Intelligence 92 (2022) 101651

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

The sexes do not differ in general intelligence, but they do in some specifics

Matthew R. Reynolds [a,*], Daniel B. Hajovsky [b], Jacqueline M. Caemmerer [c]

[a] The University of Kansas, Lawrence, KS, USA
[b] Texas A&M University, College Station, TX, USA
[c] University of Connecticut, Storrs, CT, USA

ARTICLE INFO

Keywords:
Sex differences
Cattell-Horn-Carroll theory
Intelligence
Human cognitive abilities

ABSTRACT

Reliable and meaningful sex differences exist in specific cognitive abilities despite no reliable or meaningful sex difference in general intelligence. Here we use Cattell-Horn-Carroll (CHC) theory to highlight research findings related to sex differences in intelligence, with a focus on studies of test scores from comprehensive intelligence measures that were obtained from large and representative samples of children and adolescents. Female advantages in latent processing speed and male advantages in latent visual processing are the most meaningful and consistently reported sex differences regarding CHC broad cognitive abilities. Differences have been reported in narrow and specific ability constructs such as mental rotation and object memory location. In academic achievement, the largest and most consistent findings are female advantages in writing, whereas male advantages at higher math ability levels are also found. Empirical descriptions of sex differences should consider the breadth of the construct under study and incorporate analysis beyond simple mean differences. Score analysis methods that utilize multiple-group confirmatory factor models and multiple-indicator multiple cause models are useful to address the former, and analysis methods such as quantile regression and male-female ratio calculations along score distributions are useful to address the latter. An understanding of why specific ability differences exist in combination and in the presence of similarities will improve researchers' understanding of human cognition and educational achievements.

1. Introduction

The study of sex differences in psychological variables dates back over 100 years (Hedges & Nowell, 1995; Maccoby & Jacklin, 1974; Woolley, 1914).[1] Research findings are sometimes misunderstood. Generalizations from those findings are provocative (Ceci & Williams, 2007). Because most effect sizes used to quantify mean sex differences in psychological variables are described as either non-existent or "small," the Gender Similarities Hypothesis has been promoted to replace a "differences" hypothesis (GSH; Hyde, 2005, 2014).[2]

The GSH applies to intelligence and academic achievement variables, broadly speaking. The sexes are similar in general intelligence (Jensen, 1998); however, consistent differences emerge in some specific abilities (Reynolds, Scheiber, Hajovsky, Schwartz, & Kaufman, 2015; Zell, Krizan, & Teeter, 2015). If there is no sex difference in general intelligence, is it worth studying sex differences in specific abilities? We think the answer is yes (cf. Archer, 2019; Hyde, 2014). Findings may shed light on human cognition and have implications for understanding educational progress and achievements. In this commentary we highlight sex differences in specific abilities, especially from findings related to broad cognitive abilities, discuss important matters in such research, and provide suggestions for future research. We focus on empirical *descriptions* because it has been a focus of our research. We do not address potential *explanations* for sex similarities or differences, such as those derived from evolutionary (Archer, 2019; Geary, 2021) or sociocultural (Hyde, 2007, 2014) theories, among others (see Eagly & Wood, 2013).

---

* Corresponding author at: Department of Educational Psychology, University of Kansas, 1122 W Campus Rd., Lawrence, KS 66045, USA.
E-mail address: mreynolds@ku.edu (M.R. Reynolds).

[1] We use the term sex in this commentary to indicate biological sex assignment at birth (APA, 2012). See Reilly, Neumann, and Andrews (2016) for a discussion of sex and sex-role differences in specific cognitive abilities research.

[2] Some effect sizes may not meet Hyde's medium and large thresholds, but effect size interpretation is subjective. A small effect size may not be noticeable to the naked eye at the individual level, but it may produce large differences at the societal level. Notably, effect sizes described as small by Hyde (2005) $d = 0.11–0.35$ are classified as large ($d = 0.20$) regarding the effectiveness of educational interventions (Kraft, 2020).

https://doi.org/10.1016/j.intell.2022.101651
Received 31 January 2022; Received in revised form 8 April 2022; Accepted 10 April 2022

2. IQs and *g*

Substantial and reliable mean differences between the sexes have not materialized in either general intelligence (psychometric *g*)[3] or intelligence in general (e.g., IQs). This finding is germane to the study of sex differences in specific abilities. Hypothetically, *g* differences could produce sex differences in measures of specific abilities even when no differences exist in specific abilities. Likewise, specific ability differences could produce mean differences in IQs or various aggregates, and subsequently these findings might be interpreted as *g* differences when no *g* differences exist. Specific ability differences often cancel out during score aggregation, but if a test battery weighs heavily toward specific abilities that favor one sex, composite differences emerge (see Keith, Reynolds, Patel, & Ridley, 2008 for examples). Composites, however, should not be confused with *g* (Jensen, 1998). Similarly, if a single test is used to measure *g*, sex differences in those scores may be due to specific abilities measured by that test, and not *g*. In consideration of these caveats, mean sex differences in *g* or IQs based on data from comprehensive measures of intelligence administered to large and representative samples do not consistently favor either sex (Deary, Thorpe, Wilson, Starr, & Whalley, 2003; Irwing, 2012; Jensen, 1998; Jensen & Reynolds, 1983; Keith et al., 2008; Keith, Reynolds, Roberts, Winter, & Austin, 2011; Strand, Deary, & Smith, 2006).[4] No *g* difference, however, does not equate to no specific ability difference.

3. Broad and narrow/specific cognitive abilities

We reference Cattell-Horn-Carroll (CHC) theory (Carroll, 1993; Horn & Noll, 1997; Schneider & McGrew, 2018) to highlight and describe sex differences in specific abilities. CHC theory is widely used for intelligence research, test development, and assessment (Caemmerer, Keith, & Reynolds, 2020; McGrew, 2009). CHC theory is useful for classifying and organizing cognitive abilities (Reynolds, Keith, Flanagan, & Alfonso, 2013). Likewise, it is useful for organizing research findings. In this hierarchical and multidimensional framework, general intelligence (*g*) sits at the top. Below *g* are eight or more broad cognitive abilities (e.g., visual processing, processing speed) representing a range of interrelated yet independent intellectual capacities. Grouped together within each broad ability are various narrow abilities, and more than 80 have been identified (e.g., induction, associative memory, lexical knowledge). Finally, at the very bottom are specific abilities often defined by test features (e.g., digits forward, memory for names; Schneider & McGrew, 2018).

