

What Do Undergraduates Learn About Human Intelligence? An Analysis of Introductory Psychology Textbooks

Russell T. Warne, Mayson C. Astle, and Jessica C. Hill
Utah Valley University

ABSTRACT

Human intelligence is an important construct in psychology, with far-reaching implications, providing insights into fields as diverse as neurology, international development, and sociology. Additionally, IQ scores can predict life outcomes in health, education, work, and socioeconomic status. Yet, students of psychology are often exposed to human intelligence only in limited ways. To ascertain what psychology students typically learn about intelligence, we analyzed the content of 29 of the most popular introductory psychology textbooks to learn (a) the most frequently taught topics related to human intelligence, (b) the accuracy of information about human intelligence, and (c) the presence of logical fallacies about intelligence research. We found that 79.3% of textbooks contained inaccurate statements and 79.3% had logical fallacies in their sections about intelligence. The five most commonly taught topics were IQ (93.1% of books), Gardner's multiple intelligences (93.1%), Spearman's *g* (93.1%), Sternberg's triarchic theory (89.7%), and how intelligence is measured (82.8%). Conversely, modern models of intelligence were only discussed in 24.1% of books, with only one book discussing the Carroll three-stratum model by name and no book discussing bifactor models of intelligence. We conclude that most introductory psychology students are exposed to some inaccurate information and may have the mistaken impression that nonmainstream theories (e.g., Sternberg's or Gardner's theories) are as empirically supported as *g* theory. This has important implications for the undergraduate curriculum and textbook authors. Readers should be aware of the limitations of the study, including the choice of standards for accuracy for the study and the inherent subjectivity required for some of the data collection process.

SCIENTIFIC ABSTRACT

Human intelligence is an important concept in psychology because it provides insights into many areas, including neurology, sociology, and health. Additionally, IQ scores can predict life outcomes in health, education, work, and socioeconomic status. Yet, most students of psychology do not have an opportunity to take a class on intelligence. To learn what psychology students typically learn about intelligence, we analyzed 29 textbooks for introductory psychology courses. We found that over 3/4 of textbooks contained inaccurate statements. The five most commonly taught topics were IQ (93.1% of books), Gardner's multiple intelligences (93.1%), Spearman's *g* (93.1%), Sternberg's triarchic theory (89.7%),

This article was published February 26, 2018.

Russell T. Warne, Mayson C. Astle, and Jessica C. Hill, Department of Behavioral Science, Utah Valley University.

This research was previously presented at the Utah Conference on Undergraduate Research on February 17, 2017, in Orem, UT; the National Conference on Undergraduate Research on April 7, 2017, in Memphis, TN; the annual conference of the Rocky Mountain Psychological Association on April 7, 2017, in Salt Lake City, UT; the annual conference of the International Society for Intelligence Research on July 15, 2017, in Montreal, Canada; and the annual conference of the American Psychological Association on August 3, 2017, in Washington, DC.

The authors have made available for use by others the data that underlie the analyses presented in this paper (see Warne, 2017), thus allowing replication and potential extensions of this work by qualified researchers. Next users are obligated to involve the data originators in their publication plans, if the originators so desire.

This project was financially supported by a Grant for Engaged Learning and an Undergraduate Research Scholarly and Creative Activities Grant from Utah Valley University.

Copyright of this manuscript belongs to the author(s). The author(s) grant(s) the American Psychological Association the exclusive right to publish this manuscript first, identify itself as the original publisher, and claim all commercial exploitation rights. Upon publication, the manuscript is available to the public to copy, distribute, or display under a Creative Commons Attribution-Noncommercial 3.0 Unported License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits use, distribution, and reproduction in any medium, provided that the original work is properly cited and is not used for commercial purposes. Please use APA's Online Permissions Process (Rightslink®) at <http://www.apa.org/about/contact/copyright/seek-permission.aspx> to request commercial reuse of this content.

Correspondence concerning this article should be addressed to Russell T. Warne, Department of Behavioral Science, Utah Valley University, 800 West University Parkway MC 115, Orem, UT 84058. E-mail: rwarne@uvu.edu

and how intelligence is measured (82.8%). We learned that most introductory psychology students are exposed to some inaccurate information about intelligence and may have the mistaken impression that nonmainstream theories (e.g., Sternberg's or Gardner's theories) are as empirically supported mainstream theories (such as Spearman's g).

Keywords: human intelligence, psychology textbooks, undergraduate education

Supplemental materials: <http://dx.doi.org/10.1037/arc0000038.supp>

Data repository: <http://dx.doi.org/10.3886/ICPSR36957.v1>

Psychology is one of the most popular majors in America (Halonen, 2011). The National Center for Education Statistics reports that out of 1,894,934 bachelor's degrees awarded in the 2014–2015 academic year, 117,557 were in psychology (National Center for Education Statistics, 2016, Table 322.30). The only two degree areas with more bachelor's degrees awarded were business and health-related professions.

Like the psychology major, introductory psychology courses have broad appeal, which is a unique opportunity for the field to have an educational impact across the university curriculum. These courses provide exposure to both the natural sciences (e.g., neuroanatomy, brain function) as well as the social sciences (e.g., personality, social), due to psychology's philosophical and biological origins (Barnard et al., 1970; Halonen, 2011). Thus, the courses are popular general education electives for nonpsychology majors, such as those studying medicine, business, engineering, computer science, teaching, communications, religion, and the law (Barnard et al., 1970; Halonen, 2011). Consequently, the introductory psychology course is the most frequently offered psychology course at universities across the nation (Stoloff et al., 2010), enrolling between 1.2 and 1.6 million students annually (Steuer & Ham, 2008, p. 160). Indeed, more than 99% of institutions of higher education in America offer introductory psychology (Stoloff et al., 2010). Introductory psychology courses also serve as “gateway courses” to later classes in the psychology major (Hogben & Waterman, 1997), as introductory psychology is often a prerequisite more advanced courses (Stoloff et al., 2010).

Because psychology has tremendous diversity in content, almost every university offers an introductory psychology course in which students are briefly introduced to many topics (American Psychological Association [APA], 2014; Stoloff et al., 2010) via course lecture and textbooks. To increase consistency across institutions in the content offered, the APA formed an introductory psychology Working Group to provide recommendations on critical content. The group produced five pillars that should be the foundation of any introductory course: (a) biological, (b) cognitive, (c) development, (d) social and personality, and (e) mental and physical health (APA, 2014, p. 17). Among the topics covered in the “social and personality pillar” are gender, emotion, and human intelligence.

Intelligence Is Important, but Neglected

Intelligence has been subjected to more than a century of empirical scrutiny (Detterman, 2014; Warne, 2016). One of the results today is a hierarchical theory of intelligence called the Carroll (1993) three-stratum theory, which states that all cognitive abilities are organized in a hierarchy, with specific, narrow tasks (e.g., vocabulary knowledge, arithmetic skills, visual memory) at the lowest level (see Figure 1). These tasks are subsumed by a smaller number of abilities that are broader in applicability (e.g., verbal ability, mathematical reasoning, short-term memory (STM)). At the top of the hierarchy is general intelligence, or g ,¹ which is the broadest mental ability and seems to be used in every cognitive task. Carroll's theory had the virtue of informing a variety of controversies related to cognitive abilities in the

20th century. For example, Carroll recognized the importance of g as a general ability, while also providing a space for other abilities to make a contribution to human cognition (Warne, 2016). A related model is the bifactor model, which posits that observed variables are the product of g and of broad abilities (see Figure 2). The bifactor model is mathematically related to the Carroll three-stratum theory, with the bifactor model being a generalization of Carroll's work and in accordance with Carroll's beliefs about how g influences performance on specific tasks (Beaujean, 2015; Yung, Thissen, & McLeod, 1999). In recent years the bifactor model has attracted advocates (e.g., Canivez, 2016; Cucina & Howardson, 2017; Frisby & Beaujean, 2015), mostly because the bifactor model tends to fit the data from cognitive test batteries better than competing models do (Cucina & Byle, 2017).

One contrasting hierarchical model is the Cattell-Horn-Carroll (CHC) theory (McGrew, 2009), which posits that the influence of g onto specific tasks is fully mediated by broad, midlevel cognitive abilities. The CHC theory denies any direct influence of g onto specific tasks, thus favoring a structure of mental abilities shown in Figure 1. The CHC theory forms the basis of many professionally developed cognitive test batteries, though often these same test batteries produce data that support the bifactor/Carroll three-stratum theory more (e.g., Cucina & Byle, 2017; Cucina & Howardson, 2017), mostly because the CHC model requires strict assumptions about the relative contribution of variance of each task to its factor (Gignac, 2016).

Regardless of one's preferred theoretical model, intelligence has wide-ranging implications in real-world settings. It has strong correlations with extensive variables such as income (A. R. Jensen, 1998), job prestige (Nyborg & Jensen, 2001), life expectancy (Deary, Whiteman, Starr, Whalley, & Fox, 2004), and job performance (Schmidt & Hunter, 2004). Conversely, intelligence is negatively correlated with criminal behavior (Beaver et al., 2013), long-term unemployment (Herrnstein & Murray, 1996), dementia (Deary et al., 2004), death by automobile accident (O'Toole & Stankov, 1992), and more. Exposure to the wide spectrum of human ability is beneficial to students of all backgrounds; Detterman (2014) stated that, “high ability students believe that everyone is like them. They are often shocked when told about the full range of ability and even more shocked when they encounter it in the real world” (p. 148). For example, the critical nature of understanding intelligence differences can be seen in recent exchanges between the police and those with decreased intellectual capacity, such as individuals with severe autism (e.g., Karimi, 2016). Students of criminal justice, who are more likely to have higher

¹ There is disagreement among experts about whether g is synonymous with general intelligence. Arthur Jensen was careful to distinguish the two, whereas Carroll (1993, pp. 591–599) thought they were synonymous. Many other experts fall on either side of the debate, and consensus on the issue is elusive. In this article we take the position that g is either synonymous with general intelligence, or that the two concepts are very much alike. However, we acknowledge that experts have good reasons to disagree with us. Readers who desire a comprehensive argument for why g and intelligence are not the same, see A. R. Jensen (1998, Chapter 3).

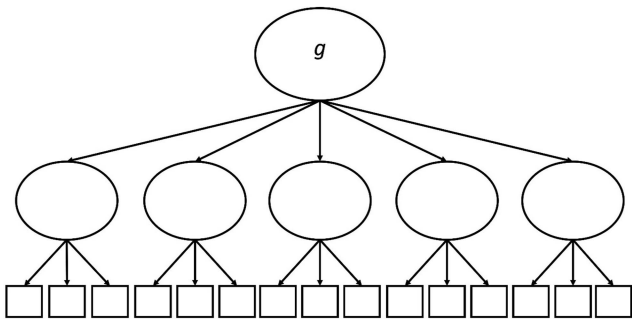


Figure 1. A representation of a hierarchical model of intelligence. The model depicts a hierarchy of cognitive abilities. At the bottom are specific, narrow abilities. Highly correlated groups of these specific abilities (represented as rectangles) coalesce into a small number of abilities that have broader impact and are represented in the middle row of ovals. These broad abilities, in turn, are all related via the general intelligence factor (labeled *g*) at the top of the hierarchy. Although many intelligence researchers subscribe to this model, the exact number of abilities in the middle and lower levels is a subject of much debate (McGrew, 2009).

intellectual ability, would be more fully prepared for their careers through encountering people with intellectual ability, who are a disproportionate share of individuals who are arrested (Beaver et al., 2013).

In contrast to the importance of intelligence and the strength of the evidence about the construct, psychology education has largely neglected the concept. In an investigation of psychology course offerings at American universities, intelligence didn't even make the list of psychology courses provided in at least 10% of universities (Stoloff et al., 2010). Earlier studies of the psychology curriculum show similar results (e.g., Perlman & McCann, 1999), and a course on intelligence has never been part of the mainstream undergraduate psychology curriculum (McGovern, 1992). Therefore, the little that college students learn about intelligence occurs in classes that mostly focus on other topics, such as cognitive or developmental psychology.

The one exception to the lack of exposure to human intelligence is the introductory psychology course. This is a course that nearly every undergraduate psychology student is required to take (Stoloff et al., 2010) and that many nonpsychology students choose to take as a general education elective (Barnard et al., 1970). To discover the depth and breadth of undergraduate experience with intelligence in introductory psychology courses, we conducted a study of introductory psychology textbooks.

Analysis of Psychology Textbooks

Textbook Coverage

To learn more about the undergraduate psychology curriculum researchers often analyze textbooks. For example, in Griggs and Marek's (2001) study of 37 introductory psychology textbooks, they found that even though general chapter topics were largely similar (e.g., abnormal, cognition), the content within the chapters showed a great degree of variation. Further, there is a wide degree of variation among texts' most cited authors and journal articles. Gorenflo and McConnell (1991) found that in the list of 37,590 citations, several prominent psychologists were missing—including Skinner, Freud, and Piaget. However, a follow-up study by Griggs, Proctor, and Cook (2004), found that books by these authors were cited in textbooks, even though their journal articles were rarely cited. This demonstrates that "there is no substantial common core either in the language used

by psychology text authors or in the psychologists cited and journal articles referenced in these textbooks" (Griggs, Proctor, & Bujak-Johnson, 2002, p. 452).

Textbook Accuracy

The introductory psychology textbook is difficult to produce with uniform accuracy, as authors have only a limited area of expertise, yet must write chapters that discuss the entire breadth of psychology. As a result, authors can unintentionally include common misconceptions or inaccurate findings. Ferguson, Brown, and Torres (in press) found that when certain issues (like the idea that humans use only 10% of their brain, the story of the murder of Kitty Genovese, and the influence of violent media on later violent behavior) were presented, textbook authors tended to discuss them in oversimplified terms that ignored controversies, often avoided discussing weaknesses in the research, or perpetuated misconceptions. Such an approach to summarizing research in introductory texts is likely due to the authors' need to aim their writing at a student audience; as a result, they present "contested research as more consistent, generalizable to socially relevant phenomena and higher quality than it was" (Ferguson et al., in press, p. 6).

Ferguson et al.'s (in press) results, unfortunately, are not an isolated finding. For instance, Skinner's seminal work on operant conditioning and his philosophy of radical behaviorism is often presented in texts as being mutually exclusive from cognition, and textbook authors sometimes make the claim that Skinner was not interested in inner behavior (R. Jensen & Burgess, 1997). Rather, Skinner proposed that the contents of "private events" (e.g., thoughts, emotions) should not be privileged beyond overt behavior (Skinner, 1953, p. 257). Regrettably, among 15 textbooks, there was not a single full and accurate account of Skinner's views (R. Jensen & Burgess, 1997). Moreover, some textbook authors even avoided any connection between behaviorism and cognition at all.

