A Meta-Analysis of the Correlations Among Broad Intelligences: Understanding their Relations

Victoria M. Bryan
University of New Hampshire

John D. Mayer
University of New Hampshire

Author Note

All correspondence regarding this manuscript may be sent to: Victoria M. Bryan Department of Psychology, 15 Academic Way, McConnell Hall, University of New Hampshire, Durham, NH 03824.

Prepublication Version of:

Bryan, V.M., & Mayer, J.D. (2020). A meta-analysis of the correlations among broad intelligences: Understanding their relations. *Intelligence, 00,* 000-000 [First published online]. https://doi.org/10.1016/j.intell.2020.101469

Abstract

The broad intelligences include a group of mental abilities such as comprehension knowledge, quantitative reasoning, and spatial reasoning that are relatively specific in their focus and fall at the second stratum of the Cattell-Horn-Carroll (CHC) model of intelligence. In recent years, the field has seen a proliferation of mental abilities being considered for inclusion among the broad intelligences, which poses challenges in terms of their effective and efficient assessment. We conducted a meta-analysis of 61 articles that reported correlations among the broad intelligences. Results indicated that the average correlation among broad intelligences fell between $r = .52$ and $r = .66$, depending upon the estimation model employed. Applying factor analysis to the composite correlation matrix, we also indicate possible dimensions that may be useful to organize the increasing number of broad intelligences. Finally, we discuss the implications of the correlations among broad intelligences as an evaluative tool for candidate intelligences.

Key words: broad intelligences, Cattell-Horn-Carroll (CHC) model, intelligence.

A Meta-Analysis of the Correlations Among Broad Intelligences: Understanding their Relations

Consider the latest hierarchical models of intelligence: They start at a singular, top-most spot occupied by general intelligence—Spearman's classical *g* (Gottfredson, 2002; Spearman, 1904; van der Maas, Dolan, Grasman, Wicherts, Huizenga, & Raijmakers, 2006). The models depict intelligence as dividing into tiers of mental abilities from there, analogous to an organizational chart with the CEO on top and vice presidents below (Carroll, 1993; McGrew, 2009). The next level below *g* (the CEO), describes a second tier of helper abilities—the *broad* intelligences—which are our focus here. The broad intelligences are wide in scope, similar to Thurstone's (1938) primary mental abilities: each with more focus relative to general mental ability. Examples include verbal-comprehension intelligence, perceptual-organizational intelligence, and quantitative reasoning (Flanagan, Alfonso, Ortiz, & Dynda, 2013; Visser, Ashton, & Vernon, 2006), with researchers currently identifying upwards of 17. Moving further down the hierarchy, each broad intelligence divides, at the next level below, into still-more specific measurable mental skills that indicate each broad ability (McGrew, 2009). This three-stratum model of intelligence, also known as the Cattell-Horn-Carroll model (CHC), is particularly influential and the most widely used at present, although alternative influential models also exist such as the Verbal-Perceptual-Image Rotation (VPR) model (Johnson & Bouchard, 2005; Major, Johnson, & Deary, 2012).

The broad intelligences found at the second tier of the CHC model represent wide-scope but distinct areas of reasoning and reflect the diversity of human problem-solving in ways that earlier models could not (Flanagan et al., 2013; MacCann, Joseph, Newman, & Roberts, 2014; Schneider & Newman, 2015; Wagner, 2011; Wai, Lubinski, & Benbow, 2009). Measuring them allows for a fairer representation of people's intellect by encouraging psychologists to assess a much wider range of a person's mental skills. Assessing multiple broad intelligences also allows for the better prediction of criteria relative to using general intelligence alone. For those reasons, one might conclude that the more broad intelligences, the better.

Yet broad intelligences also complicate measurement, because intelligence tests may require redesigns to include them. Moreover, prediction equations of a given criterion must take into account more mental abilities relative to using a single general intelligence alone. In fact, some have asked whether there are "too many intelligences" due to their increasing number (Austin & Saklofske, 2005; Hedlund & Sternberg, 2000). What matters more than the convenience of the number, though, is the accurate representation of human intellect according to how many intelligences truly exist. At its best, the three-stratum model is no more than an approximation of the more complex interrelation among human intellectual abilities. That said, the approximation appears better tailored to the realities of human cognition than *g* alone (but see Gardner, 1983 and Ree, Caretta, & Teachout, 2015). *Hierarchical models reached their current level of acceptance because of their superior fit to people's actual patterns of problem solving (e.g., McGrew, 2009).*

The Importance of the Correlational Levels Among Broad Intelligences

The present research is focused on the correlations among broad intelligences. Factor analysis can be employed to model broad intelligences and the correlations among them. This estimated level is important (a) as a benchmark for determining whether a mental ability *is*, in fact, a broad intelligence, (b) as an indicator of the kind of incremental validity one might expect among broad intelligences, and (c) to understand whether there exist subgroups among broad intelligences. No systematic study of such benchmark values has, to the best of our knowledge, been available before.

The average correlation among broad intelligences as a benchmark. The average correlation level can provide a normative benchmark for the correlations we should expect of any newly proposed broad intelligence to the CHC model. Indeed, researchers stress the importance of the modest correlations among the broad abilities as a criterion for their plausible inclusion (Carroll, 1993; Legree, Psotka, Robbins, Roberts, Putka, & Mullins 2014). If the average correlation among broad intelligences were r = .95, broad abilities arguably would be so closely related as to represent the "same entity" and be unworthy of further consideration as distinct. Or, if the average correlation were r = .05 it would raise serious suspicion that broad intelligences represent independent abilities, with little reason to postulate a general intelligence.

Current factor models suggest that the actual correlations among broad intelligences are far more moderate—but where do the values fall more exactly? A study by Keith & Kranzler (1999) of the Cognitive Assessment System, indicated that r = .75 was the approximate midpoint of correlations reported among several broad intelligences. This value was elsewhere employed as a benchmark for assessing the candidate broad intelligence—personal intelligence— by Mayer, Panter & Caruso (2017). Several years later, Burns & Nettlbeck (2003) suggested a range from r = .67 to .75.

Although researchers have reported their sense of the typical relation among broad abilities, having a clearer estimate of the level could provide a useful benchmark: for example, to ask whether a newly-proposed broad intelligence correlated within a reasonable range with other members of the group. At the extremes, if the new intelligence correlated r = .10 with other broad intelligences, it likely would not be a candidate broad intelligence, whereas if it correlated r = .90 with another broad intelligence, it would be overly similar to the already-studied ability. Establishing a normative benchmark—a typical correlation among existing broad intelligences—is worth pursuing to set a target for new broad intelligences to meet and to understand whether "too many intelligences" may have been considered in the past (Austin & Saklofske, 2005; Hedlund & Sternberg, 2000).

The average correlation among pairs of broad intelligences as an indicator of how to group them. A further purpose of collecting the correlations among broad intelligences is to explore their factor structure: Do they fall along descriptive continua? For example, Schneider and Newman (2015) speculated that a possible continua for organizing the second tier might include contrasting *Power* intelligences including knowledge, attention, and perceptual skill, from *Speed* intelligences, the rapidity with which one finds an answer to a problem (see Schneider & Newman, 2015, Fig. 4). Another possible division is between *Thing-Centered* intelligences such as quantitative and spatial intelligence and *People-Centered* intelligences such as emotional, personal, and social intelligences (Bryan & Mayer, 2017; Mayer, 2018; Mayer & Skimmyhorn, 2017)—although a relatively sparse number of studies to-date have correlated people-centered intelligences with the other broad abilities and so no such dimension was likely to emerge here.

Two Types of Correlational Estimates Among Broad Intelligences

A broad intelligence is an unobserved, hypothetical construct (with considerable evidence for its existence) that is modeled using factor analysis. Psychologists specify indicators of the broad intelligence by selecting specific intelligence tasks they believe represent the skill and then administer the relevant tasks to a sample of individuals. Those researchers then calculate the obtained correlations among the tasks and, from that and other information, create factor models that estimate the correlations between the tasks and a given (hypothesized) broad intelligence. The models may further estimate correlations among the broad intelligences themselves, and sometimes their correlations with *g*.

It is worth distinguishing between two commonly used factor models of the broad intelligences that we here refer to as two- and three-tier models, because they estimate the correlations somewhat differently.

Two-tier models of broad intelligences, and the estimates of correlations among them. The two-tier model represents intelligence as a group of interrelated broad intelligences indicated by specific tasks. At the bottom of Figure 1a., for example, the basic tasks of *concept formation*, *matrix reasoning*, and *analysis-synthesis* serve as indicators of fluid intelligence (Gf), whereas other tasks indicate comprehension (Gc), and visuospatial intelligence (Gv). The two-tier models include estimated correlations among the broad intelligences, represented by the curved lines of Figure 1a that connect Gf, Gc, and Gv. These estimates are based on an optimal weighted combination of indicators, corrected for unreliability.

Three-tier models. The other widely reported model of broad intelligences, represents all three tiers of the three-stratum model, with *g* at the top, as depicted in Figure 1b. The top-level representing *g* is added, along with associated estimated correlations between the broad abilities and *g*. In this revised depiction, the paths among the broad intelligences are replaced instead by paths between each broad intelligence and *g*. This has the effect of accounting for any and all correlations among broad intelligences as a consequence of their relation to *g* and *g* only: that is, it rules out any subsidiary relations among broad intelligence. And, in fact, researchers who report *both* two- and three-tier models on the same data find that their estimated correlations diverge somewhat—a matter we return to later (see, for example, McCann et al., 2014, Morgan, Rothlisberg, McIntosh, & Hunt, 2009, and Thaler, Barchard, Parke, Jones, Etcoff, & Allen, 2015).

In our review, we examine both kinds of models, but place an emphasis on the two-tier models because they allow for an understanding not only of the average correlation among broad intelligences, but also for the identification of possible subgroups of broad intelligences based on their relations independent of *g*. That said, it is worth recognizing that only the three-tier, *g*-inclusive models include all three levels of CHC theory.