We comment on sex differences regarding seven CHC broad cognitive abilities (processing speed, visual processing, comprehension-knowledge, short-term/working memory, long-term retrieval, fluid reasoning, and auditory processing) and several narrow and specific abilities that they subsume. Comments regarding CHC factors are based on findings from studies that utilized latent variables for studying sex differences in multiple CHC-like broad cognitive abilities with children and adolescents. These studies accounted for both the hierarchical and multi-dimensional nature of psychometric intelligence when studying sex differences via either bifactor or higher-order factor models. Although these factor models represent different conceptions of intelligence, both include a general factor with broad ability factors, which was most important to our summary.

We identified seven studies from which to summarize findings. Five studies used norming data from prominent individually administered intelligence test batteries (Woodcock Johnson Tests of Cognitive Abilities [Keith et al., 2008], Kaufman Assessment Battery for Children [Reynolds, Keith, Ridley, & Patel, 2008], Wechsler preschool and school age intelligence scales [Palejwala & Fine, 2015; Pezzuti & Orsini, 2016], and the Differential Ability Scales [Keith et al., 2011]). Four studies were based on US norms samples, and one was based on an Italian norm sample (Pezzuti & Orsini, 2016). Two additional studies used data from samples of examinees who took test batteries administered in Sweden (Härnqvist, 1997; Rosén, 1995). Some test batteries were designed explicitly with a CHC measurement model (i.e., Kaufman tests; Woodcock Johnson tests), some have scoring structures that align with CHC theory (e.g., recent Wechsler revisions; Differential Ability Scales), and others were designed to measure abilities similar to Thurstone's primary mental abilities (Härnqvist, 1997) or CHC theory (Rosén, 1995).[5] Five studies included children and adolescents (Härnqvist, 1997; Keith et al., 2008, 2011; Pezzuti & Orsini, 2016; Reynolds et al., 2008), one included children only (Rosén, 1995), and one included younger children aged 2 to 7 (Palejwala & Fine, 2015). Four studies used higher-order models (Palejwala & Fine, 2015; Pezzuti & Orsini, 2016; Reynolds et al., 2008), two used bifactor models (Härnqvist, 1997; Rosén, 1995), and one used both (Keith et al., 2008).[6] Some studies found age moderation, but we averaged effects across age unless the direction of effects changed. In addition to findings related to CHC broad factors, we include narrow and specific ability findings from these studies and others, relying heavily on recent meta-analyses or syntheses of meta-analyses for the "other" studies (e.g., Archer, 2019; Asperholm, Högman, Rafi, & Herlitz, 2019; Voyer, Postma, Brake, & Imperato-McGinley, 2007; Voyer, Saint Aubin, Altman, & Gallant, 2021; Voyer, Voyer, & Saint-Aubin, 2017; Zell et al., 2015). Narrow and specific ability findings are samples of possible differences in the plethora of narrow and specific abilities. These differences may or not be explained by CHC broad ability differences. See Fig. 1.

This summary is not an all-inclusive comprehensive review, but an attempt to report commonly found differences within the CHC framework, and especially in view of CHC broad abilities. Of the seven CHC broad cognitive abilities described here, only latent processing speed and latent visual processing show consistent differences. Although these findings are consistent, we also describe why they are incomplete.

3.1. *Processing speed*

Females scored higher on processing speed factors across all five studies that included such a factor (Härnqvist, 1997; Keith et al., 2008; Keith et al., 2011; Palejwala & Fine, 2015; Pezzuti & Orsini, 2016). Processing speed (G*s*) is the ability to perform simple and repetitive cognitive tasks quickly within the span of a few minutes (Schneider & McGrew, 2018). Average standardized Cohen's *d* effect sizes from these

---

[3] As opposed to "psychological *g*" (see Fried, 2020), psychometric *g* used here refers to the general factor measured in scores from a large battery of diverse cognitive ability tests that were administered to a sample representative of the population (Jensen, 1998).

[4] Although no *g* difference seems to be the consensus, five out of seven latent variable studies summarized here showed small female advantages in *g* along with a profile of sex differences in more specific factors in children and adolescents (Härnqvist, 1997; Keith et al., 2008; Palejwala & Fine, 2015; Reynolds et al., 2008; Rosén, 1995). Lynn (1999) proposed a developmental model of sex differences in intelligence where females have very small *g* advantages from ages 9–12 that change to small male advantages at about age 16 and into adulthood. Empirical findings related to this theory are mixed (cf. Arribas-Aguila, Abad, & Colom, 2019; Keith et al., 2008; Reynolds et al., 2008), but age is potentially a moderator.

[5] We found additional studies that used latent variable analysis and considered a hierarchical structure. However, these studies were not included here because they did not include a variety of abilities (e.g., fluid intelligence only; Lankin & Gambrell, 2014) or were based on older measurement models with factors that included tests that are no longer associated with those factors (e.g., Verbal, Performance, and memory; van der Sluis et al., 2006). We did not include studies that were based on data from adults only.

[6] Effects in the Keith et al., 2008 study were similar in direction across higher-order and bifactor models, but effect sizes were generally larger with the bifactor models.