Although Ferguson et al.'s (in press) work could be discounted as being attributable to differences in interpretation, it is regrettably common for textbooks to include errors of fact—particularly in first editions. Although many of first edition errors are corrected in later editions, some errors persist across many editions, resulting in the perpetuation of misquotations, errors of fact, omissions, and occasionally factual fabrications (Habarth, Hansell, & Grove, 2011; Thomas, 2007). Although many of these errors are small, it is possible for large errors to survive the review process. For example, descriptions of

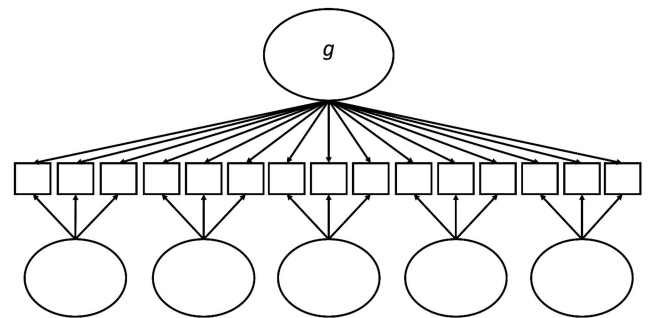


Figure 2. A representation of the bifactor model of intelligence. The model depicts each specific ability (represented as rectangles) are the product of general intelligence (labeled *g*) and the broad abilities (shown at the bottom of the figure). Like the hierarchical model shown in Figure 1, the exact number of specific and broad abilities is the subject of debate. When compared with hierarchical models, the bifactor model tends to fit the data better (Cucina & Byle, 2017).

applied psychology areas (e.g., industrial/organizational, clinical, counseling, school) are often subject to inaccuracy in textbook descriptions (Haselhuhn & Clopton, 2008).

How authors approach writing is directly related to the quantity and types of inaccuracies. Steuer and Ham (2008) randomly selected portions of introductory psychology textbooks and compared the textbook authors' explanations and interpretations with the text's references. They discovered an assortment of errors, ranging from minor citation errors, to misrepresenting the content of a source, to plagiarism. When authors engaged in inductive referencing—by becoming familiar with the literature and writing about what they learned—texts were more accurate than when authors engaged in deductive referencing, which occurs when writers start with a preconceived understanding of a topic and then search for literature supporting their views. Deductive referencing is often more error-prone, as the writing process “becomes more a matter of defending [viewpoints] than of discovering statements about scientific truth” (Steuer & Ham, 2008, p. 163).

Psychology Textbooks and Intelligence

There have been three prior studies about intelligence and psychology textbooks. In one study, Griggs (2014a) analyzed textbook coverage and course syllabi, finding that discussions on intelligence were a smaller percentage of textbook space in the 21st century than the 1980s, dropping from 6% of textbook space to 4%. Previously intelligence was covered predominantly in its own chapter, whereas in 21st century textbooks it was often combined with the language and thought sections of the book. In another study, Jackson and Griggs (2013) found that modern introductory psychology textbooks devoted 1–4% (median: 4%) of space to discussing intelligence. Although this information about the percentage of a textbook dedicated to intelligence is worthwhile information, it says nothing about accuracy of textbook information on intelligence or the topics that textbook authors introduce when discussing intelligence. To understand better what undergraduates learn about human intelligence, more research is required.

In a more detailed study of organizational psychology textbooks (Pesta, McDaniel, Poznanski, & DeGroot, 2015), intelligence was discussed in an average of 3.89 paragraphs, despite the fact that intelligence is one of the most powerful predictors of job performance, especially in more complex jobs (Schmidt & Hunter, 1998, 2004). In comparison, these organizational behavior textbooks discussed emotional intelligence—a construct with much less empirical evidence for its existence and/or utility—in almost twice as many paragraphs. Pesta et al. (2015) also found that discussions of intelligence were much less accurate than discussions of emotional intelligence.

Purpose of Current Study

Given the importance of human intelligence and the limited literature related to its inclusion in the undergraduate psychology curriculum, we investigated the presentation of intelligence in the most frequently used introductory psychology textbooks in the United States. Following the procedures of other researchers who have investigated the accuracy of psychology textbooks (e.g., Ferguson et al., in press; Habarth et al., 2011; Steuer & Ham, 2008), we conducted a study to investigate the quality of the discussion on intelligence in the most popular psychology textbooks today. Specifically, we had two research questions:

- Research Question 1: What are the most frequently discussed topics related to intelligence in introductory psychology textbooks?
- Research Question 2: How accurate are introductory psychology textbooks in their discussion of intelligence?

Because so few undergraduate students ever take a course on intelligence and the instructor's choice of textbook often dictates the content of an introductory class (Miller & Gentile, 1998), we believe that this study will provide a realistic snapshot of what undergraduates learn about intelligence.

Method

Following Habarth et al. (2011), we decided to sample the most popular introductory textbooks, based on sales of new textbooks. In August 2016 we requested the 30 most popular introductory psychology textbooks according to rankings publicly available on amazon.com. We received 29 of these books (listed in Table 1). When two versions of a book were available, we always chose the full version rather than the shorter, abridged version to follow the same procedure as Habarth et al. (2011) and because these shorter versions contained no unique information (Jackson & Griggs, 2013). A few of the textbook authors are among the most influential living psychologists (Diener, Oishi, & Park, 2014). The books represent eight different publishers, including all of the major social science textbook publishing companies. Cengage had the largest number of textbooks in the sample (10), with Pearson (6), Worth (5), McGraw-Hill (3), Norton (2), Kendall Hunt (1), Oxford University Press (1), and Wiley (1) also having textbooks in the sample. Smaller textbook publishers are absent from our sample, as are open source textbooks. However, because our purposive sampling method focused on the most popular introductory psychology textbooks and several publishers are represented, we believe that the books would provide accurate information about what many—perhaps most—introductory psychology students would learn about intelligence.

Gathering Data: Research Question 1

We recorded basic information about the section on intelligence in each book. This information consisted of whether intelligence had its own chapter in the text (or was a section of a more comprehensive chapter, such as a chapter on cognition), the number of pages devoted to intelligence, and the total number of pages in the textbook (not including references, indices, or glossaries).

To investigate Research Question 1, the first author created a coding system before the study began. Based on prior studies of textbook content (e.g., Griggs & Marek, 2001; Griggs & Mitchell, 2002; Pesta et al., 2015; Zechmeister & Zechmeister, 2000), the first author chose to code every textbook's section headings, emphasized vocabulary terms (e.g., bolded vocabulary words), and topics discussed in relationship with intelligence. Headings and emphasized vocabulary are easy to find in textbooks and are unambiguous. To ensure that the coding of topics was as objective as possible, the first author decided *a priori* that topics would be coded at the paragraph level because the point at which paragraphs begin or end is always clear. If a topic was not discussed for at least one full paragraph, it was not coded. The first author then trained the second author (an accomplished undergraduate student and veteran of the first author's class on human intelligence) in this coding scheme by coding an entire textbook chapter together. The second author then used the system to collect data about topics that textbook authors discussed in their books.

From this coding process we produced a comprehensive list of topics discussed in each textbook and also a count of the number of

Table 1
Descriptive Information for Introductory Psychology Textbooks

Citation	Full chapter on intelligence?	# of pages on intelligence/ total # of pages	# of inaccuracies	# of statements of questionable accuracy	Fallacies
Bernstein (2016)	Yes	28/690	1	2	2, 10
Bonds-Raacke (2014)	No	8/289	None	1	None
Cacioppo & Freberg (2016)	No	12/662	1	3	10, 11
Ciccarelli & White (2015)	No	16/611	3	4	10, 11
Comer & Gould (2013)	Yes	36/644	3	7	2, 3, 4
Coon & Mitterer (2016)	Yes	21/593	8	6	2, 3, 4, 13
Feist & Rosenberg (2015)	No	21/619	None	5	3
Feldman (2015)	No	14/556	5	8	1, 12
Gerrig (2013)	Yes	18/479	1	3	6, 11, 12
Gleitman, Gross, & Reisberg (2011)	Yes	34/713	1	6	1, 2, 6, 10
Gray & Bjorklund (2014)	No	19/693	1	3	1, 6
Griggs (2014b)	No	12/438	None	1	None
Grison, Heatherton, & Gazzaniga (2017)	No	17/557	1	4	6, 12
Hockenbury, Nolan, & Hockenbury (2015)	No	18/659	3	6	2, 4
Kalat (2017)	Yes	18/525	None	6	11
Lahey (2012)	No	15/595	1	2	2, 7, 8
Lilienfeld, Lynn, Namy, & Woolf (2014)	Yes	37/663	1	4	4, 13
Morris & Maisto (2016)	No	19/516	1	4	None
Myers & Dewall (2015)	Yes	20/691	2	8	4, 6
Nairne (2014)	Yes	29/536	1	6	2
Nevid (2015)	No	11/552	1	5	3
Nolen-Hoeksema, Fredrickson, Loftus, & Lutz (2014)	Yes	19/639	1	1	None
Okami (2014)	No	12/769	None	3	None
Pastorino & Doyle-Portillo (2016)	No	13/631	1	3	3, 5, 12
Rathus (2016)	No	13/376	2	2	3, 4
Schacter, Gilbert, Wegner, & Nock (2014)	Yes	26/663	1	6	1, 2, 3
Wade, Tavris, & Garry (2014)	No	12/615	None	4	1
Weiten (2017)	No	17/565	2	6	None
Zimbardo, Johnson, & McCann (2017)	No	20/581	1	10	3, 6, 12

paragraphs that each topic was discussed in. We discovered after data collection began that some topics could overlap when the topic of a paragraph was a subset of a broader topic (e.g., a paragraph about twin studies would be recorded as discussing “twin studies” and “heritability research”). Early in the coding process we chose to label these paragraphs as discussing multiple topics, and we ensured that this decision was applied uniformly to all data.

To ensure accuracy in the coding process, the first author trained an additional coder (another undergraduate student), who also conducted the same coding process independently of the first and second author on five randomly chosen textbooks. The two coders had 100% agreement for the number of pages that a textbook discussed intelligence, the total number of textbook pages, and whether the textbook devoted an entire chapter to intelligence. Additionally, the two coders were in agreement for 89.2% of vocabulary terms, 99.2% of section headings, and 85.1% of topics discussed. The first author examined every discrepancy between the two coders, and where there were discrepancies, the primary coder (i.e., the second author) was almost always the more accurate, detailed coder. Based on these results, the research team decided that the primary coder was sufficiently accurate to continue using her data without further reliability checks. Percentages of agreement for all five books are available in Supplemental File 1 (p. 23).

After coding, the first two authors compiled a detailed list of every section heading, vocabulary term, and topic in the textbooks. The first two authors and three undergraduate research assistants then classified each item in the comprehensive list into categories. After the categories were created, the first two authors and the three research assistants reexamined each category and ensured that the items within each category were truly the same. When there was doubt, the group consulted the original passages in the textbooks and reached a con-

sensus after a discussion. In cases of ambiguity we opted to maximize the number of categories so that the results would be more detailed. It is important to note that the decision to maximize the number of categories in the textbooks is an arbitrary one that could make textbooks seem to be less comprehensive than they really are. Therefore, we also performed an alternative classification analysis in which we attempted to minimize the number of categories. Both results are reported for greater transparency.

Gathering Data: Research Question 2

The first author also read the section on intelligence in every textbook and coded for (a) factual inaccuracies, (b) statements of questionable accuracy, and (c) logical fallacies that inhibit understanding of intelligence research. Before the study began we chose two standards for factual accuracy: Gottfredson’s (1997a) mainstream statement on intelligence and Neisser et al.’s (1996) summary of intelligence research. Gottfredson’s (1997a) article is a statement signed by more than 50 scholars from diverse fields related to intelligence research (e.g., psychometrics, behavioral genetics, cognitive psychology, education). Neisser et al.’s (1996) article is the official report of an APA committee to produce a summary of intelligence research. We used these articles for two reasons. First, both represent a summary of solid, noncontroversial findings in intelligence research. Second, both articles are widely cited and old enough to be commonly known.

The first author compared every statement in the textbooks with the content of the Gottfredson (1997a) and Neisser et al. (1996) articles. If a statement in a textbook contradicted one or both of these articles, then he labeled it as a factual inaccuracy. In an effort to reduce subjectivity, we decided in advance that the inaccuracy must be

explicit and in direct contradiction to either article. Inaccurate statements were noted (including a reference) and compiled so that the number of inaccurate statements per book could be calculated.

Because Gottfredson (1997a) and Neisser et al. (1996) are not fully comprehensive literature reviews of the entire field of intelligence research, we believed that it would be necessary to also identify problematic statements in a textbook which did not directly contradict anything in the Gottfredson (1997a) and Neisser et al. (1996) articles. We labeled these as “statements of questionable accuracy,” and our study design designated that the first author would search for these statements during the accuracy coding process. Just as in the coding process for inaccurate statements, the first author compiled all of the statements of questionable accuracy by noting them (including a reference) in order to report a summary of the results.

After statements of questionable accuracy were compiled, we noticed that these all fit into three categories. One type of statement of questionable accuracy were false statements that were not addressed in the Gottfredson (1997a) and Neisser et al. (1996) pieces. Sometimes these statements were trivial, such as Lilienfeld, Lynn, Namy, and Woolf’s (2014, p. 333) statement that David Wechsler was “among those classified as feeble-minded by early flawed IQ tests.” This statement (which does not have an accompanying citation) is untrue because Wechsler was too old to have been given an IQ test during his childhood. Therefore, Lilienfeld et al.’s (2014) statement was labeled as being of “questionable accuracy” because it did not contradict anything in the Gottfredson (1997a) or Neisser et al. (1996) articles. A biographical detail about Wechsler is minutia that most students will probably forget and few instructors—if any—will emphasize. On the other hand, sometimes these statements of questionable accuracy could distort students’ understanding of intelligence. For example, two different textbooks (Nairne, 2014; Zimbardo, Johnson, & McCann, 2017) report the results of the Minnesota Transracial Adoption Study (Scarr & Weinberg, 1976; Weinberg, Scarr, & Waldman, 1992) in inaccurate or overly simplified ways that make environmental influences on IQ seem more important than many experts would argue (e.g., A. R. Jensen, 1998; Lee, 2010; Levin, 1994; Lynn, 2015; Plomin & Petrill, 1997).

Another type of statement of questionable accuracy was statements which recent research would call into question. For example, stereotype threat (Steele & Aronson, 1995) was discussed in Neisser et al.’s (1996) article. But recently attempts at replicating stereotype threat effects have sometimes been disappointing (e.g., Walker & Bridgeman, 2008), and some experts have expressed doubts of the reality of the stereotype threat phenomenon (Flore & Wicherts, 2015; Ganley et al., 2013) and its applicability outside of the laboratory (Lee, 2010).