Overview of the Present Research

In the present research, we conducted a meta-analysis that included the estimated correlations among broad intelligences from both two- and three-tier models. Our primary goal was to identify the estimated correlations so as to ask, “What is the mean and dispersion of those values?” Because we are interested in the interrelation among the broad intelligences, we focus in particular on two-tier models, as indicated above. A matrix of estimated correlations from such two-tier models can be used to understand the structure of such mental abilities.

Hypotheses

We tested two hypotheses with the data collected in our meta-analysis.

Hypothesis 1. First, we hypothesized that the overall level of estimated correlations among broad intelligence would be at or near r = .75, acknowledging that it may be a bit high (e.g., Burns & Nettlbeck, 2003; Keith & Kranzler, 1999). To test the hypothesized value, we calculated the weighted mean of the correlations among broad intelligences reported in the literature.

Figure 1a. Two-Tier Model of Broad Intelligences

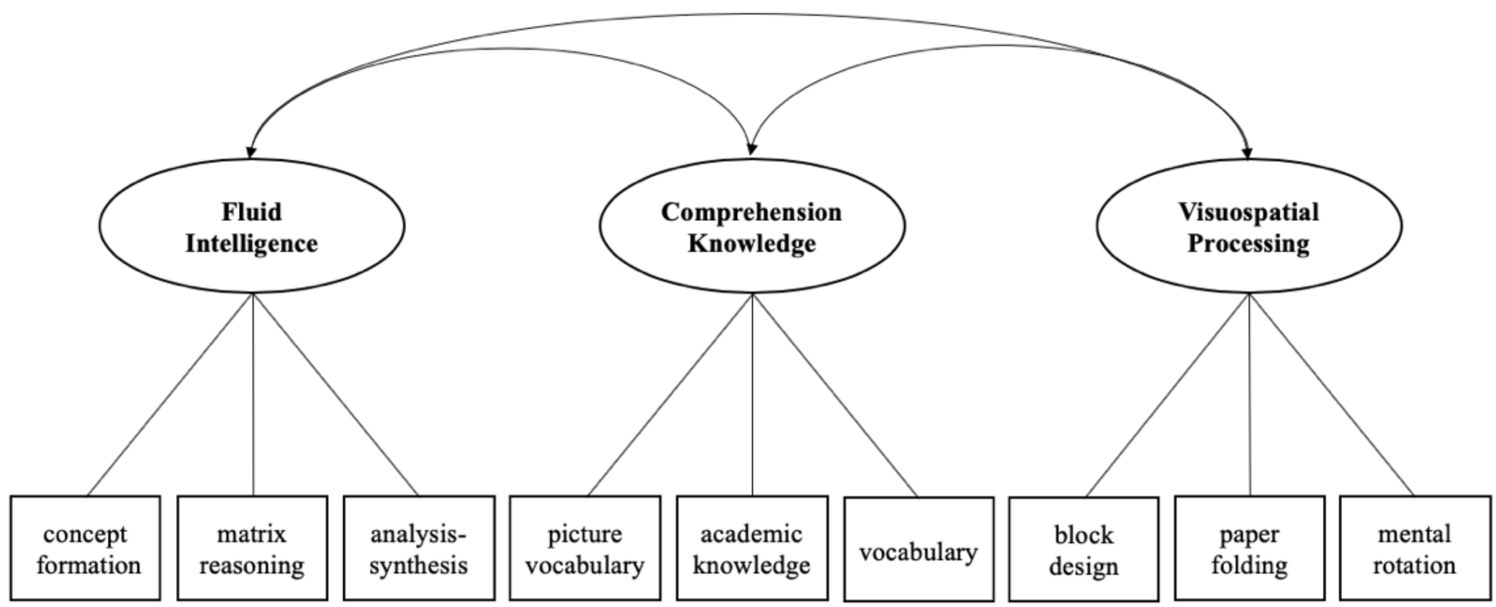

Figure 1b. Three-Tier, *g*-inclusive Model of Broad Intelligences

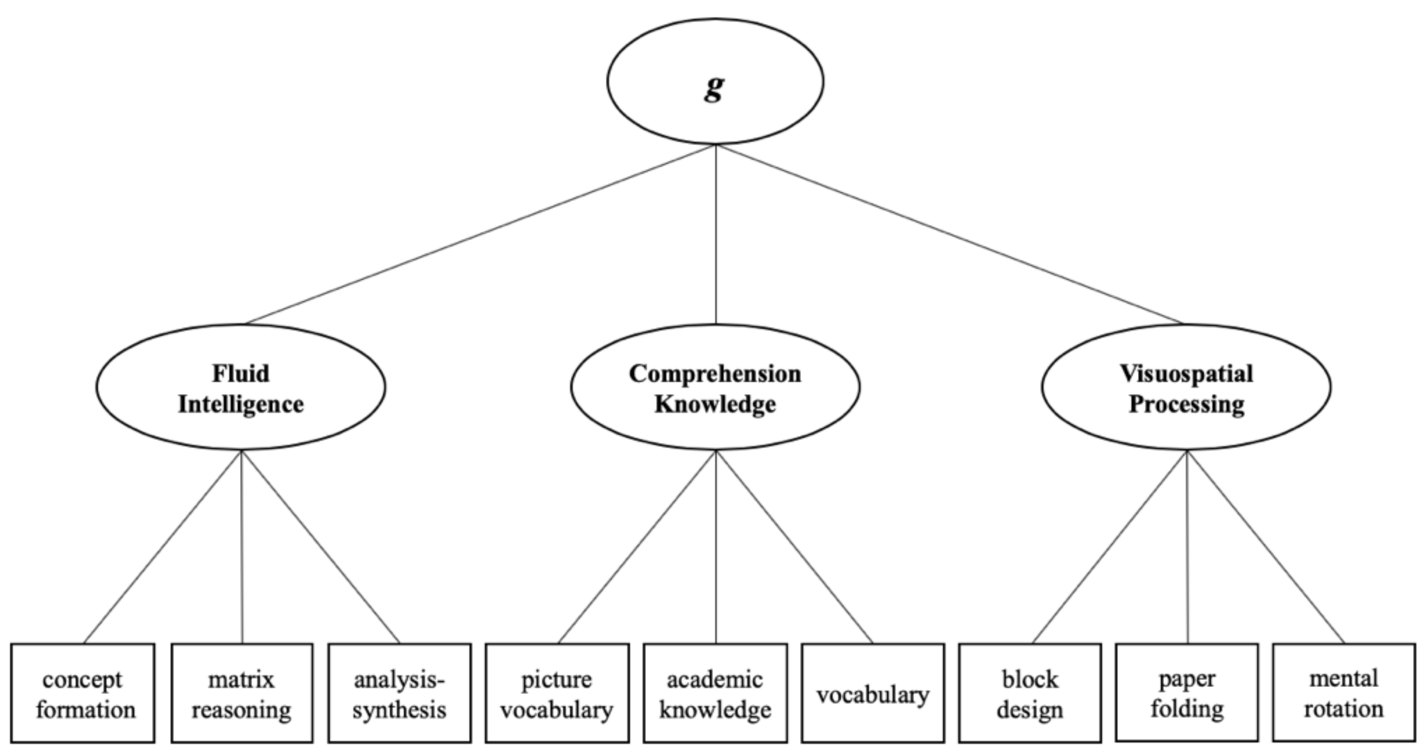

Figure 1. Comparing two- and three-tier, *g*-inclusive models of the broad intelligences.

Following up on Hypothesis 1, we conducted a further test to check for covariates of the average correlation, such as a report's year-of-publication and intelligence test(s) used.

Hypothesis 2. Second, we predicted that the resulting matrix of averaged estimated correlations of broad intelligences would possess a meaningful structure among the broad intelligences. We tested this by conducting a factor analysis of the assembled correlations. This second hypothesis is important because any emergent factors could be used to help organize the increasingly numerous group of established broad abilities.

Methods

Search Method for Identifying Relevant Studies

No comprehensive list of studies of broad intelligences existed at the beginning of our work to the best of our knowledge. Therefore, we conducted several searches using PsycINFO, employing a number of search terms to identify as many relevant articles as possible, published before December of 2019. As shown in the flow diagram (Figure 2), we began our search using terms relevant to CHC theory, including *broad intelligences, Cattell-Horn-Carroll theory,* and *three stratum model*, which collectively yielded 182 peer-reviewed journal articles. Next, the names of prominent researchers in the field of intelligence were entered into PsychINFO, including John B. Carroll, Raymond B. Cattell, John L. Horn, Dawn P. Flanagan, Timothy Z. Keith and Kevin S. McGrew, yielding a total of 545 peer-reviewed articles.

In our final round of searches using PsychINFO, we searched the names of major intelligence assessments and their various editions (e.g., Woodcock-Johnson III or Weschler Adult Intelligence Test IV). Collectively, our searches using the names and editions of major intelligence assessments yielded well over 20,000 works to review for relevant correlations—most of them easily identifiable as irrelevant. To create a more manageable set of results, we narrowed each individual search by adding to the test name the terms *cognitive ability* and/or *psychometrics* as key terms in the article, yielding 6,191 potentially relevant works.

Selection criteria. For each set of search results, the first author read through the titles and abstracts, and quickly excluded remaining irrelevant material and duplicate articles that had emerged in previous searches. Each potentially relevant article was then subject to screening based on a series of inclusion criteria. For inclusion, the work had to: (a) be a peer-reviewed journal article, (b) report an exploratory or confirmatory factor analysis that represented broad factors of intelligence, and (c) report either a two- or three-tier simple-structure oblique factor model of the relation among broad intelligences. Using the above approach, 103 relevant publications were retained for additional review (Figure 2, middle).

Coding of articles. From the 103 relevant publications in the central database, the first author read through each and made note of (a) the year of publication, (b) the journal the article was published in (if applicable), (c) the sample size of the study, (d) the age range of the sample used, (e) the type of sample used in the research (e.g., standardization sample vs. college student sample), (f) the major intelligence test employed, (g) whether the study employed more than one major intelligence test, and (h) the estimated correlations among the broad intelligences included in the study, sorted according to the pair of broad intelligences involved.