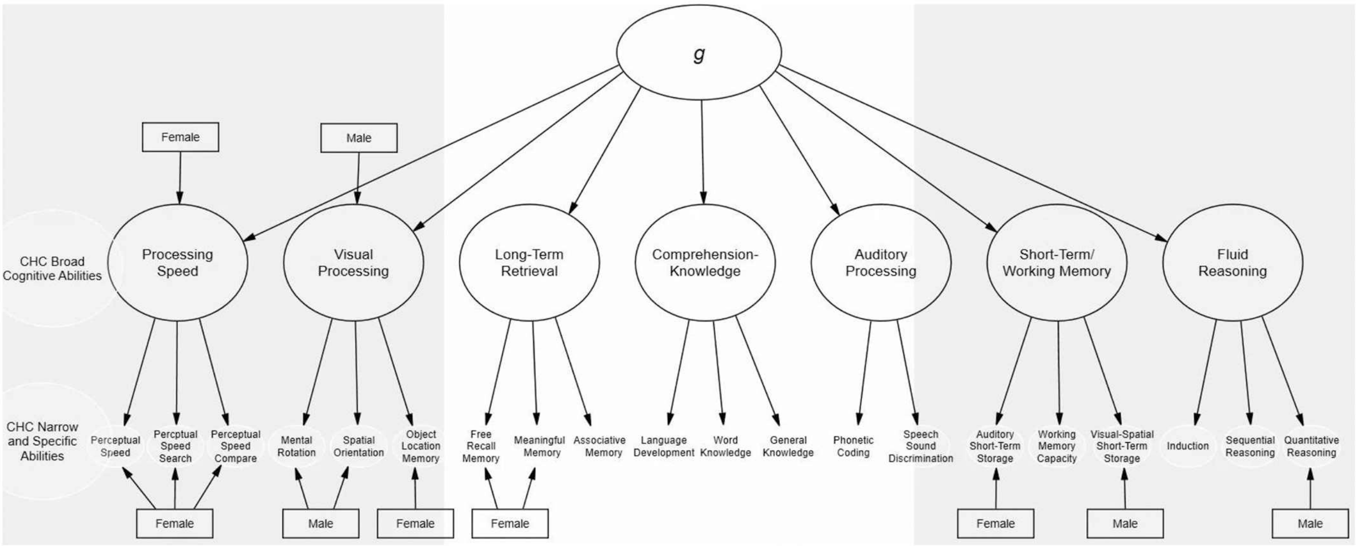

Fig. 1. A representation of sex differences in CHC broad abilities and a sample of potential differences in narrow and specific abilities.

studies ranged from 0.11–0.38 (*mdn* = 0.28) in favor of females. Findings suggest a difference in latent broad processing speed, and not simply a difference on a single test or task, especially since different tasks are used to measure latent processing speed both within and between test batteries. Female advantages in specific processing speed abilities measured by single tests have also been found (Jensen & Reynolds, 1983; van der Sluis et al., 2006), but it is unclear if these advantages are due to advantages in latent broad processing speed, a narrow/specific processing speed ability, or both. Studies with more processing speed tests per factor are needed. Nevertheless, females consistently show advantages in processing speed factors (Härnqvist, 1997; Keith et al., 2008; Keith et al., 2011; Palejwala & Fine, 2015; Pezzuti & Orsini, 2016) and in processing speed, perceptual speed, clerical speed, and rate-of-test taking measures found on psychometric intelligence tests (Camarata & Woodcock, 2006; Hedges & Nowell, 1995; Jensen & Reynolds, 1983; Johnson & Bouchard, 2007; Kaufman, Raiford, & Coalson, 2016; Roivianen, 2011).

Two additional points related to latent processing speed are needed. First, all but one finding related to processing speed factors summarized here involved whatever is in common among tests that require speed, accuracy, *and* fine motor movements. These processing speed tests are often the only "paper and pencil" tasks that remain on modern intelligence tests–the Wechsler preschool scale does not include paper and pencil, rather children mark with an ink dauber. Example paper and pencil processing speed tasks include identifying and drawing symbols in a specified box based on a key, locating and circling two numbers or letters that are alike in a row, and identifying target symbols in a row of symbols and crossing them out, all within a specific time limit. These tasks likely measure graphomotor speed in addition to what is considered processing speed. Comparisons of paper and pencil performances with newer digital administration of processing speed tests may help tease out the graphomotor component. For example, females show advantages on the traditional paper and pencil Coding test from the Wechsler scales (Jensen & Reynolds, 1983; van der Sluis et al., 2006). The traditional format requires examinees to draw symbols in boxes, whereas the newer digital format administration only requires examinees to tap the symbols (Wahlstrom, Raiford, Breaux, Zhu, & Weiss, 2018). If the digital task produces smaller sex differences, or even eliminates them, graphomotor speed may play a role in the processing speed difference.

The second point related to processing speed is how those findings fit within a recently proposed hierarchical model of speeded abilities (Schneider & McGrew, 2018). In this model, a higher-order "general speediness factor" subsumes broad psychomotor speed, broad decision speed, broad cognitive processing speed, and broad retrieval fluency. Broad cognitive processing speed comprises perceptual speed—essentially equivalent to the processing speed factor discussed so far—and academic speed. Females show advantages in perceptual speed. Females also show advantages in academic speed related to reading but not math (Camarata & Woodcock, 2006; Pargulski & Reynolds, 2017). A review of sex differences research in processing speed by Roivianen (2011), however, showed that females do not evince advantages in at least two of the other factors subsumed by the general speediness factor. Sex differences were absent from research with inspection time tasks and male advantages were found in some reaction time measures, both of which are associated with the proposed broad decision speed factor. Similarly, findings involving specific aspects of the proposed psychomotor speed factor were mixed, with male advantages in movement time, but female advantages in writing speed. Additional research is needed to support the validity of a hierarchy of speeded abilities, but the application of this framework may help to more systematically account for content, graphomotor requirements, and complexity differences across "processing speed" tasks when investigating sex differences (see Schneider & McGrew, 2018).

3.2. *Visual processing*

Sex differences in visual processing mostly favor males, but not always, and even within the same data, it may depend on how it is defined (Camarata & Woodcock, 2006; Keith et al., 2008; Rosén, 1995). Visual processing (G*v*) is the ability to use mental imagery to solve problems with visual information (Schneider & McGrew, 2018). Male advantages in latent visual processing with meaningful effect sizes (range 0.16–0.97; *Mdn* = 0.24) were found in all six of the studies with visual processing factors summarized here (Härnqvist, 1997; Keith et al., 2008; Keith et al., 2011; Palejwala & Fine, 2015; Reynolds et al., 2008; Rosén, 1995). Put differently, these six studies had identifiable visual processing factors that accounted for sex differences among visual processing subtests. One study included many visual processing tests, and beyond the broad visual processing factor there were male advantages in specific speed of closure and spatial orientation factors, but not in a visualization factor (Rosén, 1995; cf. Linn & Petersen, 1985; Malanchini et al., 2020). The largest sex difference in favor of males has been reported on tests of mental rotation, a specific ability (Archer, 2019; Linn & Petersen, 1985; Maeda & Yoon, 2013; Zell et al., 2015). It is unlikely that a male advantage in broad visual processing fully explains a male advantage in mental rotation since differences in those tests exceed all other findings. On another narrow visual processing ability, visual memory, males show advantages on tasks that involve remembering non-nameable images and routes (Asperholm et al., 2019), whereas females show advantages on tasks that ask for object location memory[7] (Archer, 2019; Asperholm et al., 2019; Voyer et al., 2017).