The last type of statement of questionable accuracy was statements which the first author (who teaches a course on human intelligence and has published multiple peer-reviewed articles on the topic) did not believe would find widespread agreement among experts on intelligence research. An example of this was Feist and Rosenberg’s (2015, p. 360) claim that “fluid intelligence is not influenced by culture or the size of your vocabulary. Instead, it simply involves how fast you learn things.” The first author classified this as a statement of questionable accuracy because of evidence that experience and crystallized intelligence can influence fluid intelligence (e.g., Lohman, 2006) and because some experts (e.g., Carroll, 1993) would question whether learning speed and fluid intelligence are synonymous.

The standard for logical fallacies was taken from a list compiled by Gottfredson (2009) of 13 logical fallacies used to dismiss research on intelligence. These fallacies cover a wide variety of specious arguments that scientists and lay people use to dismiss intelligence research. Table 2 gives a brief description of the fallacies and provides an example of each fallacy in the textbooks. Like the coding process

for textbook accuracy, the decision to use the Gottfredson (2009) list of logical fallacies was an *a priori* decision. We also decided in advance that for statements to be labeled as a fallacy, the statement had to be explicit in applying the specious reasoning to a discussion about intelligence. Statements that appeared to be fallacious were transcribed and recorded.

Whether judging a statement as inaccurate, of questionable accuracy, or as a logical fallacy, the first author was always as conservative as possible. If the implications or subtext of a statement were inaccurate, fallacious, or of questionable accuracy, the first author did not code the statement as being problematic. Only statements that explicitly met the criteria were coded as inaccurate, fallacious, or of questionable accuracy. To verify the first author’s work, the second author (who was very familiar with the sources the logical fallacies and the standards of factual accuracy) examined the lists of inaccurate, questionably accurate and fallacious statements, and removed any that she did not fully agree had met the criteria.

Results

Descriptive Statistics

Table 1 shows that descriptive statistics of the results of our analysis of introductory textbooks. The mean number of pages discussing intelligence was 19.5 ($SD = 7.80$), and the average length of a textbook was 591.0 pages ($SD = 102.71$), indicating that the average textbook author devotes 3.29% of their textbook to discussing intelligence, which is in accordance with previous research indicating that 3% to 4% of introductory psychology textbook length is dedicated to intelligence (Griggs, 2014a; Jackson & Griggs, 2013). A total of 11 of the 29 textbooks (37.9%) dedicate an entire chapter to intelligence. In the other textbooks, authors combined their discussion on intelligence with sections on language, cognition, creativity, memory, and other allied topics.

Topics Covered in Textbooks

In the analysis that maximized detail, we found 102 topics that were discussed in at least two introductory psychology textbooks. The mean number of these topics discussed in a textbook was 30.9 ($SD = 9.67$, $min = 17$, $max = 53$). The 10 most common topics in textbooks were IQ (27 books, 93.1%), Spearman’s *g* (27 books, 93.1%), Gardner’s theory of multiple intelligences (27 books, 93.1%), Sternberg’s triarchic theory of intelligence (26 books, 89.7%), the measurement of intelligence (24 books, 82.8%), psychometric validity and reliability (23 books, 79.3%), Alfred Binet and his work (22 books, 75.9%), the Stanford-Binet intelligence test (21 books, 72.4%), environmental influences on intelligence (20 books, 69.0%), and intellectual disabilities (20 books, 69.0%). For a full list of topics taught in at least two textbooks, see Table 3 and Supplemental File 2.

When we minimized the number of categories, we found 40 topics discussed in at least two textbooks ($M = 18.2$, $SD = 3.98$, $min = 13$, $max = 29$). According to this list, the 10 most common topics discussed in textbooks were psychometrics (29 books, 100%), IQ (28 books, 96.6%), Spearman’s *g* (28 books, 96.6%), specific tests (e.g., Stanford-Binet; 28 books, 96.6%), genetics and environment (28 books, 96.6%), Gardner’s theory of multiple intelligences (27 books, 93.1%), Sternberg’s triarchic theory of intelligence (26 books, 89.7%), the history of intelligence tests (24 books, 82.8%), extremes of intelligence (23 books, 79.3%), and group differences (23 books, 79.3%). Table 3 and Supplemental File 2 contain the full list of topics taught in at least two textbooks.

Table 2
Gottfredson's (2009) *Logical Fallacies Used to Dismiss Intelligence Research*

Fallacy #	Description	Citation in textbooks
1	Yardstick = Construct Item appearance and surface content reflect the actual ability an item measures.	Wade et al., 2014, pp. 320–321
2	Intelligence is a marble collection Intelligence/ <i>g</i> is a collection of specific abilities or skills that are forced to “add up” to an IQ score. Thus, <i>g</i> doesn't really exist.	Hockenbury, Nolan, & Hockenbury, 2015, p. 308
3	Nonfixedness proves malleability Because IQ can grow or change over the lifespan, it must be possible to change people's IQs.	Nevid, 2015, p. 271
4	Improvability = Equalizability Because skills and knowledge can be improved in groups, it must be possible to equalize groups.	Myers & DeWall, 2015, p. 412
5	Gene-environment interaction nullifies heritability Genes and environment work together to produce phenotypes. Therefore, it is impossible to tease apart the influence of either one.	Pastorino & Doyle-Portillo, 2016, p. 340
6	Genetic similarity among humans negates differences Humans are 99%+ alike genetically. Therefore, that last < 1% must be trivial.	Grison et al., 2017, p. 301
7	Contending definitions of intelligence negate evidence Because there are conflicting definitions of intelligence, no one really knows what (if anything) IQ tests measure.	Lahey, 2012, p. 286
8	Phenotype = genotype Physical differences among humans (e.g., IQ scores) are innate, genetically determined traits.	Lahey, 2012, p. 296
9	Biological = genetic A biological difference (e.g., brain size, reaction time) must be genetically caused.	None
10	Environmental = nongenetic Environmental influences on development (e.g., parental SES) are unaffected by individuals' genes.	Ciccarelli & White, 2015, p. 292
11	Imperfect measurement pretext IQ tests must perfectly measure intelligence and/or make predictions perfectly before they can be used or trusted.	Gerrig, 2013, pp. 256–257
12	Dangerous thoughts trigger Socially divisive or controversial ideas should be held to a higher standard of proof before dissemination.	Feldman, 2015, p. 279
13	Happy thoughts leniency Politically correct or socially acceptable ideas are the default belief or should be given less scrutiny.	Lilienfeld et al., 2014, p. 353

Logical Fallacies

Tables 1 and 2 show that almost every fallacy was mentioned at least once across the books. The textbooks contained an average of 1.76 logical fallacies ($SD = 1.21$, $min = 0$, $max = 4$). The most commonly committed fallacies were Fallacies 2 (eight books), 3 (eight books), 4 (six books), and 6 (six books).

Fallacy 2 is committed when an author indicates that intelligence does not exist because it is a collection of abilities that an IQ test creator happens to choose to include on their test. For example, Coon and Mitterer (2016) stated:

many psychologists simply accept an operational definition of intelligence by spelling out the procedures they use to measure it. . . . Thus, by selecting items for an intelligence test, a psychologist is saying in a direct way, “This is what *I* mean by intelligence.” A test that measures memory, reasoning, and verbal fluency offers a very different definition of intelligence than one that measures strength of grip, shoe size, hunting skills, or the person's best *Candy Crush* mobile game score. (p. 290)

This gives readers the idea that because “intelligence” is nothing more than the sum of whatever tasks a psychologist arbitrarily chooses to put on a test. However, this is not the case because a common *g* factor accounting for about half of variance on cognitive tasks has been found across many human cultures (e.g., Carroll, 1993; Dolan, 2000; Dolan & Hamaker, 2001; Frisby & Beaujean, 2015; Gurven et al., 2017; Reuning, 1972) and even in nonhuman primates (Fernandes, Woodley, & te Nijenhuis, 2014; Herndon, Moss, Rosene,

& Killiany, 1997), dogs (Arden & Adams, 2016), and rats (Anderson, 1993; Galsworthy, Paya-Cano, Monleón, & Plomin, 2002). Thus, the existence of *g* is not dependent on the items on an IQ test used to investigate it. Many different collections of cognitive tasks produce an overall *g* factor (Carroll, 1993), and these different *g* factors from different test batteries are highly correlated (usually $r \geq .95$), indicating their *g* factors correspond to the same ability (Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004; Johnson, te Nijenhuis, & Bouchard, 2008).

Fallacy 3 is the idea that because people learn as they age or because IQ scores can change or fluctuate over the lifetime, it is possible to change a person's IQ. The most common manifestation of Fallacy 3 in the introductory psychology textbooks was the claim that psychologists know how to raise IQ among individuals in high quality environments (e.g., in industrialized nations). As an example, Schacter, Gilbert, Wegner, and Nock (2014, p. 420) listed ways in which parents could raise their child's IQ, including enriching pregnant women's diets with polyunsaturated fatty acids and sending children to preschool. In addition to being unsupported by any citations to the scholarly literature, Schacter et al.'s (2014) list ignores the disappointing results from studies of efforts to raise intelligence in most children, with the fadeout of any IQ gains occurring quickly (Protzko, 2015; see Simons et al., 2016, for similarly disappointing results of “brain training” studies of older adults). In evaluating textbooks for their adherence to Fallacy 3, we acknowledged that eliminating environmental characteristics with a known negative im-

Table 3
Table of Topics Related to Intelligence in Textbooks

Maximum number of topics		Minimum number of topics	
Concept	Number (%) of books	Concept	Number (%) of books
Gardner and his theory of multiple intelligences	27 (93.1%)	Psychometrics	29 (100.0%)
IQ	27 (93.1%)	Genetics and environment	28 (96.6%)
Spearman and <i>g</i>	27 (93.1%)	IQ	28 (96.6%)
Sternberg and his theory of triarchic intelligence	26 (89.7%)	Spearman and <i>g</i>	28 (96.6%)
Measurement of intelligence	24 (82.8%)	Tests (specific)	28 (96.6%)
Validity and reliability	23 (79.3%)	Gardner and his theory of multiple intelligences	27 (93.1%)
Binet and his tests	22 (75.9%)	Sternberg and his triarchic theory of intelligence	26 (89.7%)
Stanford-Binet test	21 (72.4%)	History of intelligence tests	24 (82.8%)
Environment	20 (69.0%)	Extremes of intelligence	23 (79.3%)
Fluid and crystallized intelligence	20 (69.0%)	Group differences	23 (79.3%)
Intellectual disabilities	20 (69.0%)	Types of intelligence	21 (72.4%)
Wechsler and his tests	19 (65.5%)	Emotional intelligence	18 (62.1%)
Emotional intelligence	18 (62.1%)	Cultural bias in intelligence tests	17 (58.6%)
Extremes and giftedness	18 (62.1%)	Flynn effect	17 (58.6%)
Genetics and environment	18 (62.1%)	Long term outcomes	17 (58.6%)
Mental age	18 (62.1%)	Theories of intelligence	15 (51.7%)
Cultural bias on intelligence tests	17 (58.6%)	Self-fulfilling prophecies	14 (48.3%)
Flynn effect	17 (58.6%)	Politics and controversies	13 (44.8%)
Heritability	16 (55.2%)	Definitions of intelligence	12 (41.4%)
Modern intelligence tests	16 (55.2%)	Biology and brain	11 (37.9%)
Theories of intelligence	15 (51.7%)	Early theorists	11 (37.9%)
Genes	14 (48.3%)	Cognitive psychology and intelligence	10 (34.5%)
History of intelligence tests	13 (44.8%)	Evaluating intelligence tests	10 (34.5%)
Stereotype threat	13 (44.8%)	Normal distribution	10 (34.5%)
Definitions of intelligence	12 (41.4%)	Multifactor theories (including Thurstone's)	8 (27.6%)
Group differences in intelligence	12 (41.4%)	Hierarchical theories of intelligence (including Carroll three-stratum model and CHC theory)	7 (24.1%)
Race differences in intelligence	12 (41.4%)	Education system	7 (24.1%)
Standardization	12 (41.4%)	Achievement and aptitude	5 (17.2%)
Twin studies	12 (41.4%)	Causes of <i>g</i>	5 (17.2%)
Biology and the brain	11 (37.9%)	Creativity, convergent/divergent thinking, and creativity tests	5 (17.2%)
Calculating IQ	11 (37.9%)	Non-Western views/context	5 (17.2%)
Evaluating intelligence tests	10 (34.5%)	Individual differences	4 (13.8%)
Family studies of intelligence	10 (34.5%)	Statistics	4 (13.8%)
Nature vs. nurture	10 (34.5%)	Tacit knowledge	4 (13.8%)
Normal distribution	10 (34.5%)	Wisdom	4 (13.8%)
Sex differences in intelligence	10 (34.5%)	Adaptive testing	2 (6.9%)
"Termites"	10 (34.5%)	Animal intelligence	2 (6.9%)
Adoption studies	9 (31.0%)	Artificial intelligence	2 (6.9%)
Galton and his work	9 (31.0%)	Human evolution	2 (6.9%)
Interventions	9 (31.0%)	Personality	2 (6.9%)
Long-term outcomes associated with intelligence	9 (31.0%)		
<i>The Bell Curve</i>	9 (31.0%)		
Test norms	8 (27.6%)		
Savants	8 (27.6%)		
Between- and within-group differences	7 (24.1%)		
Cognitive psychology and intelligence	7 (24.1%)		
Cultural differences in intelligence	7 (24.1%)		
Culture reduced testing	7 (24.1%)		
Education	7 (24.1%)		
Factor analysis	7 (24.1%)		
Aptitudes and aptitude tests	6 (20.7%)		
<i>s</i> (specific abilities)	6 (20.7%)		
Thurstone and his work	6 (20.7%)		
Achievement and achievement tests	5 (17.2%)		
Causes of <i>g</i>	5 (17.2%)		
Creativity and creativity tests	5 (17.2%)		
General or specific intelligence?	5 (17.2%)		
Learning disabilities	5 (17.2%)		
Miscellaneous types of intelligence	5 (17.2%)		
Non-Western views/context	5 (17.2%)		
Politics and controversies	5 (17.2%)		
Psychometric approach to investigating intelligence	5 (17.2%)		
Psychometrics	5 (17.2%)		

(table continues)

Table 3 (continued)

Maximum number of topics		Minimum number of topics	
Concept	Number (%) of books	Concept	Number (%) of books
Stability of IQ	5 (17.2%)		
Twin and adoption studies	5 (17.2%)		
Working memory	5 (17.2%)		
Eugenics	4 (13.8%)		
Hierarchical theories of intelligence	4 (13.8%)		
Individual differences	4 (13.8%)		
Mental speed	4 (13.8%)		
Reaction range	4 (13.8%)		
Self-fulfilling prophecies	4 (13.8%)		
Statistics	4 (13.8%)		
Tacit knowledge	4 (13.8%)		
Wisdom	4 (13.8%)		
Army Alpha and Army Beta	3 (10.3%)		
Conclusion section	3 (10.3%)		
Dove Test	3 (10.3%)		
Group tests of intelligence	3 (10.3%)		
Intelligence and age	3 (10.3%)		
Longitudinal studies	3 (10.3%)		
Metacognition	3 (10.3%)		
Mozart and/or Beethoven	3 (10.3%)		
Reaction time	3 (10.3%)		
Adaptive testing	2 (6.9%)		
Animal intelligence	2 (6.9%)		
Artificial intelligence	2 (6.9%)		
Cattell-Horn-Carroll model of intelligence	2 (6.9%)		
Chronological age	2 (6.9%)		
Convergent thinking	2 (6.9%)		
Criticisms of intelligence tests	2 (6.9%)		
Goddard and his work	2 (6.9%)		
Human evolution	2 (6.9%)		
Meaning of IQ scores	2 (6.9%)		
Mindset theory	2 (6.9%)		
Multi-factor theories	2 (6.9%)		
Personality	2 (6.9%)		
Prodigies	2 (6.9%)		
Spearman and Thurstone	2 (6.9%)		
Terman's theories	2 (6.9%)		
Theory of mind	2 (6.9%)		
Working memory and attention	2 (6.9%)		

Note. For a detailed enumeration of topics in each textbook, see Supplemental File 2.

fact on intelligence—such as lead poisoning, iodine deficiency, and severe childhood neglect—can raise IQ (e.g., Huang et al., 2012). Therefore, we only identified this fallacy if a textbook author (a) stated that interventions to raise IQ were successful for most or all individuals, or (b) if they did not distinguish between interventions that are appropriate for very poor environments and interventions designed for middle- and upper-class individuals in industrialized nations.