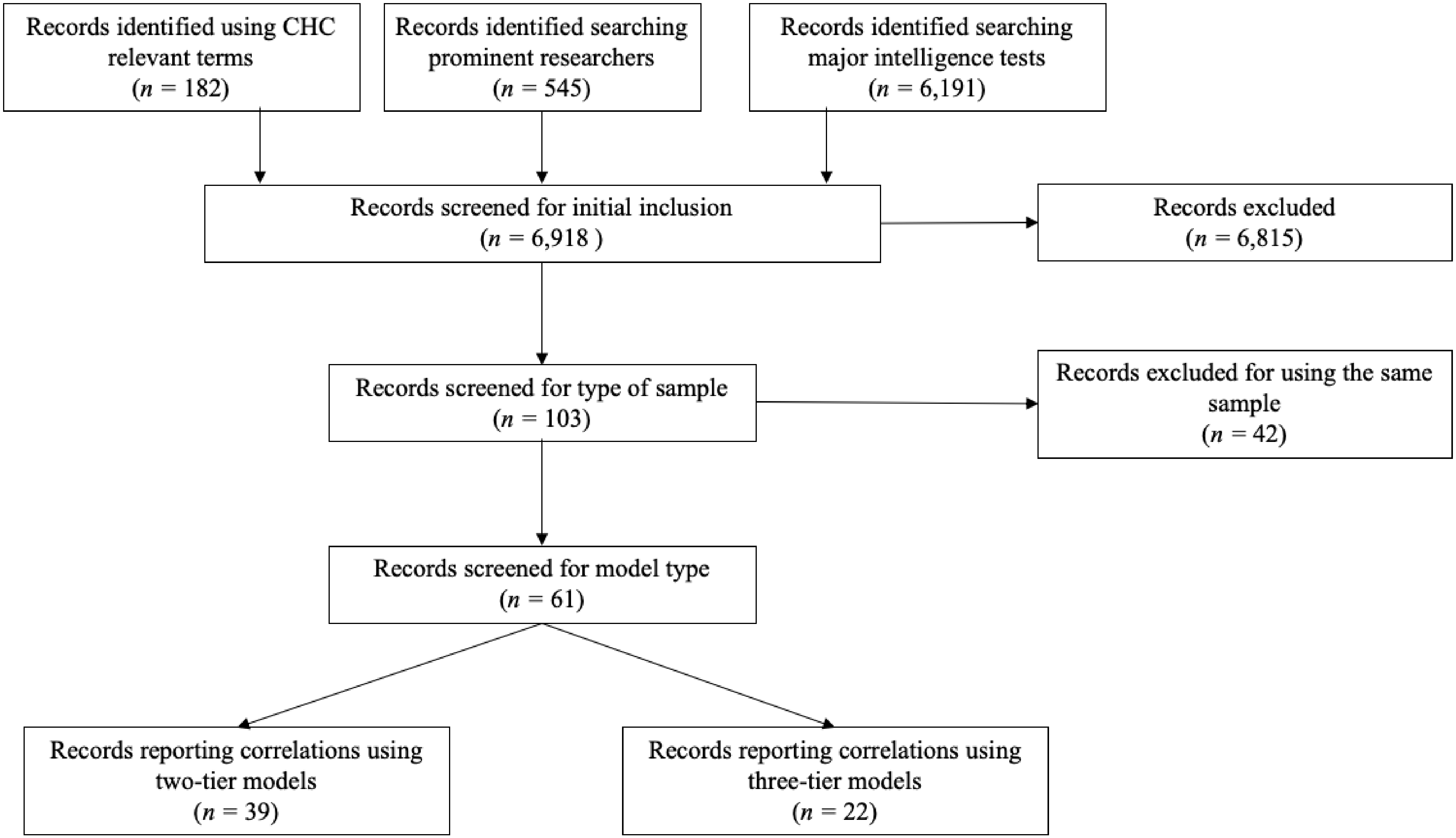

Figure 2. Flow diagram depicting article selection process.

Handling composite factors. Eight instances arose in which the factor-based intelligences examined were composites of two broad intelligences. For example, the Wechsler Adult Intelligence Scales (WAIS) includes a perceptual reasoning index (PRI) which combines the Cattell-Horn-Carroll (CHC) factors of fluid (Gf) and visuospatial (Gv) intelligences. In this instance the factor was reassigned to either fluid intelligence or visuospatial processing based on its indicator tasks and their loadings; a parallel procedure was employed for the other composites (see corresponding section in the technical supplement for additional details).

Controlling for the use of standardization samples. Forty-two of the remaining 103 relevant articles were dropped because they employed the same samples as other studies—for example, the same standardization sample—of a given test. This prevented overrepresenting specific samples. The included articles were either the first published or the most representative (e.g. most comprehensive sample size, widest age range). This left us with 61 relevant works reporting correlations among the broad intelligences.

Lastly, we split our remaining 61 articles according to whether they fitted two- or three-tier models to their data, yielding 39 two-tier articles (*k* studies = 46) and 22 three-tier articles (*k* studies = 46). See Table 1 for a detailed list of the articles included in our analyses.

Results

Study Characteristics

Number of relevant articles and their characteristics. The trend line for the number of publications for two-tier, three-tier, and total studies is indicated in Figure 3 for the period from 1963 to 2019 that they span. As might be expected, the studies appeared in a number of journals including *Intelligence, Psychological Assessment,* and *The Journal of Educational Psychology*, among others. The sample sizes varied from as few as 29 to over 2,000 for an overall sample of N = 20,498 for the two-tier models and N = 51,051 for the three-tier models.

The broad intelligences represented. Table 2 summarizes ten broad intelligences that regularly recur across the 61 articles, including their conventional abbreviations and a short description of each. In some instances, the terminology used to depict broad intelligences evolved over time (see McGrew, 2009; Phelps, McGrew, Knopik, & Ford, 2005). For example, whereas several early works in the field included a broad intelligence named fluency/broad retrieval ability (*Gr*; see Undheim & Gustafsson, 1987; Carroll, 1993), more recent treatments label these as long-term retrieval (Glr), and still more recently, Schneider & McGrew (2018) argued that long-term retrieval might represent two factors: retrieval fluency (Gr) and learning efficiency (Gl).

Commonly discussed broad intelligences such as fluid reasoning and comprehension knowledge were well-represented in the research we reviewed (31 and 41 studies, respectively), whereas less central and newer broad intelligences such as reading and writing ability (Grw) and emotional intelligence (Gei) were less common, at one apiece. The latter were included nonetheless because they have modeled persuasively in a manner that supports their inclusion among the set of broad abilities (Flanagan & McGrew, 1998; MacCann et al., 2014). Additional candidate broad abilities had too little data to include here now but show promise for the future (see Flanagan, Alfonso, & Reynolds, 2013; McGrew, 2009; MacCann et al., 2014; Mayer, Panter, & Caruso, 2019).

Table 1.

List of Included Works by Intelligence Test, Including Model Type, Sample Size, Age of Sample, Population of Sample, and Other Tests Employed

Intelligence Test and Published Works	**Model Type[a]**	**N**	**Age (in years)**	**Type of Sample**	**Cross-Battery**	**Other Test(s)**
Woodcock-Johnson-R						
Bickley, Keith, & Wolfe (1995)	1	2,201	6 to 80	Standardization sample	No	
Flanagan (2000)	1	166	8 to 11	Special validation sample	No	
Flanagan & McGrew (1998)	0	114	10 to 15	School sample	Yes	KAIT, WISC-III
Burns & Nettlbeck (2003)	0	90	18 to 40	Community sample	Yes	WAIS-R
Woodcock-Johnson III						
Keith, Kranzler, & Flanagan (2001)	0	155	8 to 11	School sample	Yes	CAS
Taub & McGrew (2004)	1	7485	6 to 90+	Standardization sample	No	
Sanders et al. (2007)	0	131	3 to 5	Standardization sample	Yes	DAS
Kaufman et al. (2012)	1	6686	4 to 19	Standardization sample for the WJ-III	Yes	KABC-II; KAIT
Strickland, Watkins, & Caterino (2015)	1	529	6 to 13	School sample	No	
Woodcock-Johnson IV						
McGrew, LaForte, & Schrank (2014)	1	6914	3 to 90+	Standardization sample (test manual)	No	
Wechsler Intelligence Scale for Children						
Undheim (1976)	0	144	10 to 12	Norwegian school sample	No	
Wechsler Intelligence Scale for Children III						
Cockshott, Marsh, & Hine (2006)	0	579	6 to 16	Australian school sample	No	
Ogata (2015)	1	105	6 to 12	Japanese sample	Yes	KABC
Wechsler Intelligence Scale for Children IV						
Bergeron & Floyd (2013)	1	85	6 to 18	Clinical sample with mild/moderate ID	Yes	KABC-II; DAS-II
Devena, Gay, & Watkins (2013)	1	297	6 to 15	Clinical sample	No	
Golay et al. (2013)	1	249	Avg. 9.84	French-speaking Swiss children	No	
Nakano & Watkins (2013)	1	176	6 to 16	School sample (Native American)	Yes	WISC-III
Weiss et al. (2013)	1	1967	6 to 16	Clinical + non-clinical standardization	No	
Cavinez (2014)	0	345	6 to 16	Learning disabled school sample	No	
Reverte et al. (2014)	1	249	Avg. 10.21	Swiss school sample	No	
Rowe, et al. (2014)	0	406	6 to 12	Gifted children	No	
Thaler et al. (2015)	0	314	6 to 16	ADHD school sample	No	
Pezzuti & Orsini (2016)	1	2200	6 to 16	Italian standardization sample	No	
Reynolds et al. (2016)	1	166	7 to 16	Shipley-2 validation sample	Yes	Shipley-2
Styck & Watkins (2017)	1	233	6 to 16	ADHD school sample	No	
Do Santos et al. (2018)	0	150	6 to 14	School sample	No	
Wechsler Intelligence Scale for Children V						
Cavinez, Watkins, & Dombrowski (2016)	0	2200	6 to 16	Standardization sample	No	
Reynolds & Keith (2017)	1	2200	6 to 16	Standardization sample	No	