There is much to learn about visual processing that goes beyond what was summarized here (e.g., Buckley, Seery, & Canty, 2018; Malanchini et al., 2020). Traditional intelligence tests do not measure at least one potentially key component of visual processing, large-scale navigation (Schneider & McGrew, 2018). Strong correlations have been reported between navigation, object manipulation, and visualization spatial ability factors, with a higher-order spatial ability factor hypothesized to partially account for those correlations (Malanchini et al., 2020). Most sex differences research with intelligence tests is limited to object manipulation and visualization components. Outside of tests typically administered in psychometric studies, males have shown advantages in other related areas such as targeting (Jardine & Martin, 1983), maze learning (Asperholm et al., 2019; Moffat, Hampson, & Hatzipantelis, 1998), and navigation (Malanchini et al., 2020; Nazareth, Huang, Voyer, & Newcombe, 2019). One advantage of psychometric research is that large samples of individuals may be tested efficiently, but one disadvantage is that the tests used in these studies limit how abilities may be measured. Technological advances with virtual reality should allow for larger scale studies of visual processing in more real world type settings (see Malanchini et al., 2020), thus providing a more comprehensive understanding of visual processing, which in turn should provide a better understanding of sex similarities or differences in visual processing.

3.3. *Long-term retrieval*

Of the studies we summarized, only two tested for differences in long-term retrieval, and neither found a sex difference (Keith et al., 2008; Reynolds et al., 2008). Long-term retrieval is defined as the ability to store, consolidate, and retrieve information over time. However, a recent study showed that the ability should split into distinct learning efficiency (G*l*) and retrieval fluency (G*r*) abilities (Jewsbury & Bowden, 2017). The distinction has been supported in additional research (Agelink van Rentergem et al., 2020; Hajovsky et al., 2018; Meyer, & Reynolds, 2018), and CHC theory has been updated to acknowledge the distinction (Schneider & McGrew, 2018). Females have demonstrated advantages in two narrow abilities (free-recall memory and meaningful memory) associated with learning efficiency (Asperholm et al., 2019; Hedges & Nowell, 1995; Johnson & Bouchard, 2007; Keith et al., 2011; Lowe, Mayfield, & Reynolds, 2003), which is the ability to learn, store, and consolidate new information over periods of time (Schneider & McGrew, 2018).

3.4. *Comprehension-knowledge*

Findings regarding sex differences in comprehension-knowledge are inconsistent. Comprehension-knowledge is the breadth and depth of acquired knowledge and language learned inside and outside of school (often referred to as crystallized intelligence; Schneider & McGrew, 2018). In the studies reviewed here that included comprehension-knowledge (or verbal) factors, two found no sex differences (Keith et al., 2011; Palejwala & Fine, 2015), four found male advantages (Härnqvist, 1997; Keith et al., 2008; Pezzuti & Orsini, 2016; Reynolds et al., 2008), and one found a female advantage (Rosén, 1995). Other findings related to comprehension-knowledge abilities are also mixed. Male advantages have been reported on measures of vocabulary, verbal analogies, and general knowledge (Jensen & Reynolds, 1983; Lynn, Irwing, & Cammock, 2001; Lynn, Fergusson, & Horwood, 2005; van der Sluis et al., 2008), whereas a recent large-scale synthesis reported female advantages in general verbal and language abilities (Archer, 2019). Inconsistent findings are likely due in part to small effect sizes, differential item functioning (Steinmayr, Bergold, Margraf-Stiksrud, & Freund, 2015), and differences in the operationalization and measurement of comprehension-knowledge across studies. Some researchers equate comprehension-knowledge with verbal ability, others include reading and writing and equate it with language, and others equate it with tests of domain specific knowledge (see Kan, Kievit, Dolan, & van der Maas, 2011; Keith & Reynolds, 2010; Schneider & McGrew, 2018).

3.5. *Short-term memory, fluid reasoning, and auditory processing*

There are either negligible or inconsistent sex differences in latent short-term or working memory (G*sm or Gwm*), fluid reasoning (G*f*), and auditory processing (G*a*) (Keith et al., 2008; Keith et al., 2011; Palejwala & Fine, 2015; Reynolds et al., 2008; Rosén, 1995). An auditory processing factor was included in only one study that we reviewed, and there was not a sex difference (Keith et al., 2008). Only one of six studies we summarized found a short-term/working memory difference, with a small female advantage from ages 5–13 that changed to a small male advantage at around age 14 (Keith et al., 2011). Although short-term working memory and fluid reasoning broad abilities have not shown consistent sex differences, some differences have been demonstrated in the narrow and specific abilities associated with them. For example, there is evidence of female advantages in auditory short-term storage, a component of short-term working memory (Voyer et al., 2021), and of male advantages in visual-spatial working memory, a component of short-term working memory (Voyer et al., 2017), and quantitative reasoning, a component of fluid reasoning (Johnson & Bouchard, 2007; Keith et al., 2008; van der Sluis et al., 2006). Additional considerations for interpreting research involving these two CHC broad abilities are that psychometric short-term/working memory tasks often differ substantially from working memory tasks used in experimental settings, and fluid reasoning and *g* factors are often not differentiated in research because they correlate with each other perfectly (e.g., Caemmerer et al., 2020; Reynolds et al., 2013).

4. Broad and specific academic achievement abilities

Sex differences in academic achievement scores are often included in descriptions of intelligence differences. The development of these achievement areas relies more on formal academic instruction, however. There are also many more large datasets available for achievement scores (e.g., National Assessment of Educational Progress). Here we discuss sex differences in writing, mathematics, and reading test scores, but do not summarize effect sizes. There are too many studies to summarize, different types of data to analyze (e.g., group vs. individually administered tests), and different ways in which sex differences emerge. However, themes emerge based on our reading of the literature and in our own research using individually administered achievement tests (e.g., Kaufman and Wechsler achievement test batteries that are standalone batteries and different from the intelligence test batteries). The summary below is based entirely on test scores. Females on average obtain higher course grades in all academic areas (Voyer & Voyer, 2014), though the relative profile pattern of differences (e.g., written language vs. math) is consistent across grades and test scores.