Whereas Gottfredson's (2009) Fallacy 3 is about interventions that raise individuals' IQ scores, Fallacy 4 concerns improving the intelligence level of groups or individuals until individual differences are eliminated. Those who subscribe to this fallacy believe that research on the environmental impact on IQ can be used to eliminate mean differences in IQ among groups or individuals. For example, Comer and Gould (2013) described research that found the mean IQ score gap between African American and White American students decreased from early adolescence through the end of college. Because of these findings, Comer and Gould (2013) concluded that gaps in these groups' mean scores can be closed completely. However, whether mean IQ score differences between racial groups are narrowing is a matter of contentious debate among experts (e.g., Nisbett et al., 2012; Murray, 2007; Rushton, 2012; Williams & Ceci, 1997). And even if

these mean score differences are narrowing, it does not indicate that the gaps will close completely or that experts understand exactly which societal changes are driving improvements for low scoring individuals. Indeed, some experts have questioned whether closing mean IQ score gaps among some demographic groups is even possible (e.g., Woodley & Meisenberg, 2012). Although it is conceivable that one day an environmental intervention could eliminate individual or group differences in IQ scores, such an intervention does not exist at this time, and there is no guarantee that it ever will (Lee, 2010). Those who subscribe to Fallacy 4 cling to the possibility of an intervention and ignore (or deny) its current nonexistence.

Fallacy 6 primarily focuses on the similarities of human beings, stating that because human genes are about 99% alike, genetic differences among people or demographic groups must be inconsequential. For example, when discussing racial differences in intelligence, Grison, Heatherton, and Gazzaniga (2017) wrote:

The first issue to consider is whether 'race' is a biologically meaningful concept. . . . The vast majority of genes—perhaps as many as 99.9%—are identical among all humans. . . . But it is unlikely that differences in skin color and hair type relate to the mental capacities that underlie intelligence. (pp. 301, 302)

Fallacy 6 is so common in the biological and social science literature that it even has its own name: *Lewontin's fallacy*, named for a biologist who popularized it (Lewontin, 1972). Briefly, what makes this reasoning fallacious is that this high degree of genetic similarity is what makes all humans belong to the same species and able to interbreed. Because these genes are identical, they cannot be responsible for any phenotypical differences among humans. (In comparison, humans and chimpanzees differ in about 4% of their genes; see The Chimpanzee Sequencing and Analysis Consortium, 2005.) But this high degree of similarity does not indicate that the relatively few genetic differences among humans are irrelevant. Indeed, these genetic differences are responsible for at least some of the interpersonal phenotypic variation in a wide variety of traits, including height (Yang et al., 2015), heart disease (Pickrell et al., 2016), aggressive behavior (Beaver, Barnes, & Boutwell, 2016), schizophrenia (Sariaslan, Larsson, & Fazel, 2016), and intelligence (Davies et al., 2015). Other explanations of why Lewontin's fallacy is incorrect are available (e.g., Edwards, 2003; Smouse, Spielman, & Park, 1982).

Extending the incorrect reasoning of Fallacy 6 to another species shows why it is unwise to rely on genetic similarity among individuals to dismiss phenotypic variation. All domesticated dog breeds differ by only 0.15% of their genes—a much lower level of genetic variation than humans. By the reasoning of Fallacy 6, the breed of a dog doesn't matter for its owner's use because such slight genetic differences are trivial. Therefore, poodles can pull Iditarod sleds, and a pug is a great police dog (see Figure 3). Indeed, many canid species (e.g., wolves, foxes, dogs, jackals, and dingoes) are so closely related that they can interbreed (Wayne & Ostrander, 1999). Thus, even a small percentage of genetic diversity can contribute greatly to a population's phenotypical variance (The Chimpanzee Sequencing and Analysis Consortium, 2005, p. 83).

Tables 1 and 2 show that authors frequently perpetuate logical fallacies related to intelligence in their introductory psychology textbooks. Indeed, nearly all the fallacies from Gottfredson's (2009) list appeared in at least one book, indicating that these specious arguments about intelligence research are commonly disseminated. Other examples of these logical fallacies in the textbooks can be found in Supplemental File 1 (pp. 3–21).



Richard Lewontin creates a dog sled team to test his theory that small genetic differences are irrelevant.

Figure 3. The incorrect reasoning of Lewontin's fallacy. Slight differences in genotypes among organisms can result in major phenotype differences. Compared with humans, dogs are much less genetically diverse. Ignoring these genotype differences among organisms in the same species is nonsensical. (Art by Bryan Johnson, caption by Russell T. Warne. Copyright: Russell T. Warne, 2017. This work is licensed under the Creative Commons CC-BY-NC-ND License. The image—with accompanying caption—may be shared for non-commercial purposes with attribution to the authors, but may not be altered without permission of the copyright holder.)

Inaccuracies

Only six books (20.7%) did not have any inaccuracies. These books were Bonds-Raacke (2014); Feist and Rosenberg (2015); Griggs (2014b); Kalat (2017); Okami (2014); and Wade, Tavris, and Garry (2014). The remaining 23 textbooks (79.3%) contained at least one factually inaccurate statement about human intelligence, a similar percentage as was found in the Pesta et al. (2015) study of organizational behavior textbooks.

By far the most common type of inaccurate statement was related to test bias (14 books, 48.2%). These inaccurate statements about test bias often were claims that mean differences in test scores among demographic groups (especially racial/ethnic groups) were due to test bias. For example, Morris and Maisto (2016) stated:

Another major criticism of intelligence tests is that their content and administration do not take into account cultural variations and, in fact, discriminate against minorities. High scores on most IQ tests require considerable mastery of standard English, thus biasing the tests in favor of middle- and upper-class White people. (p. 252)

However, this directly contradicts the mainstream statement on intelligence from Gottfredson (1997a), which stated:

Intelligence tests are not culturally biased against American Blacks or native-born, English-speaking peoples in the U.S. Rather, IQ scores predict equally accurately for all such Americans, regardless of race and social class. Individuals who do not understand English well can be given either a nonverbal test or one in their native language. (p. 14; see also Neisser et al., 1996, p. 90)

Indeed, eliminating test bias is a commonplace procedure among professional psychometrics that it is required before putting a cognitive, intelligence, or educational test on the commercial market (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; Warne, Yoon, & Price, 2014).

A related inaccurate statement that was found in the textbooks was claims that intelligence simply could not be measured in any meaningful way or that it was extremely challenging to measure intelligence (e.g., Coon & Mitterer, 2016; Feldman, 2015; Pastorino & Doyle-Portillo, 2016). This contradicts the mainstream view of intelligence research, which stated, "Intelligence . . . can be measured, and intelligence tests measure it well" (Gottfredson, 1997a, p. 13). Additionally, one of the bedrock foundations of intelligence testing is the "indifference of the indicator" (Spearman, 1927, p. 197), which states that the surface content of intelligence test items is irrelevant. Rather, any test item that requires cognitive effort measures—at least partially—intelligence (Lubinski & Humphreys, 1997). This means that it is actually easier to measure intelligence than many other psychological constructs. Indeed, some individuals trying to measure other constructs have inadvertently created intelligence tests (see Gottfredson, 2004, and Reeve & Basalik, 2014, for an example).

Another frequent topic of inaccurate statements was the claim that intelligence is only relevant in academic settings—and not in everyday life (e.g., Caciopo & Freberg, 2016, p. 385; Coon & Mitterer, 2016, pp. 297, 309). However, the APA statement (Neisser et al., 1996, pp. 82–83) and the mainstream statement on intelligence (Gottfredson, 1997a, p. 13) argued that intelligence tests scores do correlate with many nonacademic life outcomes. Indeed, there is a vast body of research on the correlates of intelligence (Gottfredson, 1997b; A. R. Jensen, 1998; Warne, 2016). This is why intelligence scholars believe that "there is little that IQ doesn't help predict" (Belasen & Hafer, 2013, p. 615). Other topics that were the subject of inaccurate

statements in the textbooks included the influence of environment and genes on intelligence, culture, and racial issues related to intelligence.

Questionable Accuracy

There were several ways that textbook authors provided statements of questionable accuracy. We will focus on the most common topics of questionably inaccurate statements: race, stereotype threat, environmental influences on intelligence, and Lewontin's seed analogy. However, readers should be aware that there were other topics that were the subject of statements of questionable accuracy, including sex differences in intelligence, culture and intelligence, *The Bell Curve* (Herrnstein & Murray, 1996), the history of intelligence research/testing, and childhood intervention programs. For a more detailed summary of the statements that we found questionably accurate, see Supplemental File 1.

Race. One of the most controversial topic in all of science is mean race differences in intelligence (Check Hayden, 2013; Cofnas, 2016). The topic is so taboo that it is one of the few topics that some scholars declare a priori off limits to empirical investigation (e.g., Kourany, 2016; Sternberg, 2005) and which is subject to regular censorship in the scientific community (Gottfredson, 1994, 2009, 2010). Perhaps for this reason, some textbook authors (e.g., Bonds-Raacke, 2014; Kalat, 2017; Nolen-Hoeksema, Fredrickson, Loftus, & Lutz, 2014) avoided the topic. Because racial differences in intelligence are not the most important issue related to intelligence research, this is a valid way to handle the controversy (Hunt, 2014).

However, many authors who chose to address the controversy had questionably accurate statements regarding race and intelligence. Many of these revolved around the persistent finding (dating back nearly a century) that—on average—American White examinees score approximately 15 points higher than African Americans (e.g., Roth, Bevier, Bobko, Switzer, & Tyler, 2001; Yerkes, 1921, p. 707), though some scholars contend that in recent decades the score gap has narrowed to as little as 10 points (Hedges & Nowell, 1999; Nisbett et al., 2012). None of the textbook authors denied the existence of these mean score differences, though in one textbook the differences were described as being “relatively small” (Gleitman, Gross, & Reisberg, 2011, p. 450). We classified this particular statement as being questionably accurate because a 10–15 point mean difference would have major consequences for individuals and society; some people may not consider that difference “relatively small.”

But the textbook authors' explanation of the cause(s) of mean score differences across groups was frequently a source of questionably inaccurate information. Although these statements took a variety of formats, they all emphasized environmental, nongenetic causes for the score differences across White American and African American examinees and de-emphasized, minimized, dismissed, or ignored possible empirically supported genetic causes. Because clear answers in this contentious nature/nurture debate elude experts, we classified statements as being questionably accurate if the authors' explanation for these race differences was largely discredited or if it completely denied the role of genetics. For example, Nevid (2015, p. 271) stated—citing evidence from the Minnesota Transracial Adoption Study—that “being raised in an environment that places a strong value on educational achievement . . . basically canceled out the oft-cited 15-point gap in IQ between these two groups.” However, the data clearly reported in the study shows that in adolescence clearly showed that differences were not “basically cancelled out.” Indeed, IQ scores of African American and multiracial adoptees were 7–16 points lower than White adoptees' scores and 10–20 points lower than the scores from White adopted parents' biological children (Weinberg et al., 1992, p. 123). Even the earlier report from the study—largely seen as more favorable to an environmental argument—still

shows a 5-point mean difference between African American and White adoptees' IQ scores (Scarr & Weinberg, 1976, p. 732). Generally, scholars in the field of intelligence see the evidence from this study—and others (e.g., Moore, 1986)—as consistent with both environmental and genetic hypotheses for the cause of Group IQ score differences (e.g., Nisbett et al., 2012).

Stereotype threat. A topic closely related to racial differences in intelligence that appeared in 13 (44.8%) textbooks was stereotype threat. Originally proposed by Steele and Aronson (1995), stereotype threat is the tendency for individuals to do more poorly on an ability test when they are reminded that their demographic group (e.g., a racial or a gender group) performs more poorly on standardized tests than other groups. In addition to recent difficulties in replicating these results, there is new circumstantial evidence of publication bias in the stereotype threat literature (Flore & Wicherts, 2015; Ganley et al., 2013), which may inflate the apparent strength of the effect.

Although it is an intriguing theory, stereotype threat is not a viable explanation for the cause of mean racial and gender group differences in intelligence test scores. Sackett, Hardison, and Cullen (2004) showed that Steele and Aronson (1995) used stereotype threat to create *new* mean score differences between African American and White examinees. Therefore, Steele and Aronson's (1995) study explained nothing about *preexisting* score gaps between demographic groups (Sackett et al., 2004). However, scholarly authors, the popular press, and psychology textbook authors interpreted stereotype threat as explaining part or all of the score difference between groups (Sackett et al., 2004). We found that misinterpretations have not lessened since the Sackett et al. (2004) study. The majority (9 of 13) textbooks in our sample that discussed stereotype threat stated that it was a partial cause and/or a contributing factor to mean differences in IQ scores across racial groups (Gerrig, 2013, p. 255; Gleitman et al., 2011, pp. 455–456; Gray & Bjorklund, 2014, p. 407; Hockenbury, Nolan, & Hockenbury, 2015, pp. 304–305; Kalat, 2017, pp. 306–307; Myers & DeWall, 2015, p. 415; Nairne, 2014, pp. 330–331; Schacter et al., 2014, pp. 418–419; Wade et al., 2014, pp. 321–322).