Lecerf & Cavinez (2018)	0	1049	6 to 16	French standardization sample	No	
Cavinez, Watkins, & McGill (2019)	0	415	6 to 16	United Kingdom standardization sample	No	
Wechsler Adult Intelligence Scale-R						
Waller & Waldman (1990)	0	1880	16 to 74	Standardization sample	No	
Davis, Massman, & Doody (2003)	0	516	73.19	Alzheimer's sample	No	
Wechsler Adult Intelligence Scale III						
Dickinson, Iannone, & Gold (2002)	0	120	35 to 44	Clinical sample and standardization	No	
McPherson & Burns (2007)	0	60	20.6	College sample	Yes	WJ-III
Taub & Benson (2013)	1	4650	16 to 90	Standardization sample	Yes	WAIS-IV
Wechsler Adult Intelligence Scale IV						
Niileksela et al. (2013)	0	400	70 to 90	Standardization sample		
Merz et al. (2019)	0	300	18 to 72	Clinical sample	No	
Kaufman Adolescent & Adult Intelligence Test						
Kaufman (1993)	0	124	11 to 12	School sample	Yes	K-ABC
Kaufman, Kaufman, & McClean (1995)	0	1901	11 to 94	Standardization sample	No	
Caruso & Jacob-Timm (2001)	0	60	11 to 14	Cross-check sample	No	
Kaufman Assessment Battery for Children						
Keith et al. (1995)	0	1299	7 to 12	Standardization and sociocultural sample	No	
Kaufman Assessment Battery for Children-II NU						
Morgan et al. (2009)	0	200	4 to 5	School sample	No	
Reynolds et al. (2013)	0	432	6 to 16	Standardization sample	Yes	WISC-III;IV; WJ-III
Differential Abilities Scale						
Keith (1990)	1	3475	3 to 17	Standardization sample	No	
Differential Abilities Scale II						
Cavinez & McGill (2016)	0	3480	2 to 17	Standardization sample	No	
Stanford-Binet Intelligence Scale IV						
Gridley & McIntosh (1991)	0	187	2 to 11	School sample	No	
Kaplan & Alfonoso (1997)	0	441	2 to 5	Preschool sample with ID	No	
Stanford-Binet Intelligence Scale V						
Williams et al. (2010)	0	201	8 to 10	School sample	Yes	WJ-III
Chang et al. (2014)	0	200	4 to 5	Preschool sample	Yes	WJ-III
Culture Fair Intelligence Test						
Cattell (1963)	0	277	13 to 14	School sample	Yes	Thurstone[b]
Undheim (1978)	0	149	12 to 14	Norwegian school sample	Yes	Thurstone; Guilford[c]
Undheim (1981)	0	148	14 to 16	Norwegian school sample	Yes	Thurstone; Guilford
Berlin Model of Intelligence Structure						
Beauducel & Kersting (2002)	1	9520	17 to 32	Community sample	No	
Conzelman & Süß, (2015)	0	301	21 to 40	College sample	Yes	Auditory Intell. Test
Situational Test of Emotion Management						
MacCann et al. (2010)	0	207	19 to 59	College sample	Yes	Educational Testing Kit[d]

Mayer-Salovey-Caruso Emotional Intelligence Test						
MacCann et al., (2014)	1	688	17 to 59	College sample	Yes	MSCEIT
Evans, Hughes, & Steptoe-Warren (2019)	1	830	18 to 71	College and convenience sample	Yes	ICAR-9; STEU; STEM
Multi-Battery/ Test Scales						
Horn & Cattell (1966)	0	297	14 to 61	Prison sample	Yes	Thurstone; Guilford
Cattell & Horn (1978)	0	883	Approx. 14	School sample	No	
Stankov (1978)	0	113	11 to 12	Yugoslavian school sample	No	
Comprehensive Ability Battery						
Hakstian & Cattell (1978)	0	280	15 to 19	Canadian school sample	No	

Note: WJ = Woodcock-Johnson; WISC = Wechsler Intelligence Scale for Children; WAIS = Weschler Adult Intelligence Scale; MSCEIT = Mayer-Salovey-Caruso Emotional Intelligence Test; DAS = Differential Abilities Scale; CAS = Cognitive Assessment System.

[a]Model type distinguishes between studies that represented correlations among the broad intelligences using two-tier (coded as 0), or three-tier, *g*-inclusive models (coded as 1).

[b]see Thurstone (1937).

[c]see Guilford & Hoepfner (1971).

[d]see Ekstrom, French, Harman, & Dermen (1976).

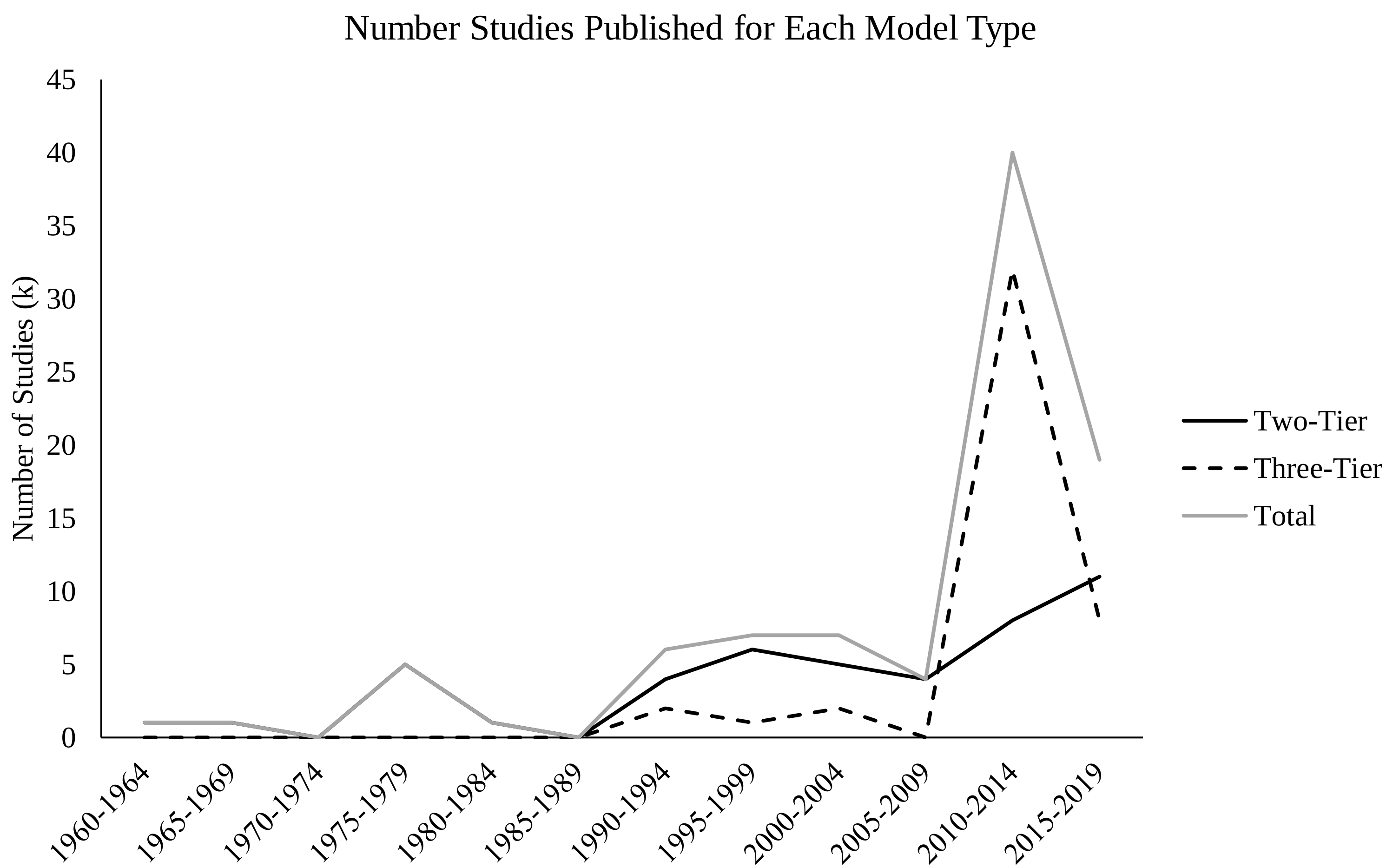

Figure 3. Number of studies published using two-tier and three-tier, *g*-inclusive models in 5-year intervals.

Table 2.

Definitions of Broad Intelligences and Selected Common Tasks Assessing Each

Broad Intelligence	**Selected Relevant Test[a] and Subtests**		**Description**
Fluid Reasoning (Gf)	WJ-II	Concept Formation; Analysis-Synthesis	The ability to use cognitive functions to solve novel problems; using mental tasks to understand concepts, draw inferences, and identify relationships among concepts.
	WISC-IV	Matrix Reasoning	
Comprehension Knowledge (Gc)	WJ-III	Picture Vocabulary; Verb. Comp.; Analogies	Accumulated knowledge, often involving an understanding of the language and knowledge related to one's culture.
	WISC-IV	Vocabulary	
Visual Processing (Gv)	WJ-III	Visual Closure Spatial Relations	The ability to perceive, and mentally represent spatial relations among objects. Visual processing further allows us to mentally transform objects.
	WISC-III	Block Design	
Long-term Retrieval (Glr)	WJ-R	Memory for Names; Delayed Recall	Broad ability facilitating the long-term storage and subsequent retrieval of new information over long periods of time.
Short-term Memory (Gsm)	WJ-II	Numbers Reversed; Memory for Words	The ability to maintain and consciously manipulate a limited amount of information that is susceptible to immediate decay if not attended too.
	WISC-IV	Digit Span	
Processing Speed (Gs)	WJ-III	Cross Out; Visual Matching; Decision Speed	Broad ability facilitating the automatic use of stored information during well-learned tasks. Mental efficiency.
	WAIS-IV	Coding	
Quantitative Reasoning (Gq)	WJ-R	Applied Problems	Quantitative knowledge, often involving the storage of learned declarative and procedurally based knowledge of numbers.
	WISC-III	Arithmetic	
Auditory Intelligence (Ga)	WJ-III	Auditory Attention	The broad ability focused on interpreting and discriminating sounds; involves the ability to cognitively manipulate, synthesize and analyze sounds and sound patterns.
	WJ-R	Sound Blending Incomplete Words	
Reading and Writing (Grw)	WJ-R	Letter-word identification Reading Vocabulary	The broad ability involved in reading and writing ability, including reading comprehension and the ability to write complex narratives.
Emotional Intelligence (Gei)	MSCEIT STEU; STEM	Emotion Blends; Emotion Management	The ability to recognize, understand, and manage emotions in one's self and others.