4.1. *Writing*

The largest and most consistent sex difference is in writing. Females

---

[7] We are not sure whether location tasks should be included under short-term/working memory or visual processing, and in research they have been described as both. However, because these tasks often do not require items to be remembered in a sequence, we included them under visual processing (see Schneider & McGrew, 2018).

demonstrate consistent advantages in writing that are often at least moderate in size according to Hyde's (2005) criteria (Archer, 2019; Hajovsky et al., 2018; Hedges & Nowell, 1995; Reilly, Neumann, & Andrews, 2019; Reynolds et al., 2015; Scheiber, Reynolds, Hajovsky, & Kaufman, 2015). To provide context, these effect sizes are often almost always larger than what is considered a large effect size ($d = 0.20$) for educational interventions (Kraft, 2020). The female advantage seems to increase as writing task complexity increases, such that female advantages in written expression are larger than advantages in spelling (Pargulski & Reynolds, 2017). Females are also more likely (at least 2:1) to be high achievers in writing, whereas males are at greater risk of writing failure (Pargulski & Reynolds, 2017; Reilly et al., 2019; Wai, Cacchio, Putallaz, & Makel, 2010). The sex difference in writing is an empirical fact.

4.2. Mathematics

The most controversial area related to sex differences and specific abilities is in mathematics–the findings are not as controversial as are hypotheses about what the findings might mean (Ceci & Williams, 2010; Geary, 1999; Halpern et al., 2007). Research findings are nuanced. Male advantages in math problem-solving that emerge at the average to upper end of math problem solving test score distributions seem to be found most reliably. These advantages have been indicated by mean differences at average to above average math score percentiles and by male overrepresentations in the right tails of the score distributions (Martin & Hoover, 1987; Pargulski & Reynolds, 2017; Parker, Van Zanden, & Parker, 2018; Stoet & Geary, 2013; Wai et al., 2010). Hence, it is important to use math-problem solving tests with high ceilings when investigating sex differences in this area.

4.3. Reading

Sex difference findings related to reading skill are mixed, sort of. Females have advantages in reading fluency (Camarata & Woodcock, 2006; Pargulski & Reynolds, 2017). Findings in other aspects of reading are not always consistent, but if there is a sex difference in reading, it favors females (Reilly et al., 2019; Reynolds et al., 2015; Stoet & Geary, 2013). Interestingly, female advantages in reading may be explained in part by female advantages at lower levels of reading ability (e.g., males tend to score lower and exist in greater numbers than females at the 20th percentile of reading; Baye & Monseur, 2016; Robinson & Lubienski, 2011; Stoet & Geary, 2013).

5. Important considerations for research in sex differences and specific abilities

Establishing empirical facts is essential to understanding sex differences–easier said than done. Study design and psychometric and participant sampling are critical, as selected measures and samples provide selected results. We address two matters in more detail. We are not the first researchers to address them, but they are worth reiterating and we offer some suggestions for future analyses.

5.1. Construct breadth

While reviewing studies of sex differences it is common to read different terms used for the same ability or the same term used for different abilities (i.e., jingle-jangle fallacy). Relatedly, construct breadth is important. We had difficulty classifying research findings as related to broad, narrow, or specific ability constructs. Part of this difficulty arises when findings from specific measures are interpreted broadly (e.g., difference in math computation equals difference in general mathematics) or vice versa (e.g., difference in a math composite equals a difference in math computation) (see Brunner et al., 2013; Johnson & Bouchard, 2007; Reynolds et al., 2015).

Some datasets, such as those from comprehensive intelligence and achievement measures that include both subtest and composite scores, allow for analysis at different "levels" to account for construct breadth. Such data, and other multivariate data, may be modeled with hierarchical factor models such as CHC-based models using multi-group confirmatory factor (MG-CFA) models (e.g., van der Sluis et al., 2006; van der Sluis et al., 2008) and multi-indicator, multiple cause (MIMIC) models (Muthén, 1989; see Reynolds et al., 2008, 2015 examples using both models). Advantages of these models are that the assumption of construct equivalence across sex may be tested, and scores from multiple measures of the same construct, and, equally informative, multiple measures of different constructs may be analyzed simultaneously with common factors and structured in theoretically meaningful ways across the sexes. An underappreciated application of MG-CFA and MIMIC models is to investigate whether differences (or similarities) are in narrow abilities, broad cognitive abilities, or *g* (or in all or none of them). The models provide specific frameworks for testing such hypotheses (Niileksela & Reynolds, 2014; Reynolds et al., 2015). For example, several studies have found male advantages in mental rotation tests. Does that advantage remain the same, change, or altogether disappear when other visual processing measures are included and broader visual processing is controlled?

This modeling approach to subtest scores within a test battery may also be applied to items within a test (i.e., differential item functioning; Steinmayr et al., 2015). Moreover, item level data may be analyzed to help understand differences found at the test level (Stewart et al., 2017; Voyer & Hou, 2006). In all, construct breadth is important and should be considered in both analysis and interpretation, including when research moves from describing score differences to explaining score differences (e.g., Burgaleta et al., 2012).

5.2. Beyond simple mean differences

The GSH is based on syntheses of simple mean differences. Nonetheless, mean differences at different ability levels, male-female ratios at different points of score distributions (including the tails), and score variance reveal meaningful sex differences (Feingold, 1992; Hedges & Friedman, 1993; Hedges & Nowell, 1995). These alternative ways to analyze data are not new, but findings based on these alternatives need to be integrated better with interpretations and conclusions from simple mean differences (i.e., GSH). For example, if males are twice as likely to develop a writing disability, describing a writing difference as "small" to "moderate" minimizes the actual importance of this difference. In other words, the implications may not be translated well by standardized effect sizes.

Research beyond mean differences is conducted with academic achievement specific abilities more often than with CHC broad cognitive abilities (e.g., processing speed, visual processing). Quantile regression, a method that allows for statistical tests of mean differences at different points along score distributions, and calculations of male-female ratios along score distributions would be useful to apply to CHC broad, narrow, and specific cognitive abilities scores to better understand sex differences (e.g., Pargulski & Reynolds, 2017). Moreover, combining findings across variables such as academic profiles and interests (e.g., Stoet & Geary, 2020) and multivariate effect sizes (cf. Del Giudice, Booth, & Irwing, 2012; Hyde, 2014) should provide additional insights.