Environment. When examining the relative importance of environmental and genetic factors, even the most strident hereditarians acknowledge that individual and group differences in intelligence are at least partially influenced by environmental variables (e.g., Gottfredson, 2005; Levin, 1994; Lynn, 2015; Rushton & Jensen, 2005).

In contrast, many textbook authors suggested in their books that environmental factors were the dominant or only cause for the differences in IQ scores among individuals. For example, Schacter et al. (2014) discussed that high socioeconomic status (SES) was the cause of higher IQ scores, even stating, “Money can't buy love, but it sure appears to buy intelligence” (p. 413), and provided three paragraphs of supporting information. Although there is no question that socioeconomic status is correlated with intelligence, evidence is clear that individuals' intelligence is at least a partial cause of their SES (Deary et al., 2005; Snyderman & Rothman, 1987; Strenze, 2007) and/or that the two variables share a common partial genetic cause (Marioni et al., 2014; Trzaskowski et al., 2014).

Lewontin's seed analogy. One way in which several textbook authors minimized or denied the importance of genetic factors was by arguing that the causes of individual IQ variability may have nothing to do with group differences in IQ. Most of these authors illustrated this claim with an analogy of two randomly selected groups of seeds, one planted in favorable soil and another planted in unfavorable soil. Because the two groups of seeds were formed randomly, any mean differences between groups must be due to the different environments, whereas any differences among individual plants within groups would be attributable to genetic variation. This analogy was usually attrib-

uted to Lewontin (1970),² who used it to argue that average IQ score differences were unrelated to genetic differences among groups. All of the textbook authors in this study who used the analogy (Bernstein, 2016; Gleitman et al., 2011; Gray & Bjorklund, 2014; Hockenbury et al., 2015; Myers & DeWall, 2015; Nolen-Hoeksema et al., 2014; Weiten, 2017) used it for the same purpose.

We did not classify the use of the seed analogy as being inaccurate because Neisser et al. (1996, p. 95) used the same analogy to state that within- and between-groups differences may not have the same origins. However, we classified the use of the analogy as “questionably accurate” because we disagreed with the logic of the analogy for two reasons. One reason is that group differences *must* be at least partially caused by individual genetic differences, unless groups are either formed randomly (as in Lewontin’s seed analogy), or are perfectly matched genetically (as in a study of identical twins raised in different homes). Neither of these scenarios applies to the formation of actual human racial/ethnic groups, which have genetic differences due to their geographically dispersed recent evolutionary histories (The 1000 Genomes Project Consortium, 2015; Tishkoff et al., 2009). Additionally, the possibility of unique influences that impact intelligence scores of one demographic group uniquely (e.g., a “legacy of slavery” among African Americans, or an influence of societal racism that only affects minority groups) has been empirically ruled out (Dalliard, 2014; Lubke, Dolan, Kelderman, & Mellenbergh, 2003; Rowe & Cleveland, 1996; Rowe, Vazsonyi, & Flannery, 1994, 1995).

The second reason we classified Lewontin’s seed analogy as being questionably accurate is that for intelligence to be heritable within groups but not between groups, the mean environmental differences between groups would be so large and/or so consistent that there would be little overlap between groups’ environments (A. R. Jensen, 1998, pp. 447–458). Indeed, assuming a within-group heritability of intelligence of 0.5 (a realistic estimate, see Deary, 2012; Gottfredson, 1997a; Neisser et al., 1996; Plomin & Petrill, 1997), environmental differences would have to be $d = 1.41$ for two groups to have a mean IQ difference of $d = 1.0$. Such a difference would indicate that (on a normally distributed variable) only 7.9% of individuals in the lower environment group would exceed the mean environment for the higher group. Although this size of a difference in environments is plausible between wealthy industrialized nations and poor undeveloped nations, it is not plausible as an estimate for the environmental differences among groups within the United States. For example, in the United States, 59.5% of White households have incomes under \$75,000; 22.5% of Black households have incomes of \$75,000 or more (Proctor, Semega, & Mollar, 2016, Table A-1). The discrepancies in the adult education levels between African Americans and White Americans are even less than the income discrepancies are (see Ryan & Bauman, 2016, Table 1).

We acknowledge that “environment” consists of much more than just education and income. As a result, some may argue that an aggregate environment variable could explain the mean IQ score differences in African Americans and White Americans. However, for this to be true, differences in environmental variables would have to (a) be remarkably consistent in their unfavorability toward African Americans, and (b) have a causal impact on intelligence levels. Assuming 10 relevant environmental variables average $r = .40$ in their intercorrelations, members of the lower scoring group would have to average 0.53 standard deviations below the higher scoring group’s mean on all variables in order to produce an aggregate environmental difference of $d = 1.41$. In a normally distributed variable, a $d = .53$ difference would indicate that 29.8% of the lower scoring group’s members exceed the higher scoring group’s mean. Although this is a possible mean group difference for some variables—such as income—it is not realistic for many others, such as

education level or many variables (e.g., vaccination rates, access to early childhood education).³

Our point in this discussion is not to say that environment is totally irrelevant in explaining mean group differences in IQ; such a claim is not supported by the evidence. Rather, we wish to show that empirical data shows that the possibility of a completely environmental explanation of group differences in IQ—as Lewontin’s (1970) analogy implies—is not plausible.

Discussion

Most Frequently Discussed Topics

The first research question in this study was, “What are the most frequently discussed topics related to intelligence in introductory psychology textbooks?” Our results indicated that many of the most frequently discussed topics related to intelligence in textbooks are important concepts to scholars in the field (e.g., IQ, the measurement of intelligence, intellectual disabilities). Although there is no objective list of the “correct” concepts that textbook authors should discuss, we believe that most intelligence researchers would find relevance in many frequently discussed topics listed in Table 3 (see also Supplemental File 2).

However, two concepts received much more attention from textbook authors than they do from the intelligence research community: Gardner’s theory of multiple intelligences and Sternberg’s triarchic theory of intelligence. Almost every textbook author included these topics in their discussion, a finding that is in accordance with previous research showing that Gardner’s *Frames of Mind* is one of the most frequently cited books in introductory psychology textbooks (Griggs et al., 2004). Gardner’s and Sternberg’s theories are popular with textbook authors, even though key aspects of both theories have little empirical support. Indeed, researchers investigating cognitive abilities using these theories often produce g in their data anyway (e.g., Pyryt, 2000). We suggest that textbook authors should either eliminate a discussion of these theories or present them in a more nuanced light, exploring both the theories’ strong points and aspects that are unsupported by empirical data. For example, it would be fair to discuss how Gardner’s linguistic, logical-mathematical, and visual-spatial intelligences align well with the broad midlevel factors in modern

² As a historic sidenote, Lewontin was not the first to publish the analogy; it originated with Thoday (1969).

³ These calculations are based on the formula provided by Schneider (2016, p. 8). The average group difference on specific variables that contribute to a composite is dependent on (a) the number of variables and (b) the intercorrelation among variables. As (a) increases or (b) decreases, there is a decrease in the average group difference needed on individual variables to produce a given group difference on a composite variable. We chose 10 variables and an average intercorrelation of $r = .40$ arbitrarily. But the results are not greatly different for other plausible estimates. For example, with 20 variables and an average intercorrelation of $r = .40$, the required average group difference for individual variables to produce a composite mean difference of $d = 1.41$ would be $d = .37$. Fifteen variables with a mean intercorrelation of $r = .35$ would require an average group difference of $d = .40$ to produce a composite mean difference of $d = 1.41$. In theory, one could suggest an extremely large number of environmental variables, each with an extremely small impact on intelligence levels (e.g., with a mean intercorrelation of $r = .50$, it would require 75 with a mean difference of $d = .20$ to produce a composite difference of $d = 1.41$). But then the challenge becomes generating a plausible list of dozens of nonredundant variables that all have a possible causal impact on intelligence after genetic influences have been controlled for. Given the lackluster results of randomized interventions to increase intelligence (e.g., Lipsey, Farran, & Hofer, 2015; U.S. Department of Health & Human Services, 2010) and the severe attenuation in effect sizes in correlational studies when genetic effects are controlled for (e.g., Bouchard, Lykken, Tellegen, & McGue, 1996), we believe that a lengthy list of variables that each have a small detrimental impact on African Americans’ intelligence levels would be extremely difficult to produce.

intelligence tests and that there is strong evidence for these types of abilities. However, a balanced approach to Gardner's theory could discuss how his denial of *g* is at odds with over a century of psychometric data and that Gardner never explained how to measure these intelligences, nor has he embarked on any sort of systematic research program to gather data to test his theory (Hunt, 2001; Lubinski & Benbow, 1995). Likewise, in discussing Sternberg's theory, most textbook authors would be justified in praising Sternberg's emphasis on creativity, a trait that many psychologists and educators value (e.g., Guilford, 1950; Subotnik, Olszewski-Kubilius, & Worrell, 2011). However, the textbooks would improve by mentioning how Sternberg's practical intelligence is not as general or important as *g*—an empirical fact that undercuts the theory (Gottfredson, 2003a). Discussions of the weaknesses of both theories are widely available (e.g., Deary, Penke, & Johnson, 2010; Gottfredson, 2003a, 2003b; Lubinski & Benbow, 1995; Waterhouse, 2006), as are rebuttals from their creators (e.g., Gardner, 1995; Sternberg, 2003; Sternberg & Hedlund, 2002). If textbook authors do decide to keep these theories in their textbooks, drawing from both types of sources would strengthen the discussion of these theories.

Spearman and *g* were mentioned in most books; however, there was often no explanation that *g* is the dominant theory in this field. Few books taught about hierarchical models, and those that did usually discussed these models in general terms. Only one textbook (Okami, 2014, pp. 455–456) mentioned the Carroll three-stratum theory by name, and none introduced the bifactor model or any other contemporary models. In contrast, nonmainstream theories (e.g., Gardner's and Sternberg's) were treated as favorably—or better—than *g* or modern theories of intelligence. Thus, we believe that some introductory psychology students would mistakenly think that Sternberg's or Gardner's theories are as scientifically supported as Spearman's *g* or the models shown in Figures 1 and 2. This situation also created an inherent contradiction in some of the textbooks: after writing positively about Gardner's and/or Sternberg's theories, some textbook authors (e.g., Ciccarelli & White, 2015; Comer & Gould, 2013; Nairne, 2014) then proceeded to discuss data from IQ research in depth—much of which was based on *g* theory.

Textbook Accuracy

The second research question we aimed to answer was, “How accurate are introductory psychology textbooks in their discussion of intelligence?” Judged solely by the number of factually inaccurate statements, the textbooks we examined were mostly accurate. Six of the textbooks had no factual inaccuracies, and no book had more than eight inaccurate statements or eight questionably accurate statements. Unfortunately, many of Gottfredson's (2009) logical fallacies were present in the textbooks. Apart from Fallacy 9, every fallacy appeared at least once. We found this unfortunate because few instructors—and likely no students—would likely have the background knowledge in psychometrics, population genetics, and intelligence research to understand the problems associated with these fallacies. We urge introductory psychology textbook authors to read Gottfredson's (2009) chapter and then scrutinize their work to eliminate these fallacies.

Textbook Themes

Although it was not a goal in our research, certain themes emerged during data analysis. These themes included the perspectives on psychological testing, an environmental bias, and a minimizing of the importance of individual differences. We will briefly discuss each of these themes.

Testing perspectives. One unexpected finding of our study was that most of the textbooks discussed basic psychometrics and the

principles of testing. Perhaps because scientifically based tests were first invented to test intelligence (Fancher, 1985), many authors used their discussion of IQ tests to teach foundational psychometric principles, like reliability and validity. Most of these discussions were at an appropriate level of complexity for introductory students. Moreover, these digressions into psychometric theory were usually integrated seamlessly into the text.

But the psychometric discussions in some books were not ideal. Several textbooks gave a distorted or oversimplified view of test bias to readers. Indeed, test bias was the most common topic of the inaccurate statements, and readers of nearly half of the textbooks in our study would receive the impression that intelligence tests are *prima facie* biased against diverse examinees. This flatly contradicts mainstream opinion about the topic, which is that “the issue of test bias is scientifically dead” (Hunter & Schmidt, 2000, p. 151).

Environmental bias. Although some books provided a balanced explanation of the existence interindividual intelligence differences, textbook authors almost always favored environmental explanations for the differences. Prominent authors (e.g., Sternberg, Kamin, Gould, Lewontin) who advocate a purely environmental explanation for intelligence differences (both among individuals and among groups) were usually cited uncritically and approvingly. But scholars who posit a moderate or strong genetic role in intelligence differences (e.g., Deary, Gottfredson, A. R. Jensen, Plomin, Rushton, Lynn) were often dismissed and rarely cited in a positive way—if they were mentioned at all. As a result, we doubt that most students would believe that mainstream scholars view both genetics and environment as important determinants of individual and mean group differences in intelligence. Although we do not know why textbook authors take this approach to the nature-nurture debate in intelligence, our literature review raised a few possibilities. One possibility is the deductive referencing that Steuer and Ham (2008) identified. Another explanation could be authors' desire that Ferguson et al. (in press) identified to present research on intelligent as being less ambiguous and more consistent than it really is.

Rather than shying away from genetic influences on intelligence, we challenge textbook authors to use intelligence research to introduce the basic principles of behavioral genetics into their textbooks (see Plomin, deFries, Knopik, & Neiderhiser, 2016, for an accessible introduction to the topic). In recent years, the evidence that genes are at least a partial influence of every human behavior and psychological trait has mounted so quickly that the early 21st century may be the dawn of a behavioral genetics revolution in psychology. Such a revolution may be as important—or more important—for psychology than the cognitive revolution was in the mid-20th century. Introducing behavioral genetics concepts in introductory psychology may help students and instructors handle these new advances in psychological science. Because the evidence that intelligence is influenced by genes is strong and dozens of alleles associated with IQ have already been identified, the textbook section on intelligence research may be an ideal place for authors to introduce behavioral genetics to psychology undergraduates.

Minimizing the importance of individual differences in intelligence. Probably the most surprising theme we observed in the introduction was the tendency of some authors to minimize the importance of individual differences in intelligence. Most frequently this appeared in the form of a tacit acknowledgment that IQ test scores correlate with academic success, followed by a quick denial that the scores are important for anything else in life (e.g., Coon & Mitterer, 2016, pp. 297, 309; Lahey, 2012, p. 289; Nolen-Hoeksema et al., 2014, p. 418; Weiten, 2017, p. 281). Although no one would say that IQ test scores are a perfect predictor of any life outcome, we found it

striking that the textbook authors who discussed the relationship between personality traits and life outcomes did not mention similar caveats, even though personality traits are also variables that correlate imperfectly long-term outcomes.