Screening for Outliers

We next checked the data set for studies that produced outlier values for the average correlations among broad intelligences. The unweighted average correlation within each study ranged from as low as r = .02 to as high as .86, with an unweighted mean of .57 (SD = .20). Studies were flagged as reporting outlying values if their unweighted correlation was above or below | ±3.0 | standard deviations from the mean; no such outliers were present, although one study was close to the designated cut off. Therefore, the full dataset was retained for further analyses.

Test of Hypotheses: The Nature of Correlations among Broad Intelligences

Was the average estimated correlation among broad intelligences near r = .75? (Hypothesis 1). To address Hypothesis 1, we first tested whether the average correlation among broad intelligences would be in the vicinity of r = .75. To do so, we modified a script in the open software project R, drawing on the *meta* package (Balduzzi, Rücker, & Schwarzer, 2019), which uses inverse variance weighting to calculate the average fixed and random effects estimates for the correlations between pairs of broad intelligences (see Technical Supplement for details). We used a random-effects model based on the heterogeneity of correlations across studies, as indicated by large I_2 statistics for both the two- (I_2 = 95.8%, 95% CI [95.1, 96.5]) and three-tier models (I_2 = 98.5%, 95% CI [98.3, 98.6]; e.g., Higgins, Thompson, Deeks, & Altman, 2003). Using this method, average correlations as well as confidence intervals were calculated for each pair of broad intelligences included in our work.

Finding the overall average correlation among broad intelligences involved averaging correlations first within a study—so that each study yielded one average (see Hedges, Tipton, & Johnson, 2010, p. 40). The unweighted averages for each study were then entered into the R script outlined above, to produce a weighted, overall average across studies.

The average estimated correlation for two-tier models. The ten broad intelligences yielded 45 possible pairs of weighted average correlations ((N x (N-1))/2), but 11 of these values were missing in Table 3 (indicated as empty cells) as a consequence of relatively incomplete data for the recently introduced abilities of emotional intelligence and quantitative knowledge. The estimated weighted average correlation among broad intelligences for the two-tier focused models was r = .58 (SE = .03), 95% CI [.53, .64]. Correlations among pairs of broad abilities ranged from r = .22 for processing speed and auditory intelligence, to r = .81 between fluid intelligence and quantitative reasoning. The r = .58 value across studies was in the moderate range we expected, albeit noticeably lower than the hypothesized population mean of ρ = .75.

The average estimated correlation for three-tier models. A somewhat different estimate of the average correlations can be obtained from the three-tier models (46 studies). Recall that all correlations among broad intelligences are due to their relations with g in these models. Imputing the correlations between any pair of intelligences involves multiplying their path coefficients to-and-from g (Leohlin, 2004). This value will differ from two-tier models in that all shared variance among the broad intelligences will be attributable to g. Indeed, the estimated overall average for the three-tier models alone was r = .65 (SE = .01), 95% CI [.62, .68], significantly higher than the estimated overall average for the two-tier models, t(71,547) = 2.19, p = .029, 95% CI [.01 to .12]. (The value for both two- and three-tier models together was r = .62 (SE = .02) , 95% CI [.59, .65]). See the specific correlations for the three-tier models in the corresponding section of the technical supplement.

Table 3.

The Number of Studies Including Each Broad Intelligence, Participants Observed, and the Average Weighted Correlations Among Broad Intelligences[a]

	Fluid Intelligence	Comp. Knowledge	Visuospatial Processing	Short-Term Memory	Long-Term Retrieval	Processing Speed	Quantitative Reasoning	Auditory Intelligence	Reading and Writing	Emotional Intelligence	Totals
Study Characteristics and Number of Participants											
k Studies	31	41	30	24	7	23	4	4	1	1	46
Total N across Studies	11,274	18,637	15,320	9,009	1,460	7,867	829	508	114	207	20,498
Averaged Weighted Correlations (in Bold) Among Pairs of Broad Intelligences and Their Confidence Intervals[b]											
Fluid Intelligence	1.00										
Comprehension-Knowledge	**.64** [.57, .71]	1.00									
Visuospatial Processing	**.58** [.51, .66]	**.60** [.55, .64]	1.00								
Short-Term Memory	**.67** [.57, .76]	**.68** [.63, .74]	**.64** [.57, .70]	1.00							
Long-Term Retrieval	**.46** [.21 .70]	**.56** [.43, .68]	**.48** [.25, .70]	**.53** [.41, .64]	1.00						
Processing Speed	**.54** [.46, .62]	**.36** [.29, .43]	**.47** [.40, .55]	**.48** [.41, .55]	**.37** [.11, .64]	1.00					
Quantitative Reasoning	**.81** [.78, .84]	**.73** [.62, .85]	**.68** [.61, .75]	**.73** [.62, .83]	--	--	1.00				
Auditory Intelligence	**.36** [.16, .55]	**.46** [.22, .69]	**.26** [.16, .35]	**.28** [-.15, .70]	**.38** [.27, .48]	**.22** [.05, .39]	--	1.00			
Reading and Writing	**.46** [.31, .61]	**.85** [.80, .90]	**.42** [.27, .57]	**.45** [.30, .60]	**.62** [.51, .73]	**.25** [.08, .42]	--	**.37** [.21, .53]	1.00		
Emotional Intelligence	**.45** [.34, .56]	**.71** [.64, .78]	--	--	--	--	--	--	--	1.00	
Overall Average[c]	**.60** [.53, .67]	**.60** [.54, .66]	**.57** [.49, .65]	**.60** [.52, .68]	**.48** [.31, .65]	**.43** [.36, .51]	**.74** [.70, .77]	**.31** [.17, .44]	**.49** [.35, .63]	**.58** [.49, .67]	***r* = .58** [.53, .64]

[a]Weighted average correlations are in boldface and were taken from the random-effects model produced from the *meta* package in R. 95% confidence intervals for each weighted average are found below, in brackets.

[b]Only one correlation per pair of broad intelligences was reported per study, so the confidence intervals for the correlations between pairs of broad intelligences are based on independent observations.

[c]The overall average correlation for a given broad intelligence (e.g., for fluid) was calculated first by averaging within study if there was more than one correlation reported, and then running those averages in the R script to find an across study overall average.

Covariates of the average estimated correlations among broad intelligences. We wondered whether any of several additional factors we coded for might influence the correlation among broad intelligences. Therefore, we examined whether the estimated average correlation among broad intelligences differed as a function of the year of publication and intelligence assessment administered.

Average correlation based on year of publication. The 46 two-tier studies included in our review spanned several decades, beginning in the mid 1960's with the advent of confirmatory factor analysis (Jöreskog, 1969) and ending with several studies published in 2019.

A key landmark during this time was the publication of John Carroll's 1993 work, "Human Cognitive Abilities", which promoted further work in the field. Therefore, we split the studies into two groups: two-tier studies published on or before 1993 (*k* studies = 11), and those post-1993 (*k* studies = 35). The weighted average correlation among broad intelligences for the pre-1993 works was r_{mean} = .48 (*SE* = .07), 95% CI [.34, .62], whereas the average for studies published post-1993 was r_{mean} = .62 (*SE* = .03), 95% CI [.56, .67], indicating that studies published on or prior to 1993 had significantly lower correlations among broad intelligences than studies published after 1993 ($t(20{,}496) = 2.27$, $p = .023$, 95% CI [.02 to .25]). We note further that all the three-tier studies reported here had been published after Carroll's 1993 work with the exception of Keith (1990). The higher average weighted correlation estimate from those three-tier studies might also, therefore, be due in part to a year-of-publication influence.

Average correlation based on the intelligence test administered. To investigate the effect of intelligence test on results, we divided them into the 7 major intelligence tests employed and an eighth "Research-Based" test group, that included tasks such as those employed by researchers, including those drawn from Guilford's *Structure of Intellect* model and those drawn from the *Berlin Intelligence Structure* Model (e.g., Guilford & Hoepfner 1971, Thurstone, 1937; Conzelman & Süß, 2015). The weighted average and standard error were then calculated per group (see Table 4). The Research-Based test group yielded far lower average correlations at r = .36 (*SE* = .05), relative to the such tests as the Stanford-Binet (SB) and the Differential Ability Scale (DAS) at r = .77, and .74, respectively; one-way ANOVA, $F(8, 20{,}490) = 7.17$, $p < .001$.

Tukey's post-hoc analyses suggested significant between group differences for all comparisons with the Research-Based tests group (all *p's* < .05), with the exception of the WISC (*M* = .54, *SE* = .06) and Woodcock-Johnson (*M* = .49, *SE* = .04). Excluding comparisons between the WISC and the DAS (*M* = .74, *SE* = .02), all other between group comparisons were not significant (all *p's* > .05). Perhaps researchers are more attentive to employing distinctive tasks than those tests that serve clinical practitioners.

Was there an identifiable structure among the correlations of pairs of broad intelligences? (Hypothesis 2). To explore whether there might exist one or more continua that could be used to characterize the relation among broad intelligences (our Hypothesis 2), we factor analyzed the composite correlation matrix (Table 3). Using maximum likelihood extraction for the exploratory model and an oblimin oblique rotation in Mplus 8.0 (Muthén & Muthén, 1998-2017), we replaced missing correlations with the overall average correlation for each (e.g., Gei r =.58; Grw r = .49; Gq r = .74). We sought a standard "good fit" of an RMSEA less than or equal to .06, and both Comparative and Tucker-Lewis Fit Indices of close to .95 (Boomsma, Hoyle, & Panter, 2012).