6. Conclusion

The GSH shifted focus from differences to similarities between sexes. The shift, with the intent to promote more equality among the sexes, is noble. Regrettably, the GSH is at times interpreted to mean that no important sex differences exist in psychological variables. Likewise, similarities in general intelligence may be interpreted incorrectly to mean no important sex differences exist in human cognitive abilities. Here, using CHC theory as a guide to structure findings, we highlighted consistent findings of female advantages in broad processing speed and male advantages in broad visual processing, among other narrow and specific ability and achievement differences. We also described model-based approaches that correspond to the latent structure of human abilities and analyses that move beyond simple mean differences as important considerations for future research involving sex similarities and differences. Going forward, although CHC theory provides a useful framework for organizing research findings, it is important to remember that CHC theory is not "complete" either. Moreover, research should not be limited to traditional psychometric tests, and statistical models can incorporate a wide range of measurements (Malanchini et al., 2020).

Finally, research interpretations and debates about the meaning of sex differences should not simply be isolated to single abilities. For example, sex differences in mathematics are often the focus. However, in our opinion, consistent and meaningful female advantages in processing speed (Camarata & Woodcock, 2006) and writing (Reynolds et al., 2015) are rarely discussed. Systematic research that uncovers why these advantages exist for females may produce new insights, and especially when viewed in combination with other findings such as those related to male advantages in visual processing or math problem-solving (e.g., Stoet & Geary, 2013). Ultimately, a reliable and comprehensive empirical account of sex similarities and differences in specific abilities is important. Understanding why those similarities and differences exist together will only contribute to researchers' understanding of human cognition and the combinations of internal and social forces that affect educational performances and achievements.

References

Agelink van Rentergem, J. A., de Vent, N. R., Schmand, B. A., Murre, J. M., Staaks, J. P., & Huizenga, H. M. (2020). The factor structure of cognitive functioning in cognitively healthy participants: A meta-analysis and meta-analysis of individual participant data. *Neuropsychology Review, 30*(1), 51–96. https://doi.org/10.1007/s11065-019-09423-6

American Psychological Association. (2012). Guidelines for psychological practice with lesbian, gay, and bisexual clients. *American Psychologist, 67*(1), 10–42. https://doi.org/10.1037/a0024659

Archer, J. (2019). The reality and evolutionary significance of human psychological sex differences. *Biological Reviews, 94*(4), 1381–1415. https://doi.org/10.1111/brv.12507

Arribas-Aguila, D., Abad, F. J., & Colom, R. (2019). Testing the developmental theory of sex differences in intelligence using latent modeling: Evidence from the TEA Ability Battery (BAT-7). *Personality and Individual Differences, 138*(1), 212–218. https://doi.org/10.1016/j.paid.2018.09.043

Asperholm, M., Högman, N., Rafi, J., & Herlitz, A. (2019). What did you do yesterday? A meta-analysis of sex differences in episodic memory. *Psychological Bulletin, 145*(8), 785–821. https://doi.org/10.1037/bul0000197

Baye, A., & Monseur, C. (2016). Gender differences in variability and extreme scores in an international context. *Large-Scale Assessments in Education, 4*, 1–16. https://doi.org/10.1186/s40536-015-0015-x

Brunner, M., Gogol, K. M., Sonnleitner, P., Keller, U., Krauss, S., & Preckel, F. (2013). Gender differences in the mean level, variability, and profile shape of student achievement: Results from 41 countries. *Intelligence, 41*, 378–395. https://doi.org/10.1016/j.intell.2013.05.009

Buckley, J., Seery, N., & Canty, D. (2018). A heuristic framework of spatial ability: A review and synthesis of spatial factor literature to support its translation into STEM education. *Educational Psychology Review, 30*, 947–972. https://doi.org/10.1007/s10648-018-9432-z

Burgaleta, M., Head, K., Álvarez-Linera, J., Martínez, K., Escorial, S., Haier, R., & Colom, R. (2012). Sex differences in brain volume are related to specific skills, not to general intelligence. *Intelligence, 40*(1), 60–68. https://doi.org/10.1016/j.intell.2011.10.006

Caemmerer, J. M., Keith, T. Z., & Reynolds, M. R. (2020). Beyond individual intelligence tests: Application of Cattell-Horn-Carroll Theory. *Intelligence, 79*, Article 101433. https://doi.org/10.1016/j.intell.2020.101433

Camarata, S., & Woodcock, R. (2006). Sex differences in processing speed: Developmental effects in males and females. *Intelligence, 34*(3), 231–252. https://doi.org/10.1016/j.intell.2005.12.001

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies.* Cambridge University Press.

Ceci, S. J., & Williams, W. M. (2007). *Why Aren't More Women in Science? Top Researchers Debate the Evidence.* Washington, DC: American Psychological Association.

Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions in Psychological Science, 19*(5), 275–279. https://doi.org/10.1177/0963721410383241

Deary, I. J., Thorpe, G., Wilson, V., Starr, J. M., & Whalley, L. J. (2003). Population sex differences in IQ at age 11: The Scottish mental survey 1932. *Intelligence, 31*(6), 533–542. https://doi.org/10.1016/S0160-2896(03)00053-9

Del Giudice, M., Booth, T., & Irwing, P. (2012). The distance between Mars and Venus: Measuring global sex differences in personality. *PLoS One, 7*(1), Article e29265. https://doi.org/10.1371/journal.pone.0029265

Eagly, A. H., & Wood, W. (2013). The nature–nurture debates: 25 years of challenges in understanding the psychology of gender. *Perspectives on Psychological Science, 8*(3), 340–357. https://doi.org/10.1177/1745691613484767

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research, 62*(1), 61–84. https://doi.org/10.3102/00346543062001061

Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry, 31*(4), 271–288. https://doi.org/10.1080/1047840X.2020.1853461

Geary, D. C. (1999). Sex differences in mathematical abilities: Commentary on the math-fact retrieval hypothesis. *Contemporary Educational Psychology, 24*, 267–274. https://doi.org/10.1006/ceps.1999.1007

Geary, D. C. (2021). *Male, Female: The Evolution of Human Sex Differences* (3rd ed.). American Psychological Association.

Hajovsky, D. B., Villeneuve, E. F., Reynolds, M. R., Niileksela, C. R., Mason, B. A., & Shudak, N. J. (2018). Cognitive ability influences on written expression: Evidence for developmental and sex-based differences in school-age children. *Journal of School Psychology, 67*, 104–118. https://doi.org/10.1016/j.jsp.2017.09.001

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*, 1–51. https://doi.org/10.1111/j.1529-1006.2007.00032.x

Härnqvist, K. (1997). Gender and grade differences in latent ability variables. *Scandinavian Journal of Psychology, 38*(1), 55–62. https://doi.org/10.1111/1467-9450.00009

Hedges, L. V., & Friedman, L. (1993). Computing gender difference effects in tails of distributions: The consequences of differences in tail size, effect size, and variance ratio. *Review of Educational Research, 63*(1), 110–112. https://doi.org/10.3102/00346543063001110

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*, 41–45. https://doi.org/10.1126/science.7604277

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 53–91). The Guilford Press.