The attitude that some authors had toward the importance of intelligence in determining life outcomes is a noticeable contrast to the opinions of scholars of intelligence, who often explain that “measures of individual differences in . . . intelligence . . . are the most accurate predictors that we have of success in academic achievement, industrial and professional competence and military performance” (Hunt, 2014, p. 156; see also Belasen & Hafer, 2013, p. 615). Despite their enthusiasm, intelligence scholars are realistic in their attitudes toward the topic they study; frank admissions of the limits of intelligence are common (e.g., Schmidt & Hunter, 1998). It is true that *if* intelligence were relevant only in an academic setting, then it would be relatively unimportant. Nevertheless, intelligence seems to extend to many aspects of human life, which underscores its inclusion in introductory psychology textbooks. Hence, we find the tendency to minimize the importance of individual differences in intelligence puzzling.

Implications for Psychology

Some readers may question why psychologists should attend to the content of introductory textbooks. In response, we wish to remind readers of the importance of the introductory psychology course. With over a million students taking introductory psychology every year, the introductory psychology course and the assigned textbook are important for educating nonexperts about psychological science. Furthermore, because fewer than 10% of all psychology departments offer a course on intelligence (Stoloff et al., 2010), most psychology students will never have the chance to correct the misconceptions that they learn in their introductory course. Nonpsychology majors—who are a majority of introductory psychology students (Miller & Gentile, 1998)—would be even less likely to learn correct information about intelligence if their introductory psychology course contains inaccurate information. Thus, as a field, psychologists should remain watchful of the content of introductory psychology texts, as psychologists are ethically bound to present factually correct information to their students.

Beyond higher education implications, this study highlights the mismatch between scholarly consensus on intelligence and the beliefs of the general public (e.g., Cronbach, 1975; Freeman, 1923; Gottfredson, 1994; Snyderman & Rothman, 1987). After reading 43 inaccurate statements, 129 questionably accurate statements about intelligence, and 51 logical fallacies about intelligence in introductory psychology textbooks, the reason for this mismatch became obvious to us. We believe that members of the public likely learn some inaccurate information about intelligence in their psychology courses. The good news about this implication is that reducing the public’s mistaken beliefs about intelligence will not take a massive public education campaign or public relations blitz. Instead, improving the public’s understanding about intelligence starts in psychology’s own backyard with improving the content of undergraduate courses and textbooks.

Limitations

Although we feel that our findings are generally robust, some limitations must be considered in the context of our study. First, our sample size was limited to only 29 textbooks. We tried to ameliorate this limitation by including the bestselling introductory psychology textbooks. We also wish to note that this sample of textbooks is larger than most studies of introductory psychology textbooks (e.g., Griggs et al., 2004).

It is also possible that textbook content may not reflect what an instructor teaches in class and additional assigned readings. However, Griggs (2014a) stated “it is clear that introductory textbooks greatly impact the structure of the introductory course” (p. 9). Therefore, it can be assumed that textbooks are—at minimum—an important source of structure and content for introductory psychology courses. Although most instructors are free to correct or supplement their textbook’s content on intelligence, we doubt this happens often because most instructors probably do not have extensive education in intelligence research. Courses on intelligence are rare (Stoloff et al., 2010), and many instructors will have never had extensive education on the topic. For instructors trained in cognitive psychology, developmental psychology, social psychology, or other branches that are more concerned with the generalities of human behavior, the foundation that intelligence research has in individual differences may be somewhat alien.

Another shortcoming is that we were conservative in our criticism or the textbooks. For example, we limited our definition of a “factual accuracy” to anything that was contradicted by the Gottfredson (1997a) and Neisser et al. (1996) articles. This may make textbooks appear relatively accurate; comparing textbooks to a more recent article written by an expert (e.g., Deary, 2012) would have likely made textbooks seem less accurate. However, we thought that the two articles represented consensus among experts and had a foundation in robust empirical findings. We believed these standards would reduce the influence of our own professional opinions when judging accuracy.⁴ Another example of our conservative standards was that topics, inaccuracies, and logical fallacies had to be explicit. This reduced the subjectivity of our judgments, but it also made the fallacies appear rare. An example of this is Fallacy 13, which is that politically correct or socially acceptable ideas are treated more leniently than controversial ones. Despite the proenvironmental bias in explanations of intelligence differences in many books and the frequent use of the seed analogy, only two textbooks (Coon & Mitterer, 2016, p. 309; Lilienfeld et al., 2014, p. 353) had explicit statements to this effect.

Some may question the standards of accuracy in this study, which were two articles from the mid-1990s. However, we chose these standards because we did not expect introductory textbooks to reflect the most cutting-edge intelligence research. This is a reasonable expectation because it takes time for findings to be replicated, theories to develop, and for the importance of some empirical work to become apparent. We believed that two extensively cited, mainstream articles were sound summaries of the intelligence literature and were reasonable standards for textbook authors to meet. The choice of older articles also seemed fair considering research by Gorenflo and McConnell (1991), who found that it “typically takes 20 years or so before article is perceived as being ‘classic’ by most authors of introductory psychology texts” (p. 10).

Finally, we want readers to recognize that the data regarding inaccurate statements, questionably accurate statements, and logical fallacies were not checked for accuracy, as indicated by interrater reliability analysis. Rather, the second author examined the information compiled by the first author and removed any statements that she did not believe met the criteria. However, this is not a completely independent evaluation of the first author’s work because the first author provided the second author with her education on intelligence. We adopted this strategy for checking our data because of the limitations of the skills of our research team: the first author is the only expert on intelligence at the university, and an accomplished under-

⁴ After all, we were interested in the accuracy of textbooks compared with mainstream scholarly thought—not compared with our own professional opinions.

graduate was the most qualified person available to check the first author's work. Of course, this is not an ideal situation.

Conclusion

Despite these limitations, we believe that the findings of our study are important for psychology educators, textbook authors, and the field of psychology. Although introductory psychology texts incorporate many central concepts related to intelligence (e.g., intelligence testing, Spearman's g), they also include some nonmainstream topics—specifically Gardner's theory of multiple intelligences and Sternberg's triarchic theory of intelligence. These nonmainstream topics can be removed from textbooks, de-emphasized, or presented in a more critical fashion to improve the accuracy of textbooks.⁵ Moreover, an added emphasis on the Carroll three-stratum model and/or bifactor models would greatly strengthen many textbooks.

For those interested in knowing more about intelligence, we suggest starting with the Gottfredson (1997a) and Neisser et al. (1996) articles, which provide useful summaries of the scholarly literature through the mid-1990s. The short books by Deary (2001) and Ritchie (2015) are not only accurate, but they also have a breezy writing style that makes them easily digestible. Those willing to tackle lengthier tomes should consult Herrnstein and Murray's *The Bell Curve* (1996) or A. R. Jensen's (1998) book on g . Both books stand up well to the test of time and contain very little information that has since come into question by mainstream scholars. Some readers will also be surprised to find that *The Bell Curve* is not as controversial as its reputation would lead one to believe (and most of the book is not about race at all). More recent books that some may find helpful were written by Hunt (2011) and Mackintosh (2011), both of which could also serve as a textbook for an advanced undergraduate or basic graduate-level course on human intelligence. Readers interested in the historical predecessors and development of hierarchical models of intelligence should consult Carroll (1993, Chapter 2) for a technical explanation of various factor analysis models that found favor during the 20th century. Beaujean (2015) provides a useful historical perspective on bifactor models, and Canivez (2016) provides an accessible comparison of hierarchical and bifactor models, whereas Cucina and Byle (2017) make the same comparison across dozens of data sets. Lohman (1997) gives an evenhanded account of the philosophical and scientific debates that occupied historical figures in intelligence research, though modern debates in intelligence research are concerned with other matters (Gottfredson, 2009; Gottfredson & Saklofske, 2009). An article from Nisbett et al. (2012) provides a contrasting view to some of our opinions, especially in relation to genetics. Regardless of the resources that readers use to learn more about intelligence, a sincere perusal of the intelligence literature will help any reader learn more about this important psychological construct.

⁵ To facilitate these changes, we intend to send all textbook authors a copy of the article's Supplemental File 1, which lists the problematic statements from every textbook.

References

*References marked with an asterisk indicate textbooks included in the analysis.

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Psychological Association [APA]. (2014). *Strengthening the common core of the introductory psychology course*. Washington, DC: American Psychological Association, Board of Educational Affairs. Retrieved from <http://www.apa.org/ed/governance/bea/intro-psych-report.pdf>
- Anderson, B. (1993). Evidence from the rat for a general factor that underlies cognitive performance and that relates to brain size: Intelligence? *Neuroscience Letters*, *153*, 98–102. [http://dx.doi.org/10.1016/0304-3940\(93\)90086-Z](http://dx.doi.org/10.1016/0304-3940(93)90086-Z)
- Arden, R., & Adams, M. J. (2016). A general intelligence factor in dogs. *Intelligence*, *55*, 79–85. <http://dx.doi.org/10.1016/j.intell.2016.01.008>
- Barnard, C., Carmichael, L., French, T., Hunter, W., Thurstone, L., Bronk, D., . . . Whitehorn, J. (1970). The place of psychology in an ideal university: The report of the university commission to advise on the future of psychology at Harvard. *American Psychologist*, *25*, 391–410. <http://dx.doi.org/10.1037/h0029489>
- Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher-order models. *Journal of Intelligence*, *3*, 121–136. <http://dx.doi.org/10.3390/jintelligence3040121>
- Beaver, K. M., Barnes, J. C., & Boutwell, B. B. (2016). Exploring the relationship between violent behavior and participation in football during adolescence. *Youth & Society*, *48*, 786–809. <http://dx.doi.org/10.1177/0044118X13515279>
- 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. <http://dx.doi.org/10.1016/j.intell.2013.05.001>
- Belasen, A., & Hafer, R. W. (2013). IQ and alcohol consumption: International data. *Intelligence*, *41*, 615–621. <http://dx.doi.org/10.1016/j.intell.2013.07.019>
- *Bernstein, D. A. (2016). *Psychology: Foundations and frontiers*. Boston, MA: Cengage Learning.
- *Bonds-Raacke, J. (2014). *Psychology in a complex world*. Dubuque, IA: Kendall Hunt.
- Bouchard, T. J., Jr., Lykken, D. T., Tellegen, A., & McGue, M. (1996). Genes, drives, environment, and experience: EPD theory revised. In C. P. Benbow (Ed.), *Intellectual talent: Psychometric and social issues* (pp. 5–43). Baltimore, MD: Johns Hopkins University Press.
- *Cacioppo, J. T., & Freberg, L. A. (2016). *Discovering psychology: The science of the mind* (2nd ed.). Boston, MA: Cengage Learning.
- Canivez, G. L. (2016). Bifactor modeling in construct validation of multifaceted tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schwizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements* (pp. 247–271). Gottingen, Germany: Hogrefe.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511571312>
- *Ciccarelli, S. K., & White, J. N. (2015). *Psychology* (4th ed.). Upper Saddle River, NJ: Pearson.
- Cofnas, N. (2016). Science is not always “self-correcting”: Fact-value conflation and the study of intelligence. *Foundations of Science*, *21*, 477–492. <http://dx.doi.org/10.1007/s10699-015-9421-3>
- *Comer, R., & Gould, E. (2013). *Psychology around us*. Hoboken, NJ: Wiley.
- *Coon, D., & Mitterer, J. O. (2016). *Introduction to psychology: Gateways to mind and behavior* (14th ed.). Boston, MA: Cengage Learning.
- Cronbach, L. J. (1975). Five decades of public controversy over mental testing. *American Psychologist*, *30*, 1–14. <http://dx.doi.org/10.1037/0003-066X.30.1.1>
- Cucina, J., & Byle, K. (2017). The bifactor model fits better than the higher-order model in more than 90% of comparisons for mental abilities test batteries. *Journal of Intelligence*, *5*, 27. <http://dx.doi.org/10.3390/jintelligence5030027>
- Cucina, J. M., & Howardson, G. N. (2017). Woodcock-Johnson-III, Kaufman Adolescent and Adult Intelligence Test (KAIT), Kaufman Assessment Battery for Children (KABC), and Differential Ability Scales (DAS) support Carroll but not Cattell-Horn. *Psychological Assessment*, *29*, 1001–1015. <http://dx.doi.org/10.1037/pas0000389>
- Dalliard, M. (2014). The elusive x -factor: A critique of J. M. Kaplan's model of race and IQ. *Open Differential Psychology*, *27*. Retrieved from <https://openpsych.net/paper/27>

