in both tails of the distribution. Transnational mathematics analyses (TIMMS, PISA) show boys' higher variance ratios (VRs) – the male variance divided by the female variance (e.g., Else-Quest et al., 2010; Penner, 2008). Else-Quest and colleagues report VRs ~1:1.19 in the United States, 1:1.06 in the UK, 1:0.99 in Denmark, and 1:0.95 in Indonesia. In representative studies, VRs average 1.15, and on average there is at least a 2-to-1 ratio favoring males among the top 1% of math scorers. These sex differences are real. However, transnational and trans-state differences suggest that something more than mathematical potential is driving the higher male variability. Yet VRs may underestimate population variance because more males are developmentally delayed and not included in assessments (see Halpern, 2012). Thus, state-by-state, transnational, cohort, and ethnic data all indicate that sex differences at the right tail are fluid; these ratios can and do change.

Moreover, data sources used in these analyses are vulnerable to variations in context, (e.g., changes over time in test content that favor one sex,

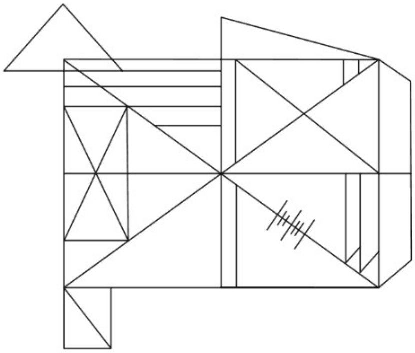

Figure 3.7a Spatial task used by Huguet and Regner (2009); can be framed as geometry or art.

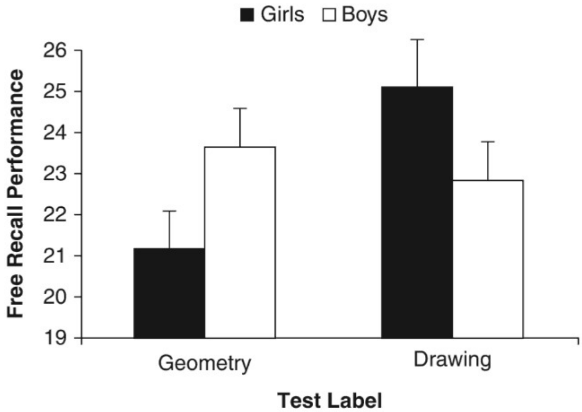

Figure 3.7b Shows the interaction between framing and sex differences.

is framed as an art task. Huguet and Regner (2009) presented Figure 3.7a to middle school students for 90 seconds and gave them five minutes to reproduce it from memory.

Boys outperformed girls when the task was presented as a geometry problem, but the reverse was true when it was presented as an artistic task (Figure 3.7b).

Various interventions to teach spatial processing demonstrate that the sex gap can be narrowed, though not fully closed within the confines of the training durations, which have been one semester or less (Ceci et al., 2009).

If boys' spatial superiority is due to playing dynamic video games, Lincoln logs, erector sets, Legos, etc., then exposing girls to these activities could narrow the gap. Some research shows that spatial activities between the ages of two and four years (e.g., shaping clay, drawing, and cutting 2-D figures) predicts mathematical skills at age four and a half (Grissmer et al., 2013). They showed that an intervention based on transforming spatial materials to 2-D and 3-D elevated disadvantaged children on the Woodcock-Johnson Applied Problems and KEYMATH3-Numeration tests from ~ 32nd to the 48th percentile; visual-spatial ability was also elevated due to play activity, from 33% to 47%.

There is some evidence, however, of sex differences in spatial processing prior to the onset of activities. Four studies have shown that male infants outperform females on rotation tasks. However, this depends on whether they are speeded and/or entail rigid surface transformations (see Miller & Halpern, 2014). And there is some evidence that crawling experience fosters spatial ability. Moore and Johnson (2011) employed a habituation paradigm using the spatial task in Figure 3.6. Following habituation to an object, when infants are shown it in a new perspective, three-month-old boys prefer the novel display over the rotated version of the familiar display, whereas girls look at the familiar and novel objects equally, indicating that only the boys mentally rotated objects. This suggests that boys' spatial intelligence is evident somewhat earlier than girls'. (Quinn and Liben, 2008, found a similar result for three–four-month-olds with 2-D rotation though, as noted, sex differences are primarily found for 3-D rotations.) Although these studies strongly suggest an early biological basis of early sex differences, some argue they cannot rule out environmental causes (see Miller & Halpern, 2014, p. 39).

As noted, 3-D mental rotation is linked to seemingly small differences favoring infants who engage in early crawling and manual manipulation. Researchers presented nine-month-olds with a 3-D rotation task; half had been crawling for nine weeks, and some were likely to manipulate five objects presented to them. The infants were habituated to a video of an object rotating back and forth through a 240° angle around its longitudinal axis. When tested with the same object rotating through the unseen 120° angle, the crawlers focused longer at the novel (mirror) object, regardless of their manual manipulation scores. In contrast, the non-crawlers' rotation was influenced by their manual manipulation (Schwarzer, Freitag, & Schum, 2013). These findings indicate that subtle environmental differences, such as early crawling and object manipulation, influence spatial cognition. Of course this does not preclude a biological role in male spatial superiority (perhaps early crawling is biologically determined and occurs earlier for male babies), but it suggests an intervention to induce infant girls to manipulate and crawl.

A Seeming Paradox

There is a seeming paradox: females outperform males on classroom mathematics achievement. Yet males are more numerous at the right tail of mathematics performance on standardized tests such as the SAT-Math, the NAEP, PISA, and GRE-Q. Males are also more likely to major in GEEMP disciplines, obtain GEEMP masters and PhDs, and work in GEEMP fields.