Table 4.

Weighted Average Correlation and Standard Deviation Among Broad Intelligences by Major Intelligence Test

Intelligence Test	**Total N**	***k* Studies**	**Mean Corr.**	**95% CI**	**SE**
Woodcock-Johnson	490	4	.49	[.42, .56]	.04
Wechsler Intelligence Scale for Children	5602	10	.54	[.43, .65]	.06
Wechsler Adult Intelligence Scale	3276	6	.71	[.65, .77]	.03
Kaufman Adolescent & Adult Intelligence Test	2085	5	.70	[.66, .74]	.02
Kaufman Assessment Battery for Children	1931	3	.70	[.63, .77]	.03
Stanford Binet	1029	5	.77	[.72, .83]	.03
Differential Abilities Scale	3480	3	.74	[.69, .78]	.02
Research-Based or Other Tests[a]	2605	10	.36	[.25, .46]	.10
Overall Weighted Average Correlation	20498	46	.58	[.53, .64]	.03

[a]Research-based or other tests included assessments used in earlier intelligence work, prior to the development of the other major tests listed, such as those developed by Thurstone (1937), Guilford & Hoepfner (1971), or Ekstrom, French, Harman, & Dermen (1976).

Our initial exploratory analyses were marred by the presence of estimated correlations above r = 1.0—which can distort the integrity of a solution (i.e., Heywood cases; see de Winter, Dodou, & Wieringa, 2009; Hoyle & Duval, 2004; van Driel, 1978; Velicer & Jackson, 1990). We took two different approaches to ameliorate the problem. The results of both approaches are indicated in Table 5. In the sequential-empirical approach, we removed Heywood cases on an empirical basis, beginning with Gq and, as other Heywood cases emerged, removing them one after the other until no further cases emerged (Table 5, left). In the theoretical approach, we drew on the idea that crystallized and fluid intelligences overlap highly with *g* and removed those two to start, followed by Gq (which still exhibited a Heywood case). These two methods yielded highly similar results for the one- and two-factor solutions (Table 5, center).

In each case, the one-factor, general-intelligence models exhibited a marginal fit, at best (e.g., for the conceptual model, RMSEA = .12, CFI = .92, TLI = .89). By comparison, the two-factor solutions both fit well (the sequential model: RMSEA= .10, CFI = .97, TLI = .93; the theoretical model, RMSEA= .07, CFI = .99, TLI = .97). In both sets of solutions, the first factor of the two-factor solutions represented a fluid-like intelligence, loading fluid intelligence (Gf) where it was included in the sequential model, as well as (in both models) short-term memory (Gsm), visuospatial processing (Gv), processing speed (Gs), and emotional intelligence (Gei). The second factor represented a crystallized-like intelligence, loading long-term retrieval (Glr) and reading and writing (Grw), and for the sequential model only, comprehension knowledge (Gc).

A three-tier version of the factor analysis. The three-tier model ought to yield a powerful one-factor solution because the model assigns any common variance shared by broad intelligences to *g*. We modified the three-tier correlation matrix as we had the two-tier matrix, replacing 4 missing correlations with emotional intelligence with its overall average correlation (Gei r =.60). The data converged on one factor but yielded a relatively poor fit, which Mplus flagged as owing to the large negative residual variance for Gei. Muthén (2005) recommended removing such variables. Removing emotional intelligence led to a one factor model with a far better fit, RMSEA = .09, CFI = .97, TLI = .96. The superior fit of the one-factor solution from the three-tier models relative to the two-tier models provided a striking confirmation of the effects of the different allocations of covariance produced by these two models (Table 5, right). The model indicates that fluid intelligence is most representative of a one-factor model; processing speed is least representative.

Examination of Publication Bias

Analyses related to publication bias are controversial at present for many reasons, including whether interpretations of bias are always warranted (see van Aert, Wicherts, & van Assen, 2019). Moreover, bias-detecting software tailored to correlation coefficients is designed to work with actual correlations, for which the sampling distribution is understood. In our meta-analysis, however, we analyzed *estimated* correlations, which have a less-well-understood distribution (Yuan et al., 2010, p. 633). For those interested, however, we report a funnel-plot, a widely used method for visualizing publication bias in meta-analyses (Sterne, Sutton, Ioannidis, Terrin, Jones, Lau, et al., 2011). The plot is shown in Figure 4 and was created using the *meta* package (Balduzzi et al., 2019). We note that relatively few studies included in our meta-analysis contained small sample sizes, as shown by the small number of studies found towards the middle and bottom of our plot.

Table 5.

Fit Statistics and Factor Loadings for the 1- and 2-Factor Exploratory Solutions using the Two-Tier (N = 20,498) and Three-Tier (N = 51,051) Models of Broad Intelligences

	Two-Tier Models of Broad Intelligences						***Three-Tier Models of Broad Intelligences***		
	Fit Statistics – Sequential Approach[a]			**Fit Statistics – Theoretical Approach**[b]			**Fit Statistics – Three-Tier Model**		
	RMSEA	CFI	TLI	RMSEA	CFI	TLI	RMSEA	CFI	TLI
One Factor	.13	.91	.87	.12	.92	.89	.09	.97	.96
Two Factors	.10	.97	.93	.07	.99	.97	--	--	--
Three Factors[c]	--	--	--	.03	1.00	.99	--	--	--
	Factor Loadings						***Factor Loadings***		
Broad Intelligence	One-Factor Solution	Two-Factor Solution		One-Factor Solution	Two-Factor Solution		One-Factor Solution	Two-Factor Solution	
	I	I	II	I	I	II	I	I	II
Fluid Intelligence (Gf)	**.77**	**.78**	.03	--	--	--	**.92**	--	--
Comp. Knowledge (Gc)	--	--	--	--	--	--	**.83**	--	--
Visuo-Spatial Processing. (Gv)	**.75**	**.72**	.06	**.74**	**.78**	.00	**.82**	--	--
Short-term Memory (Gsm)	**.80**	**.77**	.06	**.77**	**.76**	.06	**.78**	--	--
Long-term Retrieval (Glr)	**.72**	.17	**.67**	**.75**	.21	**.64**	**.91**	--	--
Processing Speed (Gs)	**.61**	**.74**	-.13	**.57**	**.69**	.35	**.59**	--	--
Quant. Knowledge (Gq)	--	--	--	--	--	--	**.82**	--	--
Auditory Intelligence (Ga)	**.53**	.24	.35[d]	**.54**	.23	.35	**.87**	--	--
Reading and Writing (Grw)	**.63**	-.06	**.84**	**.65**	-.07	**.84**	**.81**	--	--
Emotional Intelligence (Gei)	**.69**	**.40**	.35	**.72**	**.49**	.27	--	--	--
	Intercorrelations for the Two-Tier Models						***Intercorrelations for the Three-Tier Models***		
	I	**II**	**III**	**I**	**II**	**III**	I	II	III
Factor I	1.00			1.00			1.00		
Factor II	.71	1.00		.70	1.00		--	1.00	
Factor III	--	--	1.00	--	--	1.00	--	--	1.00

[a]The sequential approach involved the stepwise removal of Heywood cases. In earlier iterations of the model, both quantitative knowledge and compression knowledge had factor loadings greater than 1. Quantitative knowledge was removed first, followed by comprehension knowledge in order to produce the above fits and factor loadings.

[b]The theoretical approach involved sought an acceptable-fitting model by removing broad intelligences that have previously demonstrated exceptionally high loadings on *g*. Both fluid intelligence, quantitative reasoning, and comprehension knowledge have at times been suggested to be indistinguishable from *g*.

[c]Although the three-factor solution converged in the conceptual model, an ultra-Heywood case for Reading and Writing on the second factor (loading = 1.82) and is not included in the solutions presented in the table.

[d]Auditory intelligence failed to load above .40 on either factor of our two-factor sequential model. Removal of auditory intelligence from the model results in additional Heywood cases.

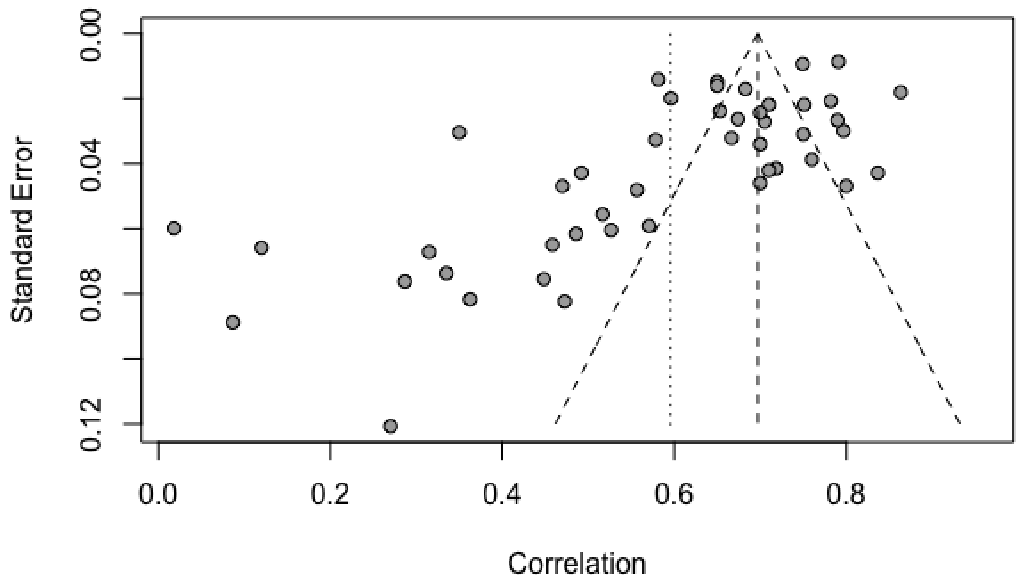

Figure 4. Funnel plot for meta-analysis exploring the correlation among broad intelligence factors. The dark gray, closed dots represent the observed studies collected in the literature search. The small, dotted vertical line in the plots represents the average, unweighted correlation among broad intelligences at the study level.