- Davies, G., Armstrong, N., Bis, J. C., Bressler, J., Chouraki, V., Giddaluru, S., . . . the Generation Scotland. (2015). Genetic contributions to variation in general cognitive function: A meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). *Molecular Psychiatry*, *20*, 183–192. <http://dx.doi.org/10.1038/mp.2014.188>
- Deary, I. J. (2001). *Intelligence: A very short introduction*. New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/actrade/9780192893215.001.0001>
- Deary, I. J. (2012). Intelligence. *Annual Review of Psychology*, *63*, 453–482. <http://dx.doi.org/10.1146/annurev-psych-120710-100353>
- Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11*, 201–211.
- Deary, I. J., Taylor, M. D., Hart, C. L., Wilson, V., Smith, G. D., Blane, D., & Starr, J. M. (2005). Intergenerational social mobility and mid-life status attainment: Influences of childhood intelligence, childhood social factors, and education. *Intelligence*, *33*, 455–472. <http://dx.doi.org/10.1016/j.intell.2005.06.003>
- 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. <http://dx.doi.org/10.1037/0022-3514.86.1.130>
- Detterman, D. K. (2014). You should be teaching intelligence! *Intelligence*, *42*, 148–151. <http://dx.doi.org/10.1016/j.intell.2013.07.021>
- Diener, E., Oishi, S., & Park, J. (2014). An incomplete list of eminent psychologists of the modern era. *Archives of Scientific Psychology*, *2*, 20–31. <http://dx.doi.org/10.1037/arc0000006>
- Dolan, C. V. (2000). Investigating Spearman's hypothesis by means of multi-group confirmatory factor analysis. *Multivariate Behavioral Research*, *35*, 21–50. http://dx.doi.org/10.1207/S15327906MBR3501_2
- Dolan, C. V., & Hamaker, E. L. (2001). Investigating Black-White differences in psychometric IQ: Multi-group confirmatory factor analyses of the WISC-R and K-ABC and a critique of the method of correlated vectors. In F. H. Columbus (Ed.), *Advances in psychology research* (Vol. 6, pp. 31–59). Huntington, NY: Nova Science Publishers.
- Edwards, A. W. F. (2003). Human genetic diversity: Lewontin's fallacy. *BioEssays*, *25*, 798–801. <http://dx.doi.org/10.1002/bies.10315>
- Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy*. New York, NY: Norton.
- *Feist, G. J., & Rosenberg, E. L. (2015). *Psychology: Perspectives and connections* (3rd ed.). New York, NY: McGraw-Hill Education.
- *Feldman, R. S. (2015). *Essentials of understanding psychology* (11th ed.). New York, NY: McGraw-Hill Education.
- Ferguson, C. J., Brown, J. M., & Torres, A. V. (in press). Education or indoctrination? The accuracy of introductory psychology textbooks in covering controversial topics and urban legends about psychology. *Current Psychology*.
- Fernandes, H. B. F., Woodley, M. A., & te Nijenhuis, J. (2014). Differences in cognitive abilities among primates are concentrated on G: Phenotypic and phylogenetic comparisons with two meta-analytical databases. *Intelligence*, *46*, 311–322. <http://dx.doi.org/10.1016/j.intell.2014.07.007>
- 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. <http://dx.doi.org/10.1016/j.jsp.2014.10.002>
- Freeman, F. N. (1923). A referendum of psychologists. *The Century*, *107*, 237–245.
- Frisby, C. L., & Beaujean, A. A. (2015). Testing Spearman's hypotheses using a bi-factor model with WAIS-IV/WMS-IV standardization data. *Intelligence*, *51*, 79–97. <http://dx.doi.org/10.1016/j.intell.2015.04.007>
- Galsworthy, M. J., Paya-Cano, J. L., Monleón, S., & Plomin, R. (2002). Evidence for general cognitive ability (g) in heterogeneous stock mice and an analysis of potential confounds. *Genes, Brain, & Behavior*, *1*, 88–95. <http://dx.doi.org/10.1034/j.1601-183X.2002.10204.x>
- 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. <http://dx.doi.org/10.1037/a0031412>
- Gardner, H. (1995). A response on four fronts. *Contemporary Psychology*, *40*, 938–939.
- *Gerrig, R. J. (2013). *Psychology and life* (20th ed.). Upper Saddle River, NJ: Pearson.
- Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence*, *55*, 57–68. <http://dx.doi.org/10.1016/j.intell.2016.01.006>
- *Gleitman, H., Gross, J., & Reisberg, D. (2011). *Psychology* (8th ed.). New York, NY: Norton.
- Gorenflo, D. W., & McConnell, J. V. (1991). The most frequently cited journal articles and authors in introductory psychology textbooks. *Teaching of Psychology*, *18*, 8–12. http://dx.doi.org/10.1207/s15328023top1801_2
- Gottfredson, L. S. (1994). Egalitarian fiction and collective fraud. *Society*, *31*, 53–59. <http://dx.doi.org/10.1007/BF02693231>
- Gottfredson, L. S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*, 13–23. [http://dx.doi.org/10.1016/S0160-2896\(97\)90011-8](http://dx.doi.org/10.1016/S0160-2896(97)90011-8)
- Gottfredson, L. S. (1997b). Why g matters: The complexity of everyday life. *Intelligence*, *24*, 79–132. [http://dx.doi.org/10.1016/S0160-2896\(97\)90014-3](http://dx.doi.org/10.1016/S0160-2896(97)90014-3)
- Gottfredson, L. S. (2003a). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence*, *31*, 343–397. [http://dx.doi.org/10.1016/S0160-2896\(02\)00085-5](http://dx.doi.org/10.1016/S0160-2896(02)00085-5)
- Gottfredson, L. S. (2003b). On Sternberg's "Reply to Gottfredson." *Intelligence*, *31*, 415–424. [http://dx.doi.org/10.1016/S0160-2896\(03\)00024-2](http://dx.doi.org/10.1016/S0160-2896(03)00024-2)
- Gottfredson, L. S. (2004). Intelligence: Is it the epidemiologists' elusive "fundamental cause" of social class inequalities in health? *Journal of Personality and Social Psychology*, *86*, 174–199. <http://dx.doi.org/10.1037/0022-3514.86.1.174>
- Gottfredson, L. S. (2005). What if the hereditarian hypothesis is true? *Psychology, Public Policy, and Law*, *11*, 311–319. <http://dx.doi.org/10.1037/1076-8971.11.2.311>
- Gottfredson, L. S. (2009). Logical fallacies used to dismiss the evidence on intelligence testing. In R. P. Phelps (Ed.), *Correcting fallacies about educational and psychological testing* (pp. 11–65). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/11861-001>
- Gottfredson, L. S. (2010). Lessons in academic freedom as lived experience. *Personality and Individual Differences*, *49*, 272–280. <http://dx.doi.org/10.1016/j.paid.2010.01.001>
- Gottfredson, L., & Saklofske, D. H. (2009). Intelligence: Foundations and issues in assessment. *Canadian Psychology/Psychologie Canadienne*, *50*, 183–195. <http://dx.doi.org/10.1037/a0016641>
- *Gray, P., & Bjorklund, D. F. (2014). *Psychology* (7th ed.). New York, NY: Worth Publishers.
- Griggs, R. A. (2014a). Topical coverage in introductory textbooks from the 1980s through the 2000s. *Teaching of Psychology*, *41*, 5–10. <http://dx.doi.org/10.1177/0098628313514171>
- *Griggs, R. A. (2014b). *Psychology: A concise introduction* (4th ed.). New York, NY: Worth Publishers.
- Griggs, R. A., & Marek, P. (2001). Similarity of introductory psychology textbooks: Reality or illusion? *Teaching of Psychology*, *28*, 254–256. http://dx.doi.org/10.1207/S15328023TOP2804_03
- Griggs, R. A., & Mitchell, M. C. (2002). In search of introductory psychology's classic core vocabulary. *Teaching of Psychology*, *29*, 144–147.
- Griggs, R. A., Proctor, D. L., & Bujak-Johnson, A. (2002). The nonexistent common core. *American Psychologist*, *57*, 452–453. <http://dx.doi.org/10.1037/0003-066X.57.6-7.452b>
- Griggs, R. A., Proctor, D. L., & Cook, S. M. (2004). The most frequently cited books in introductory texts. *Teaching of Psychology*, *31*, 113–116.
- *Grison, S., Heatherton, T. F., & Gazzaniga, M. S. (2017). *Psychology in your life*. New York, NY: Norton.
- Guilford, J. P. (1950). Creativity. *American Psychologist*, *5*, 444–454. <http://dx.doi.org/10.1037/h0063487>
- Gurven, M., Fuerstenberg, E., Trumble, B., Stieglitz, J., Beheim, B., Davis, H., & Kaplan, H. (2017). Cognitive performance across the life course of Bolivian forager-farmers with limited schooling. *Developmental Psychology*, *53*, 160–176. <http://dx.doi.org/10.1037/dev0000175>
- Habarth, J., Hansell, J., & Grove, T. (2011). How accurately do introductory psychology textbooks present psychoanalytic theory? *Teaching of Psychology*, *38*, 16–21. <http://dx.doi.org/10.1177/0098628310390912>

- Halonon, J. (2011). *Are there too many psychology majors?* Retrieved from <https://pdfs.semanticscholar.org/0bac/a87447cc5df53095cecb3a9299f7c428343.pdf>
- Haselhuhn, C. W., & Clopton, K. L. (2008). The representation of applied psychology areas in introductory psychology textbooks. *Teaching of Psychology, 35*, 205–209. <http://dx.doi.org/10.1177/009862830803500310>
- Hayden, E. C. (2013). Ethics: Taboo genetics. *Nature, 502*, 26–28. <http://dx.doi.org/10.1038/502026a>
- Hedges, L. V., & Nowell, A. (1999). Changes in the black-white gap in achievement test scores. *Sociology of Education, 72*, 111–135. <http://dx.doi.org/10.2307/2673179>
- Herndon, J. G., Moss, M. B., Rosene, D. L., & Killiany, R. J. (1997). Patterns of cognitive decline in aged rhesus monkeys. *Behavioural Brain Research, 87*, 25–34. [http://dx.doi.org/10.1016/S0166-4328\(96\)02256-5](http://dx.doi.org/10.1016/S0166-4328(96)02256-5)
- Herrnstein, R. J., & Murray, C. (1996). *The bell curve: Intelligence and class structure in American life* (2nd ed.). New York, NY: Free Press.
- *Hockenbury, S. E., Nolan, S. A., & Hockenbury, D. H. (2015). *Psychology* (7th ed.). New York, NY: Worth Publishers.
- Hogben, M., & Waterman, C. K. (1997). Are all of your students represented in their textbooks? A content analysis of coverage of diversity issues in introductory psychology textbooks. *Teaching of psychology, 24*, 95–100. http://dx.doi.org/10.1207/s15328023top2402_3
- Huang, P.-C., Su, P.-H., Chen, H.-Y., Huang, H.-B., Tsai, J.-L., Huang, H.-I., & Wang, S.-L. (2012). Childhood blood lead levels and intellectual development after ban of leaded gasoline in Taiwan: A 9-year prospective study. *Environment International, 40*, 88–96. <http://dx.doi.org/10.1016/j.envint.2011.10.011>
- Hunt, E. (2001). Multiple views of multiple intelligence [Review of *Intelligence reframed: Multiple intelligence in the 21st century*]. *Contemporary Psychology, 46*, 5–7.
- Hunt, E. (2011). *Human intelligence*. New York, NY: Cambridge University Press.
- Hunt, E. (2014). Teaching intelligence: Why, why it is hard and perhaps how to do it. *Intelligence, 42*, 156–165. <http://dx.doi.org/10.1016/j.intell.2013.06.018>
- Hunter, J. E., & Schmidt, F. L. (2000). Racial and gender bias in ability and achievement tests: Resolving the apparent paradox. *Psychology, Public Policy, and Law, 6*, 151–158. <http://dx.doi.org/10.1037/1076-8971.6.1.151>
- Jackson, S. L., & Griggs, R. A. (2013). Psych lite: Great price, less filling. *Teaching of Psychology, 40*, 217–221. <http://dx.doi.org/10.1177/0098628313487415>
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Jensen, R., & Burgess, H. (1997). Mythmaking: How introductory psychology texts present B. F. Skinner's analysis of cognition. *The Psychological Record, 47*, 221–232. <http://dx.doi.org/10.1007/BF03395221>
- Johnson, W., Bouchard, T. J., Jr., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence, 32*, 95–107. [http://dx.doi.org/10.1016/S0160-2896\(03\)00062-X](http://dx.doi.org/10.1016/S0160-2896(03)00062-X)
- Johnson, W., te Nijenhuis, J., & Bouchard, T. J., Jr. (2008). Still just 1 g: Consistent results from five test batteries. *Intelligence, 36*, 81–95. <http://dx.doi.org/10.1016/j.intell.2007.06.001>
- *Kalat, J. W. (2017). *Introduction to psychology* (11th ed.). Boston, MA: Cengage Learning.
- Karimi, F. (2016, July 25). Man with autism not sleeping, eating after North Miami shooting, family says. *CNN*. Retrieved from <http://www.cnn.com/2016/07/24/health/north-miami-police-charles-kinsey-shooting/>
- Kourany, J. A. (2016). Should some knowledge be forbidden? The case of cognitive differences research. *Philosophy of Science, 83*, 779–790. <http://dx.doi.org/10.1086/687863>
- *Lahey, B. B. (2012). *Psychology: An introduction* (11th ed.). New York, NY: McGraw-Hill.
- Lee, J. J. (2010). Review of *Intelligence and how to get it: Why schools and cultures count* [Book review]. *Personality and Individual Differences, 48*, 247–255. <http://dx.doi.org/10.1016/j.paid.2009.09.015>
- Levin, M. (1994). Comment on the Minnesota Transracial Adoption Study. *Intelligence, 19*, 13–20. [http://dx.doi.org/10.1016/0160-2896\(94\)90049-3](http://dx.doi.org/10.1016/0160-2896(94)90049-3)
- Lewontin, R. C. (1970). Race and intelligence. *Bulletin of the Atomic Scientists, 26*, 2–8. <http://dx.doi.org/10.1080/00963402.1970.11457774>
- Lewontin, R. C. (1972). The apportionment of human diversity. In T. Dobzhansky, M. K. Hecht, & W. C. Steere (Eds.), *Evolutionary biology* (Vol. 6, pp. 381–398). Boston, MA: Springer US. http://dx.doi.org/10.1007/978-1-4684-9063-3_14
- *Lilienfeld, S. O., Lynn, S. J., Namy, L. L., & Woolf, N. J. (2014). *Psychology: From inquiry to understanding* (3rd ed.). Upper Saddle River, NJ: Pearson Education.
- Lipsey, M. W., Farran, D. C., & Hofer, K. G. (2015). *A randomized control trial of a statewide voluntary prekindergarten program on children's skills and behaviors through third grade*. Nashville, TN: Vanderbilt University, Peabody Research Institute.
- Lohman, D. F. (1997). The history of intelligence testing in context: The impact of personal, religious, and scientific beliefs on the development of theories and tests of human abilities. In R. F. Dillon (Ed.), *Handbook on testing* (pp. 82–106). Westport, CT: Greenwood Press.
- Lohman, D. F. (2006). Beliefs about differences between ability and accomplishment: From folk theories to cognitive science. *Roeper Review, 29*, 32–40. <http://dx.doi.org/10.1080/02783190609554382>
- Lubinski, D., & Benbow, C. P. (1995). An opportunity for empiricism. *Contemporary Psychology, 40*, 935–938.
- Lubinski, D., & Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence, 24*, 159–201. [http://dx.doi.org/10.1016/S0160-2896\(97\)90016-7](http://dx.doi.org/10.1016/S0160-2896(97)90016-7)
- Lubke, G. H., Dolan, C. V., Kelderman, H., & Mellenbergh, G. J. (2003). On the relationship between sources of within- and between-group differences and measurement invariance in the common factor model. *Intelligence, 31*, 543–566. [http://dx.doi.org/10.1016/S0160-2896\(03\)00051-5](http://dx.doi.org/10.1016/S0160-2896(03)00051-5)
- Lynn, R. (2015). *Race differences in intelligence: An evolutionary analysis* (2nd ed.). Arlington, VA: Washington Summit Publishers.
- Mackintosh, N. J. (2011). *IQ and human intelligence* (2nd ed.). New York, NY: Oxford University Press.
- Marioni, R. E., Davies, G., Hayward, C., Liewald, D., Kerr, S. M., Campbell, A., . . . Deary, I. J. (2014). Molecular genetic contributions to socioeconomic status and intelligence. *Intelligence, 44*, 26–32. <http://dx.doi.org/10.1016/j.intell.2014.02.006>
- McGovern, T. V. (1992). Evolution of undergraduate curricula in psychology, 1892–1992. In A. E. Puente, J. R. Matthews, & C. L. Brewer (Eds.), *Teaching psychology in America: A history* (pp. 13–38). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10120-001>
- 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>
- Miller, B., & Gentile, B. F. (1998). Introductory course content and goals. *Teaching of Psychology, 25*, 89–96. http://dx.doi.org/10.1207/s15328023top2502_2
- Moore, E. G. (1986). Family socialization and the IQ test performance of traditionally and transracially adopted Black children. *Developmental Psychology, 22*, 317–326. <http://dx.doi.org/10.1037/0012-1649.22.3.317>
- *Morris, C. G., & Maisto, A. A. (2016). *Understanding psychology* (11th ed.). New York, NY: Pearson Education.
- Murray, C. (2007). The magnitude and components of change in the black-white IQ difference from 1920 to 1991: A birth cohort analysis of the Woodcock-Johnson standardizations. *Intelligence, 35*, 305–318. <http://dx.doi.org/10.1016/j.intell.2007.02.001>
- *Myers, D. G., & DeWall, C. N. (2015). *Psychology* (11th ed.). New York, NY: Worth Publishers.
- *Nairne, J. S. (2014). *Psychology* (6th ed.). Belmont, CA: Cengage Learning.
- National Center for Education Statistics. (2016). *2016 Digest of education statistics*. Retrieved from https://nces.ed.gov/programs/digest/2016menu_tables.asp
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77–101. <http://dx.doi.org/10.1037/0003-066X.51.2.77>
- *Nevid, J. S. (2015). *Essentials of psychology: Concepts and applications* (4th ed.). Stamford, CT: Cengage Learning.