Hyde, J. S. (2005). The gender similarities hypothesis. *The American Psychologist, 60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Hyde, J. S. (2007). New directions in the study of gender similarities and differences. *Current Directions in Psychological Science, 16*(5), 259–263. https://doi.org/10.1111/j.1467-8721.2007.00516.x

Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology, 65*, 373–398. https://doi.org/10.1146/annurev-psych-010213-115057

Irwing, P. (2012). Sex differences in *g*: An analysis of the US standardization sample of the WAIS-III. *Personality and Individual Differences, 53*(2), 126–131. https://doi.org/10.1016/j.paid.2011.05.001

Jardine, R., & Martin, N. G. (1983). Spatial ability and throwing accuracy. *Behavior Genetics, 13*, 331–340. https://doi.org/10.1007/BF01065771

Jensen, A. R. (1998). *The g Factor.* CT: Prager.

Jensen, A. R., & Reynolds, C. R. (1983). Sex differences on the WISC-R. *Personality and Individual Differences, 4*, 223–226. https://doi.org/10.1016/0191-8869(83)90029-6

Jewsbury, P. A., & Bowden, K. D. (2017). The Cattell-Horn-Carroll model of cognition for clinical assessment. *Journal of Psychoeducational Assessment, 35*, 547–567. https://doi.org/10.1177/0734282916651360

Johnson, W., & Bouchard, T. J. (2007). Sex differences in mental abilities: *g* masks the dimensions on which they lie. *Intelligence, 35*(1), 23–39. https://doi.org/10.1016/j.intell.2006.03.012

Kan, K.-J., Kievit, R. A., Dolan, C., & van der Maas, H. (2011). On the interpretation of the CHC factor Gc. *Intelligence, 39*(5), 292–302. https://doi.org/10.1016/j.intell.2011.05.003

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

Keith, T. Z., & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools, 47*(7), 635–650. https://doi.org/10.1002/pits.20496

Keith, T. Z., Reynolds, M. R., Patel, P. G., & Ridley, K. P. (2008). Sex differences in latent cognitive abilities ages 6 to 59: Evidence from the Woodcock–Johnson III tests of cognitive abilities. *Intelligence, 36*(6), 502–525. https://doi.org/10.1016/j.intell.2007.11.001

Keith, T. Z., Reynolds, M. R., Roberts, L. G., Winter, A. L., & Austin, C. A. (2011). Sex differences in latent cognitive abilities ages 5 to 17: Evidence from the Differential Ability Scales—second edition. *Intelligence, 39*(5), 389–404. https://doi.org/10.1016/j.intell.2011.06.008

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Lakin, J. M., & Gambrell, J. L. (2014). Sex differences in fluid reasoning: Manifest and latent estimates from the cognitive abilities test. *Journal of Intelligence, 2*(2), 36–55. https://doi.org/10.3390/jintelligence2020036

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*, 1479–1498. https://doi.org/10.2307/1130467

Lowe, P. A., Mayfield, J. W., & Reynolds, C. R. (2003). Gender differences in memory test performance among children and adolescents. *Archives of Clinical Neuropsychology, 18*, 865–878. https://doi.org/10.1016/S0887-6177(02)00162-2

Lynn, R. (1999). Sex differences in intelligence and brain size: A developmental theory. *Intelligence, 27*, 1–12. https://doi.org/10.1016/S0160-2896(99)00009-4

Lynn, R., Fergusson, D. M., & Horwood, L. J. (2005). Sex differences on the WISC-R in New Zealand. *Personality and Individual Differences, 39*(1), 103–114. https://doi.org/10.1016/j.paid.2004.12.009

Lynn, R., Irwing, P., & Cammock, T. (2001). Sex differences in general knowledge. *Intelligence, 30*, 27–39. https://doi.org/10.1016/S0160-2896(01)00064-2

Maccoby, E. E., & Jacklin, C. N. (1974). *The Psychology of Sex Differences*. Stanford University Press.

Maeda, Y., & Yoon, S. Y. (2013). A meta-analysis on gender differences in mental rotation ability measured by the Purdue spatial visualization tests: Visualization of rotations (PSVT:R). *Educational Psychology Review, 25*, 69–94. https://doi.org/10.1007/s10648-012-9215-x

Malanchini, M., Rimfeld, K., Shakeshaft, N. G., McMillan, A., Schofield, K. L., Rodic, M., … Plomin, R. (2020). Evidence for a unitary structure of spatial cognition beyond general intelligence. *npj Science of Learning, 5*(1), 1–13. https://doi.org/10.1038/s41539-020-0067-8

Martin, D. J., & Hoover, H. D. (1987). Sex differences in educational achievement: A longitudinal study. *The Journal of Early Adolescence, 7*(1), 65–83. https://doi.org/10.1177/0272431687071007

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Meyer,, E. M., & Reynolds, M. R. (2018). Scores in space: Multidimensional scaling of the WISC-V. *Journal of Psycheducational Assessment, 36*(6), 562–575. https://doi.org/10.1177/0734282917696935

Moffat, S. D., Hampson, E., & Hatzipantelis, M. (1998). Navigation in a “virtual” maze: Sex differences and correlation with psychometric measures of spatial ability in humans. *Evolution and Human Behavior, 19*, 73–87. https://doi.org/10.1016/S1090-5138(97)00104-9

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54*(4), 557–585. https://doi.org/10.1007/bf02296397

Nazareth, A., Huang, X., Voyer, D., & Newcombe, N. (2019). A meta-analysis of sex differences in human navigation skills. *Psychonomic Bulletin & Review, 26*, 1503–1528. https://doi.org/10.3758/s13423-019-01633-6

Niileksela, C. R., & Reynolds, M. R. (2014). Global, broad, or specific cognitive differences? Using a MIMIC model to examine differences in CHC abilities in children with learning disabilities. *Journal of Learning Disabilities, 47*(3), 224–236. https://doi.org/10.1177/0022219412453172