A standard interpretation of the funnel plot would be to note its slight asymmetry, given its distinct cluster of studies towards the top right side of the plot. An Egger's test (Egger, Smith, Schneider, & Minder, 1997) suggested statistically significant funnel plot asymmetry was present (bias = -6.07, 95% CI = -7.02 to -3.11, $p < .001$). A standard bias-focused interpretation might be that lower correlation estimates were favored by reviewers and their choices led to some publication bias among studies; yet other explanations aside from publication bias are possible as well—both conceptual and statistical; these are considered in the Discussion that follows.

Discussion

Intelligence researchers regard the Cattell-Horn-Carroll model of intelligence as an appealing contemporary representation of intelligences—and it fits empirical data well. At the same time, however, much remains to be understood about the general characteristics of the model. One unknown, until the present research, was the average correlation among broad intelligences one might expect. Here, we distinguished two-tier factor models, which focus on the relation among broad intelligences, and three-tier factor models, which focus on the relation between broad intelligences and *g*. The actual average correlations among broad intelligences were, for the two-tier models, $r = .58$, 95% CI [.53, .64] and for the three-tier models, $r = .65$, 95% CI [.62, .68]. Using the estimated $r = .58$ from the two-tier estimates, for example, a given pair of broad intelligences should have average estimated correlations of between $r = .53$ and .64, 95% of the time. That said, there are "individual differences" in the relation among these broad abilities. The correlation matrix indicates, for example, that although auditory intelligence and processing speed correlate $r = .22$, the correlation between short-term memory and visuospatial processing was $r = .64$, and between reading and writing and comprehension knowledge was $r = .85$.

Correlations of about $r = .60$ represent a moderate degree of relationship between variables, allowing both for overlap *and* distinct interpretations of the variables' meanings and predictions as well. Parallels may be drawn to the treatment of overlapping socio-affective traits in personality and psychopathology. For example, a recent meta-analysis found a correlation of $r = .61$ between extraversion and self-acceptance (Anglim, Horwood, Smillie, Marrero & Wood, 2020, Table 4). Extraversion and self-acceptance are theoretically distinct, of course: People can be extraverted but feel badly toward themselves, or introverted and self-accepting, yet the $r = .61$ correlation also makes sense because both extraversion and self-acceptance reflect more general positive affect. As a second example, the correlation between anger and anxiety—distinct but overlapping elements of negative affect—is about $r = .56$ across studies, and the comparable value for anxiety and depression is $r = .72$ (Ng, Sorensen, Zhang, & Yim, 2019).

Ascertaining the average correlation among broad intelligences provides a benchmark for understanding which mental abilities—old and new—are reasonably considered candidates for inclusion in the model. New candidate mental abilities that fall within the range of correlations among other broad intelligences may be considered similar yet distinct enough to include, whereas mental abilities with correlations too low or too high may be less plausible candidates.

Additional influences on the estimates. Certain additional variables appeared to impact the estimated correlations among broad intelligences including whether the article was published before or after Carroll's (1993) development of the CHC model, with earlier studies exhibiting somewhat lower estimates, and differences in correlations associated with the intelligence test employed. Regarding the latter, we found that the "Research-Based" group of tests that included the Berlin Model of Intelligence Structure and others exhibited lower correlations among the broad intelligences than other assessments like the Stanford-Binet or the Wechsler scales. It may be that research studies, which arguably place greater emphasis on the careful specification of distinct tasks than more applied clinical assessments, better distinguish among the intelligences. Alternatively, however, perhaps greater confidence could be placed in the large more representative samples used to standardize clinical instruments.

The Different Estimates Between Two- and Three-Tier Models

The three-tier models yielded an overall correlational of r = .65, about .07 higher than the two-tier model estimate of r = .58. These values are different, albeit fairly close together. The *g*-inclusive, three-tier models may simply have yielded slightly higher correlations because, like other more recent studies, the estimates among broad intelligences rose since 1993—or because the preponderance of them were developed on the Weschler and Kaufman scales, which yield higher estimates than other measures.

Alternatively, three-tier models may mistakenly allocate some reliable covariance among the broad intelligences to general intelligence. As a consequence, imputing correlations from the elevated relations might have led to an overestimate of the size of the correlations among broad intelligences. Similarly, the reliable covariance among subsets of broad intelligences may have been mis-allocated to error terms, which would, first, reduce the estimated reliabilities of the measures, and consequently, overcompensate by raising the estimated correlations between them to compensate.

Our sense is that although the three-tier models have the advantage of including all three levels of the CHC theory, such models are less useful to estimating correlations among broad intelligences. By explaining the correlations among broad intelligences strictly as a consequence of *g*, the possibility that there are subsets of broad intelligences is obviated. The indirect imputation from the three-tier models seem less compelling to us than the results from the two-tier models which were designed to provide (relatively) direct estimates of these correlations. That said, we acknowledge MacCallum and Austin's (2000) point that:

> "there is no true model…all models are wrong to some degree…the best one can hope for is to identify a parsimonious, substantively meaningful model that fits observed data adequately well." (MacCallum & Austin, 2000, p. 218)

And, we view the two- and three-tier models as complementary given that the three-tier approach represents all three levels of the CHC model but the two-tier allows for additional information about broad intelligences.

Estimated versus Actual Correlations Among Broad Intelligences

The estimated correlations among broad intelligences we studied here can be used to predict the actual, obtained correlations researchers might expect among factor-based scales. To be sure, most obtained correlations are between tasks, but some factor-based scales based on task composites also are employed in the literature. Recall that estimates of correlations within a factor-analytic context correct for errors of measurement (i.e., lack of reliability) of the original measures. To transform the estimates we work with here to predict real-life correlations with their less-than-perfect reliability, we can use the correction for attenuation due to unreliability (solving for the original correlation rather than for the corrected value—the reverse of its more common application). For example, if the estimated correlation between two broad intelligences is, on average r = .58, and if the measures employed have reliabilities of α = .75 each, then the obtained correlation would be r = .44; the comparable values for two tests with reliabilities of α = .80 would be r = .46 and α = .90 would be .52.

Using the logic above, we would expect that, for example, reading and writing ability (Grw), which has a (two-tier) average estimated correlation with short-term memory (Gsm) of *r* = .45, 95% CI [.30, .60], should exhibit obtained correlations of about .32 given reliabilities of measures of α = .70. Consistent with this, a recent meta-analysis by Peng, Barnes, Wang, Wang, Li, Swanson, Dardick, & Tao (2018) report task-based correlations between reading ability and working memory (a narrow indictor of short-term memory) between .22 and .37 (average *r* = .29), closely approximating our estimate; other findings are similarly within range (see Peng, Lin, Ünal, Lee, Namkung, Chow, & Sales, 2020; Peng, Namkung, Barnes, & Sun, 2015).

The Estimated Average Correlation Among Intelligences and their Incremental Validity

Whether one uses the two-tier *r* = .58 average correlation of the broad-intelligence-focused models or something higher (e.g., from the three-tier *g*-inclusive models), there is some room for incremental validity of one intelligence to another in predicting an outcome. Kenny (2016) notes, for example, that "A correlation of .85 or larger in absolute value indicates poor discriminant validity," but sees less cause for concern in values below that. Indeed, Schneider and Newman (2015) argued that broad intelligences add 2% to 6% of the variance over *g*: That is, incremental predictions of *r* = .14 to .24. And, comparing a big trait such as extraversion to its facets of sociability and talkativeness (roughly parallel to second-stratum abilities), Anglim et al. (2020, p. 308) noted "There is good empirical evidence that, collectively, narrow traits are better predictors of outcomes than broad traits…particularly when the outcome is narrow."

Applications

Decisions as to the viability of an intelligence as a candidate *broad* intelligence. The benchmark for correlations among broad intelligences provided here is useful for understanding estimates among proposed new broad intelligences relative to those already widely accepted as members of the set of such second-tier intelligences. For example, the finding by Keith & Kranzler (1999) that the *Cognitive Assessment Battery* exhibited an estimated correlation of *r* = .90 in a two-stratum hierarchical model between its Planning and Attention measurement areas across age groups, suggested that those two areas might be considered for merging (as the authors then proceeded to do in their three-stratum model in order to improve the fit).

Recent work assessing the inclusion of emotional intelligence indicates it is near such benchmark values, arguing for its inclusion. McCann et al.'s (2014) found that the three factors of emotional intelligence, as measured by the Mayer-Salovey-Caruso Emotional Intelligence Test (Mayer, Salovey, & Caruso, 2002), correlated between *r* = .41 to .74 with other broad intelligences in their 8-factor oblique model, mostly within the range indicated by the present research of other broad intelligences.

Organizing broad intelligences by examining their structure. Using exploratory factor analyses with our composite two-tier correlation matrix, we obtained a reasonable fit for a two-factor model of broad intelligences, dividing them between something like fluid (Gf) and crystallized (Gc) groups. This will not come as shocking news to most intelligence researchers, but it does highlight the continued relevance of the earlier Gf-Gc model of mental abilities proposed by Cattell and Horn (see Cattell, 1963 and Cattell & Horn, 1978). Psychologists may be able to examine organizations of broad intelligences within these two areas to provide some empirically supported method for organizing the still proliferating number of broad intelligences proposed.

Strengths and Limitations

This work represented a first meta-analysis of studies of broad intelligences with a focus on exploring their relation to one another and the potential implications of these findings.

Although we were able to find 60 articles that published studies that collectively represent important findings in the area, it is possible that we unknowingly omitted studies of potential relevance, despite our efforts to carry out a reasonably thorough search of the intelligence literature. That said, we believe the results reported here provide a valid estimate of the relations among broad abilities.

Similarly, although the current work represents ten of the most studied broad intelligences and their interrelations, researchers continue to evaluate a number of candidate abilities for inclusion at the second stratum of the CHC model, including psychomotor speed (Gps), kinaesthetic abilities (Gk), and personal intelligence (Gpi; see MacCann et al. 2014, Mayer, 2008, and Schneider, Mayer, & Newman, 2016). The limit of the correlational estimates to ten broad intelligences likely reduced the possibility that we fully accounted for the dimensionality of broad intelligences; that is, there may yet be more than just fluid and crystallized groups. The continued study of newly proposed mental abilities is likely to enhance our understanding of whether certain subgroups of broad intelligences exist, and the shared underlying nature of the mental abilities that make up those groups.