- 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. <http://dx.doi.org/10.1037/a0026699>
- *Nolen-Hoeksema, S., Fredrickson, B. L., Loftus, G. R., & Lutz, C. (2014). *Atkinson & Hilgard's introduction to psychology* (16th ed.). Andover, UK: Cengage Learning.
- Nyborg, H., & Jensen, A. R. (2001). Occupation and income related to psychometric *g*. *Intelligence*, *29*, 45–55. [http://dx.doi.org/10.1016/S0160-2896\(00\)00042-8](http://dx.doi.org/10.1016/S0160-2896(00)00042-8)
- *Okami, P. (2014). *Psychology: Contemporary perspectives*. New York, NY: Oxford University Press.
- O'Toole, B. I., & Stankov, L. (1992). Ultimate validity of psychological tests. *Personality and Individual Differences*, *13*, 699–716. [http://dx.doi.org/10.1016/0191-8869\(92\)90241-G](http://dx.doi.org/10.1016/0191-8869(92)90241-G)
- *Pastorino, E., & Doyle-Portillo, S. D. (2016). *What is psychology? Foundations, applications, and integration*. Boston, MA: Cengage Learning.
- Perlman, B., & McCann, L. I. (1999). The most frequently listed courses in the undergraduate psychology curriculum. *Teaching of Psychology*, *26*, 177–182. <http://dx.doi.org/10.1207/S15328023TOP260303>
- Pesta, B. J., McDaniel, M. A., Poznanski, P. J., & DeGroot, T. (2015). Discounting IQ's relevance to organizational behavior: The "somebody else's problem" in management education. *Open Differential Psychology*, *35*, 1–11. Retrieved from https://openpsych.net/files/papers/Pesta_2015a.pdf
- Pickrell, J. K., Berisa, T., Liu, J. Z., Ségurel, L., Tung, J. Y., & Hinds, D. A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nature Genetics*, *48*, 709–717. <http://dx.doi.org/10.1038/ng.3570>
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science*, *11*, 3–23. <http://dx.doi.org/10.1177/1745691615617439>
- Plomin, R., & Petrill, S. A. (1997). Genetics and intelligence: What's new? *Intelligence*, *24*, 53–77. [http://dx.doi.org/10.1016/S0160-2896\(97\)90013-1](http://dx.doi.org/10.1016/S0160-2896(97)90013-1)
- Proctor, B. D., Semega, J. L., & Mollar, M. A. (2016). Income and poverty in the United States: 2015 (U.S. Census Bureau Current Populations Reports No. P60-256(RV)). Washington, DC: U.S. Government Printing Office. Retrieved from <https://www.census.gov/content/dam/Census/library/publications/2016/demo/p60-256.pdf>
- Protzko, J. (2015). The environment in raising early intelligence: A meta-analysis of the fadeout effect. *Intelligence*, *53*, 202–210. <http://dx.doi.org/10.1016/j.intell.2015.10.006>
- Pyryt, M. C. (2000). Finding "g": Easy viewing through higher order factor analysis. *Gifted Child Quarterly*, *44*, 190–192. <http://dx.doi.org/10.1177/001698620004400305>
- *Rathus, S. A. (2016). *PSYCH* (4th ed.). Boston, MA: Cengage Learning.
- Reeve, C. L., & Basalik, D. (2014). Is health literacy an example of construct proliferation? A conceptual and empirical evaluation of its redundancy with general cognitive ability. *Intelligence*, *44*, 93–102. <http://dx.doi.org/10.1016/j.intell.2014.03.004>
- Reuning, H. (1972). Psychological studies of Kalahari Bushmen. In L. J. Cronbach & P. J. D. Drenth (Eds.), *Mental tests and cultural adaptation* (pp. 171–181). The Hague, The Netherlands: Mouton.
- Ritchie, S. (2015). *Intelligence: All that matters*. London, UK: John Murray Learning.
- Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., III, & Tyler, P. (2001). Ethnic group difference in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, *54*, 297–330. <http://dx.doi.org/10.1111/j.1744-6570.2001.tb00094.x>
- Rowe, D. C., & Cleveland, H. H. (1996). Academic achievement in Blacks and Whites: Are the developmental processes similar? *Intelligence*, *23*, 205–228. [http://dx.doi.org/10.1016/S0160-2896\(96\)90004-5](http://dx.doi.org/10.1016/S0160-2896(96)90004-5)
- Rowe, D. C., Vazsonyi, A. T., & Flannery, D. J. (1994). No more than skin deep: Ethnic and racial similarity in developmental process. *Psychological Review*, *101*, 396–413. <http://dx.doi.org/10.1037/0033-295X.101.3.396>
- Rowe, D. C., Vazsonyi, A. T., & Flannery, D. J. (1995). Ethnic and racial similarity in developmental process: A study of academic achievement. *Psychological Science*, *6*, 33–38. <http://dx.doi.org/10.1111/j.1467-9280.1995.tb00301.x>
- Rushton, J. P. (2012). No narrowing in mean Black-White IQ differences—Predicted by heritable *g*. *American Psychologist*, *67*, 500–501. <http://dx.doi.org/10.1037/a0029614>
- Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*, 235–294. <http://dx.doi.org/10.1037/1076-8971.11.2.235>
- Ryan, C. L., & Bauman, K. (2016). *Education attainment in the United States: 2015*. (U.S. Census Bureau Current Populations Reports No. P20-578). Washington, DC: U.S. Government Printing Office. Retrieved from <https://www.census.gov/content/dam/Census/library/publications/2016/demo/p20-578.pdf>
- 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*, 7–13. <http://dx.doi.org/10.1037/0003-066X.59.1.7>
- Sariaslan, A., Larsson, H., & Fazel, S. (2016). Genetic and environmental determinants of violence risk in psychotic disorders: A multivariate quantitative genetic study of 1.8 million Swedish twins and siblings. *Molecular Psychiatry*, *21*, 1251–1256. <http://dx.doi.org/10.1038/mp.2015.184>
- Scarr, S., & Weinberg, R. A. (1976). IQ test performance of Black children adopted by White families. *American Psychologist*, *31*, 726–739. <http://dx.doi.org/10.1037/0003-066X.31.10.726>
- *Schacter, D. L., Gilbert, D. T., Wegner, D. M., & Nock, M. K. (2014). *Psychology* (3rd ed.). New York, NY: Worth Publishers.
- 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. <http://dx.doi.org/10.1037/0033-2909.124.2.262>
- 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. <http://dx.doi.org/10.1037/0022-3514.86.1.162>
- Schneider, W. J. (2016). Why are WJ IV cluster scores more extreme than the average of their parts? A gentle explanation of the composite score extremity effect. Itasca, IL: Houghton Mifflin Harcourt. Retrieved from http://www.hmco.com/~media/sites/home/hmh-assessments/clinical/woodcock-johnson/pdf/wjiv/wjiv_asb_7.pdf
- Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest*, *17*, 103–186. <http://dx.doi.org/10.1177/1529100616661983>
- Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Macmillan.
- Smouse, P. E., Spielman, R. S., & Park, M. H. (1982). Multiple-locus allocation of individuals to groups as a function of the genetic variation within and differences among human populations. *American Naturalist*, *119*, 445–463. <http://dx.doi.org/10.1086/283925>
- Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist*, *42*, 137–144. <http://dx.doi.org/10.1037/0003-066X.42.2.137>
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York, NY: The Macmillan Company.
- 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. <http://dx.doi.org/10.1037/0022-3514.69.5.797>
- Sternberg, R. J. (2003). Our research program validating the triarchic theory of successful intelligence: Reply to Gottfredson. *Intelligence*, *31*, 399–413. [http://dx.doi.org/10.1016/S0160-2896\(02\)00143-5](http://dx.doi.org/10.1016/S0160-2896(02)00143-5)
- Sternberg, R. J. (2005). There are no public-policy implications: A reply to Rushton and Jensen (2005). *Psychology, Public Policy, and Law*, *11*, 295–301. <http://dx.doi.org/10.1037/1076-8971.11.2.295>
- Sternberg, R. J., & Hedlund, J. (2002). Practical intelligence, *g*, and work psychology. *Human Performance*, *15*, 143–160. <http://dx.doi.org/10.1080/08959285.2002.9668088>

- Steuer, F. B., & Ham, K. W., II. (2008). Psychology textbooks: Examining their accuracy. *Teaching of Psychology, 35*, 160–168. <http://dx.doi.org/10.1177/009862830803500302>
- Stoloff, M., McCarthy, M., Keller, L., Varfolomeeva, V., Lynch, J., Makara, K., . . . Smiley, W. (2010). The undergraduate psychology major: An examination of structure and sequence. *Teaching of Psychology, 37*, 4–15. <http://dx.doi.org/10.1080/00986280903426274>
- Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*, 401–426. <http://dx.doi.org/10.1016/j.intell.2006.09.004>
- 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. <http://dx.doi.org/10.1177/1529100611418056>
- The 1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature, 526*, 68–74. <http://dx.doi.org/10.1038/nature15393>
- The Chimpanzee Sequencing and Analysis Consortium. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature, 437*, 69–87. <http://dx.doi.org/10.1038/nature04072>
- Thoday, J. M. (1969). Limitations to genetic comparison of populations. *Journal of Biosocial Science, 1*, 3–14. <http://dx.doi.org/10.1017/S002193200002318X>
- Thomas, R. K. (2007). Recurring errors among recent history of psychology textbooks. *The American Journal of Psychology, 120*, 477–495.
- Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., . . . Williams, S. M. (2009). The genetic structure and history of Africans and African Americans. *Science, 324*, 1035–1044. <http://dx.doi.org/10.1126/science.1172257>
- Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., . . . Plomin, R. (2014). Genetic influence on family socioeconomic status and children's intelligence. *Intelligence, 42*, 83–88. <http://dx.doi.org/10.1016/j.intell.2013.11.002>
- U.S. Department of Health and Human Services. (2010). *Head Start impact study. Final report*. Washington, DC: U.S. Department of Health and Human Services Administration for Children and Families Office of Planning, Research and Evaluation.
- *Wade, C., Tavis, C., & Garry, M. (2014). *Psychology* (11th ed.). Upper Saddle River, NJ: Pearson Education.
- Walker, M. E., & Bridgeman, B. (2008). *Stereotype threat spillover and SAT scores* (Research Report No. 2008–2). New York, NY: College Board.
- Warne, R. T. (2016). Five reasons to put the *g* back into giftedness: An argument for applying the Cattell–Horn–Carroll theory of intelligence to gifted education research and practice. *Gifted Child Quarterly, 60*, 3–15. <http://dx.doi.org/10.1177/0016986215605360>
- Warne, R. (2017). *Introductory psychology textbooks' coverage of intelligence research* (ICPSR36957-v1) [Data set]. Ann Arbor, MI: Inter-University Consortium for Political and Social Research. <http://dx.doi.org/10.3886/ICPSR36957.v1>
- Warne, R. T., Yoon, M., & Price, C. J. (2014). Exploring the various interpretations of “test bias.” *Cultural Diversity and Ethnic Minority Psychology, 20*, 570–582. <http://dx.doi.org/10.1037/a0036503>
- Waterhouse, L. (2006). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. *Educational Psychologist, 41*, 207–225. http://dx.doi.org/10.1207/s15326985ep4104_1
- Wayne, R. K., & Ostrander, E. A. (1999). Origin, genetic diversity, and genome structure of the domestic dog. *BioEssays, 21*, 247–257. [http://dx.doi.org/10.1002/\(sici\)1521-1878\(199903\)21:3<247::aid-bies9>3.0.co;2-z](http://dx.doi.org/10.1002/(sici)1521-1878(199903)21:3<247::aid-bies9>3.0.co;2-z)
- Weinberg, R. A., Scarr, S., & Waldman, I. D. (1992). The Minnesota Transracial Adoption Study: A follow-up of IQ test performance at adolescence. *Intelligence, 16*, 117–135. [http://dx.doi.org/10.1016/0160-2896\(92\)90028-P](http://dx.doi.org/10.1016/0160-2896(92)90028-P)
- *Weiten, W. (2017). *Psychology: Themes and variations* (10th ed.). Boston, MA: Cengage Learning.
- Williams, W. M., & Ceci, S. J. (1997). Are Americans becoming more or less alike? Trends in race, class, and ability differences in intelligence. *American Psychologist, 52*, 1226–1235. <http://dx.doi.org/10.1037/0003-066X.52.11.1226>
- Woodley, M. A., & Meisenberg, G. (2012). Ability differentials between nations are unlikely to disappear. *American Psychologist, 67*, 501–502. <http://dx.doi.org/10.1037/a0029650>
- Yang, J., Bakshi, A., Zhu, Z., Hemani, G., Vinkhuyzen, A. A. E., Lee, S. H., . . . the LifeLines Cohort Study. (2015). Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nature Genetics, 47*, 1114–1120. <http://dx.doi.org/10.1038/ng.3390>
- Yerkes, R. M. (Ed.). (1921). *Psychological examining in the United States army*. Washington, DC: U.S. Government Printing Office.
- Yung, Y.-F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika, 64*, 113–128. <http://dx.doi.org/10.1007/BF02294531>
- Zechmeister, J. S., & Zechmeister, E. B. (2000). Introductory textbooks and psychology's core concepts. *Teaching of Psychology, 27*, 6–11. http://dx.doi.org/10.1207/S15328023TOP2701_1
- *Zimbardo, P., Johnson, R., & McCann, V. (2017). *Psychology: Core concepts* (8th ed.). Hoboken, NJ: Pearson Education.

Received March 27, 2017

Revision received October 10, 2017

Accepted October 16, 2017 ■