Palejwala, M. H., & Fine, J. G. (2015). Gender differences in latent cognitive abilities in children aged 2 to 7. *Intelligence, 48*, 96–108. https://doi.org/10.1016/j.intell.2014.11.004

Pargulski, J. R., & Reynolds, M. R. (2017). Sex differences in achievement: Distributions matter. *Personality and Individual Differences, 104*, 272–278. https://doi.org/10.1016/j.paid.2016.08.016

Parker, P. D., Van Zanden, B., & Parker, R. B. (2018). Girls get smart, boys get smug: Historical changes in gender differences in math, literacy, and academic social comparison and achievement. *Learning and Instruction, 54*, 125–137. https://doi.org/10.1016/j.learninstruc.2017.09.002

Pezzuti, L., & Orsini, A. (2016). Are there sex differences in the Wechsler Intelligence Scale for Children—Fourth edition? *Learning and Individual Differences, 45*, 307–312. https://doi.org/10.1016/j.lindif.2015.12.024

Reilly, D., Neumann, D. L., & Andrews, G. (2016). Sex and sex-role differences in specific cognitive abilities. *Intelligence, 54*, 147–158. https://doi.org/10.1016/j.intell.2015.12.004

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender differences in reading and writing achievement: Evidence from the national assessment of educational progress (NAEP). *The American Psychologist, 74*(4), 445–458. https://doi.org/10.1037/amp0000356

Reynolds, M. R., Keith, T. Z., Flanagan, D. F., & Alfonso, V. C. (2013). A cross-battery, reference variable, confirmatory factor analytic investigation of the CHC taxonomy. *Journal of School Psychology, 51*(4), 535–555. https://doi.org/10.1016/j.jsp.2013.02.003

Reynolds, M. R., Keith, T. Z., Ridley, K. P., & Patel, P. G. (2008). Sex differences in latent general and broad cognitive abilities for children and youth: Evidence from higher-order MG-MACS and MIMIC models. *Intelligence, 36*(3), 236–260. https://doi.org/10.1016/j.intell.2007.06.003

Reynolds, M. R., Scheiber, C., Hajovsky, D. B., Schwartz, B., & Kaufman, A. S. (2015). Gender differences in academic achievement: Is writing an exception to the gender similarities hypothesis? *The Journal of Genetic Psychology, 176*(4), 211–234. https://doi.org/10.1080/00221325.2015.1036833

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal, 48*(2), 268–302. https://doi.org/10.3102/0002831210372249

Roivianen, E. (2011). Gender differences in processing speed: A review of recent research. *Learning and Individual Differences, 21*(2), 145–149. https://doi.org/10.1016/j.lindif.2010.11.021

Rosén, M. (1995). Gender differences in structure, means and variances of hierarchically ordered ability dimensions. *Learning and Instruction, 5*(1), 37–62. https://doi.org/10.1016/0959-4752(95)00002-K

Scheiber, C., Reynolds, M. R., Hajovsky, D. B., & Kaufman, A. S. (2015). Gender differences in achievement in a large, nationally representative sample of children and adolescents. *Psychology in the Schools, 52*, 335–348. https://doi.org/10.1002/pits.21827

van der Sluis, S., Derom, C., Thiery, E., Bartels, M., Polderman, T. J. C., Verhulst, F. C., … Posthuma, D. (2008). Sex differences on the WISC-R in Belgium and the Netherlands. *Intelligence, 36*(1), 48–67. https://doi.org/10.1016/j.intell.2007.01.003

van der Sluis, S., Posthuma, D., Dolan, C. V., de Geus, E. J. C., Colom, R., & Boomsma, D. I. (2006). Sex differences on the Dutch WAIS-III. *Intelligence, 34*(3), 273–289. https://doi.org/10.1016/j.intell.2005.08.002

Schneider, W. J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan, & E. M. McDonough (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (4th ed., pp. 73–163). Guilford.

Steinmayr, R., Bergold, S., Margraf-Stiksrud, J., & Freund, P. A. (2015). Gender differences on general knowledge tests: Are they due to differential item functioning? *Intelligence, 50*, 164–174. https://doi.org/10.1016/j.intell.2015.04.001

Stewart, C., Root, M. M., Koriakin, T., Choi, D., Luria, S. R., Bray, M. A., … Courville, T. (2017). Biological gender differences in students’ errors on mathematics achievement tests. *Journal of Psychoeducational Assessment, 35*(1–2), 47–56. https://doi.org/10.1177/0734282916669231

Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA data. *PLoS One, 8*(3), Article e57988. https://doi.org/10.1371/journal.pone.0057988

Stoet, G., & Geary, D. C. (2020). Sex-specific academic ability and attitude patterns in students across developed nations. *Intelligence, 81*, Article 101453. https://doi.org/10.1016/j.intell.2020.101453

Strand, S., Deary, I. J., & Smith, P. (2006). Sex differences in cognitive abilities test scores: A UK national picture. *British Journal of Educational Psychology, 76*(3), 463–480. https://doi.org/10.1348/000709905X50906

Voyer, D., & Hou, J. (2006). Type of items and the magnitude of gender differences on the Mental Rotations Test. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 60*(2), 91–100. https://doi.org/10.1037/cjep2006010

Voyer, D., Postma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory: A meta-analysis. *Psychonomic Bulletin & Review, 14*(1), 23–38. https://doi.org/10.3758/bf03194024

Voyer, D., Saint Aubin, J., Altman, K., & Gallant, G. (2021). Sex differences in verbal working memory: A systematic review and meta-analyses. *Psychological Bulletin, 147*(4), 352–398. https://doi.org/10.1037/bul0000320

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Voyer, D., Voyer, S. D., & Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. *Psychonomic Bulletin and Review, 24*, 307–334. https://doi.org/10.3758/s13423-016-1085-7

Wahlstrom, D., Raiford, S. E., Breaux, K. C., Zhu, J., & Weiss, L. G. (2018). The Wechsler preschool and primary scale of intelligence—Fourth edition, Wechsler intelligence scale for children— Fifth edition, and Wechsler individual achievement test—Third edition. In D. P. Flanagan, & E. M. McDonough (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 245–282). The Guilford Press.

Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30-year examination. *Intelligence, 38*(4), 412–423. https://doi.org/10.1016/j.intell.2010.04.006

Woolley, H. T. (1914). The psychology of sex. *Psychological Bulletin, 11*(10), 353–379. https://doi.org/10.1037/h0070064

Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist, 70*(1), 10–20. https://doi.org/10.1037/a0038208