Explanations for this paradox may lie in gender stereotypes that associate math with boys and reading with girls. Cvencek, Meltzoff, and Greenwald (2011) report that by second grade, boys and girls demonstrated implicit and explicit stereotypes associating math with maleness and reading with femaleness.

Paradoxically, these stereotypes are incongruent with what children observe in classrooms, including math and science. Part of the paradox may be related to girls associating in early grades brilliance with being male. This stereotype has a sudden onset, sometime between kindergarten and second grade. Bian, Leslie, and Cimpian (2017) demonstrated that children's sense of what it means to be brilliant changes between five and seven. In their study, children had to guess which of two boys and two girls was "really, really smart" and most likely to solve a hard problem. Five-year-old boys *and* girls associated brilliance with their own gender. Yet between five and seven, girls become less likely than boys to associate brilliance with their gender. This also extended to children rating adults, they begin to exhibit a bias in favor of males between six and seven. However, there is no bias in favor of males in predicting who had better academic achievement; there was an expectation that girls will get better grades, consistent with the actual data showing girls do get better grades. Finally, additional experiments in Bian and colleagues (2017) found that girls were less interested than boys in a game they associated with smart children: "Many children assimilate the idea that brilliance is a male quality at a young age. This stereotype begins to shape children's interest as soon as it is acquired and is thus likely to narrow the range of careers they will one day contemplate" (Bian et al., 2017, pp. 390–391).

There is some evidence that a belief that math ability can be developed is self-fulfilling, and that girls are less likely to have this so-called growth until later in high school and they do not predict the later specific gender segregation observed in college majors (Legewie & DiPrete, 2012), they nevertheless reveal early career leanings.

However, several enigmas remain. Even if sex differences in math and science orientation/identification begin as early as ages six and seven and solidify by the end of middle school, it is unclear why they should result in the particular gendered pattern of career aspirations observed. Among high school students, sex differences in STEM courses and plans to major in STEM fields are well-established and demonstrated by the lower participation rates (23% to 42% female) in AP exams such as Calculus BC, computer science, and Physics C, and between 20% and 60% more males receive top scores of 5 (Ceci et al., 2014). Yet, even among those women who escape stereotypical influences, who take and excel in math-intensive advanced coursework in high school and college (where almost half of baccalaureates in mathematics are awarded to women), we still see far fewer women entering GEEMP careers; instead, they choose careers in microbiology, medicine, or statistics, fields in which women have achieved significant presence.

Thus, research converges on the following three conclusions. a) Stereotypes are important: girls/boys learn them early, although it does not translate into lower average math ability in testing until puberty. b) There is higher variance and more representation at the right tail for males. This appears to be mutable to some degree, although we do not know fully whether this can be eliminated. Yet this alone is not enough to explain the difference in GEEMP representation. And c) in addition to stereotypes and math ability differences (that may or may not be due to biology), there are gender differences in interests, with females more interested in people-related careers and males more interested in nonsocial things (e.g. Auyeung, Lombardo, & Baron-Cohen, 2013; Thorndike, 1911). Lippa has repeatedly documented very large sex differences in occupational interests, including in transnational surveys, with men more interested in "thing"-oriented activities and occupations, such as engineering and mechanics, and women more interested in people-oriented occupations, such as nursing, counseling, and elementary school teaching (e.g., Lippa, 1998, 2001, 2010). In an extensive meta-analysis of more than half a million people, Su, Rounds, and Armstrong (2009) reported a sex difference on this dimension of a full standard deviation (see also Su & Rounds, 2015). However, the extent to which these gendered interests and outcomes are influenced by early biases and stereotypes remains to be demonstrated.

Else-Quest, N., Hyde, J., & Linn, M. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136, 103–127.

Fryer, R. G., Jr. & S. D. Levitt. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics,* 2(2), April, 210–240. www.jstor.org/stable/25760212.

Good, C., Aronson, J., & Harder J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology,* 29, 17–28.

Grissmer, D. W., Mashburn, A., Cottone, E., Chen, W.-B., Brock, L. L., Murrah, W. M., … Cameron, C. E. (2013). Effects of a play-based after-school curriculum for high risk K–1 children. In N. Newcombe (Ed.), *Educating spatial skills at varied ages with varied approaches: Are STEM outcomes affected?* Seattle, WA: Society for Research in Child Development.

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th edn.). New York: Psychology Press.

Hedges, L. V. & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science,* 269, 41–45.

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics.* Washington, DC: American Association of University Women.

Huguet, P. & Regner, I. (2009). Counter-stereotypic beliefs in math do not protect school girls from stereotype threat. *Journal of Experimental Social Psychology*, 45, 1024–1027.

Hyde, J. S., Fennema, E., & Lamon, S. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139–155.

Hyde, J. S., Lindberg, S., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, 321, 494–495.

Hyde, J. S. & Mertz, J. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences, USA*, 106, 8801–8809.

Kahn, S. & Ginther, D. K. (2015). Are recent cohorts of women with engineering bachelors less likely to stay in engineering? *Frontiers in Psychology,* https://doi.org/10.3389/fpsyg.2015.01144.

Lavy, V. & Sand, E. (2015). On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases. NBER Working Paper 20909. www.nber.org/papers/w20909.pdf.

Legewie, J. & DiPrete, T. A. (2012). High school environments, STEM orientations, and the gender gap in science and engineering degrees. http://ssrn.com/abstract=2008733.

Lindberg, S. M. et al. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psychological Bulletin,* 136, 1123–1135.

Linn, M. C. & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, 56, 1479–1498.