A further limitation is our uncertainty regarding the interpretation of funnel plots as indicators of publication bias (van Aert et al., 2019). Our funnel plot suggested that editors *might* favor the publication of works with lower correlations among broad intelligences over findings of higher values. That said, such a conclusion seems questionable in this instance. First, pressure on researchers to report low or high correlations among broad intelligences seems minimal (in many cases) given that most the estimated correlations would be part of a more global model that was being tested. Moreover, several possibilities aside from publication bias may account for the asymmetry. For example, perhaps the asymmetry was due to the limit of $r = 1.0$ on estimated correlations (excepting Heywood cases). In addition, true differences in the size of the effects according to sample size might be the case given that many such studies with smaller N used research-developed measures: Those more carefully-culled measures, in turn, might better have distinguished broad intelligences through better measurement. (Egger et al., 1997).

Lastly, it is important to question whether the difference in correlations among broad abilities based on the two different types of models employed in the research are truly as different as our findings suggest. Undoubtedly, the inclusion of additional works as they become available may enhance our understanding and paint a clearer picture as to how different (or similar) these models may be in terms of their predictions regarding how human mental abilities relate to one another. We note at least three additional relevant works published since our final search in December of 2019, which we have included in a master list of relevant studies for future research (see Relevant Works in technical supplement).

Conclusions

We began this article by pointing out the growing influence of the CHC model of intelligence, and the growing interest in the broad intelligences that comprise the backbone of the model. The growing number of broad intelligences, however, draws into question what criteria must be met in order for a proposed intelligence to be included within the set of such broad intelligences. Here, we proposed one possible benchmark for evaluating newly-proposed broad intelligences in the context of the model, suggesting that the average correlation among broad intelligence factors, estimated chiefly within commonly accepted factor models, may help us distinguish between proposed intelligences which are indeed distinct from one another and those that are a subclass of another, already-existing broad intelligence—or not a mental ability at all. In addition, we examined potential organizing continua for some of the more traditional intelligences studied between the mid-20th century and the present. We believe this direction bears considerable promise for evaluating broad intelligences and for better understanding hierarchical models of intelligence in the future.

References

Anglim, J., Horwood, S., Smillie, L.D., Marrero, R.J., & Wood, J. K. (2020). Predicting psychological and subjective well-being from personality: A meta-analysis. *Psychological Bulletin, 146*, 279-323.

Austin, E. J., & Saklofske, D. H. (2005). Far too many intelligences? On the communalities and differences between social, practical, and emotional intelligences. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 107-128). Ashland, OH: Hogrefe & Huber Publishers.

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: a practical tutorial. *Evidence-Based Mental Health.*

Bergeron, R., & Floyd, R.G. (2013). Individual part score profiles of children with intellectual disability: A descriptive analysis across three intelligence tests. *School Psychology Review, 42*, 22-38.

Beauducel, A., & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence Structure (BIS). *European Journal of Psychological Assessment, 18,* 97-112.

Bickley, P.G., Keith, T.Z., & Wolfle, L.M. (1995). The three-stratum theory of cognitive abilities: Test of the structure of intelligence across the life span. *Intelligence, 20,* 309-328.

Boomsma, A., Hoyle, R.H., & Panter, A.T. (2012). In R.H. Hoyle & H. Hoyle (Eds.), The structural equation modeling research report, (pp. 341-358). New York, NY, US: Guilford Press.

Bryan. V.M., & Mayer, J.D. (2017). People versus thing intelligences? Presentation at the 2017 Meeting of the *Association for Research in Personality*, June 8-10. Sacramento, CA.

Burns, N. R., & Nettelbeck, T. (2003). Inspection time in the structure of cognitive abilities: Where does IT fit. *Intelligence, 31*, 237-255. doi:10.1016/S0160-2896(02)00120-4

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY, US: Cambridge University Press. doi:10.1017/CBO9780511571312

Caruso, J.C., & Jacob-Timm, S. (2001). Confirmatory factor analysis of the Kaufman Adolescent and Adult Intelligence Test with young adolescents. *Assessment, 8*, 11-17. doi:10.1177/107319110100800102

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

Cattell, R.B., & Horn, J.L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement, 15,* 139-164.

Cavinez, G.L. (2014). Construct validity of the WISC-IV with a referred sample: Direct versus indirect hierarchical structures. *School Psychology Quarterly*, *29*, 38-51. doi:10.1037/spq0000032

Cavinez, G.L., & McGill, R.J. (2016). Factor structure of the Differential Ability Scales – Second Edition: Exploratory and hierarchical factor analysis with the core subtests. *Psychological Assessment, 28,* 1475-1488. doi:10.1037/pas0000279

Cavinez, G.L., Watkins, M.W., & Dombrowski, S.C. (2016). Factor structure of the Wechsler Intelligence Scale for Children – Fifth Edition: Exploratory factor analysis with the 16 primary and secondary subtests. *Psychological Assessment, 28,* 975-986. doi: 10.1037/pas0000238

Cavinez, G.L., Watkins, M.W., & McGill, R.J. (2019). Construct validity of the Wechsler Intelligence Scale for Children – Fifth Edition: Exploratory and Confirmatory factor analysis of the 16 primary and secondary subtests. *British Journal of Educational Psychology, 89,* 195-224. doi: 10.1111/bjep.12230

Chang, M., Paulson, S.E., Finch, W.H., McIntosh, D.E., & Rothlisberg, B.A. (2014). Joint confirmatory factor analysis of the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition, and the Stanford-Binet Intelligence Scales, Fifth Edition, with a preschool population. *Psychology in the Schools, 51,* 32-57. doi: 10.1002/pits.21734

Cockshott, F.C., Marsh, N.V., & Hine, D.W. (2006). Confirmatory factor analysis of the Wechsler Intelligence Scale for Children – Third Edition in an Australian clinical sample. *Psychological Assessment, 18,* 353-357. doi:10.1037/1040-3590.18.3.353

Conzelmann, K., & Süß, H. (2015). Auditory intelligence: Theoretical considerations and empirical findings. *Learning and Individual Differences, 40,* 27-40. doi:10.1016/j.lindif.2015.03.029

Davis, R.N., Massman, P.J., & Doody, R.S. (2003). WAIS-R factor structure in Alzheimer's disease patients: A comparison of alternative models and an assessment of their generalizability. *Psychology and Aging, 18,* 836-843. doi: 10.1037/0882-7974.18.4.836

Devena, S.E., Gay, C.E., & Watkins, M.W. (2013). Confirmatory factor analysis of the WISC-IV in a hospital referral sample. *Journal of Psychoeducational Assessment, 31*, 591-599. doi:10.1177/0734282913483981

de Winter, J.C.F., Dodou, D., & Wierniga, P.A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research, 44,* 147-181. doi:10.1016/j.beproc.2012.11.016

Dickinson, D., Iannone, V.N., Gold, J.M. (2002). Factor structure of the Wechsler Adult Intelligence Scale – III in Schizophrenia. *Assessment, 9,* 171-180. doi: 10.1177/10791102009002008

dos Santos, A.A.A., Muniz, M., Rueda, F.J.M., Martins, R.M.M. (2018). Validity evidence for the 4th edition of the Wechsler Intelligence Scale for Children. *Psico, 49,* 1-11.

Egger, M., Smith, G.D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315,* 629-634. doi:10.1136/bmj.315.7109.629

Ekstrom, R. B., French, J. W., Harman, H. H., & Derman, D. D. (1976). Manual for kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service.

Evans, T.R., Hughes, D.J., & Steptoe-Warren, G. (2019). A conceptual replication of emotional intelligence as a second stratum factor of intelligence. *Emotion.* doi: 10.1037/emo0000569.

Flanagan, D.P., Alfonso, V.C., & Reynolds, M.R. (2013). Broad and narrow CHC abilities measured and not measured by the Wechsler scales: Moving beyond within-battery factor analysis. *Journal of Psychoeducational Assessment, 31,* 202-223. doi: 10.1177/0734282913478047

Flanagan, D.P., & McGrew, K.S. (1998). Interpreting intelligence tests from contemporary gf-gc theory: Joint confirmatory factor analysis of the WJ-R and KAIT in non-white sample. *Journal of School Psychology, 36,* 151-182.

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

http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2004-18831-000&site=ehost-live.
Golay, P., Reverte, I., Rossier, J., Favez, N., & Lecerf, T. (2013). Further insights into the French WISC-IV factor structure through Bayesian structural equation modeling. *Psychological Assessment, 25,* 496-508. doi: 10.1037/a0030676
Gottfredson, L.S. (2002). *g:* Highly general and highly practical. In R.J. Sternberg & E.L. Grigorenko (Eds.), *The general factor of intelligence: How general is it?* Mahwah, NJ: Erlbaum.
Gridley, B.E., & McIntosh, D.E. (1991). Confirmatory factor analysis of the Stanford-Binet: Fourth Edition for a normal sample. *Journal of School Psychology, 29,* 237-248.
Guilford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence.* McGraw-Hill Book Co.
Hakstian, A. R., & Cattell, R. B. (1978). Higher-stratum ability structures on a basis of twenty primary abilities. *Journal of Educational Psychology, 70*, 657-669. doi:10.1037/0022-0663.70.5.657
Hedges, L.V., Tipton, E., & Johnson, M.C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39-65.
Hedlund, J., & Sternberg, R.J. (2000). Too many intelligences? Integrating social, emotional, and practical intelligence. In R. Bar-On & J.D.A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (p.136-167). Jossey-Bass.
Higgin, J.P.T., Thompson, S.G., Deeks, J.J., & Altman, D.G. (2003). Measuring inconsistency in meta-analysis. *BMJ, 327*, 557-560. doi: 10.1136/bmj.327.7414.557
Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*, 253-270. doi:10.1037/h0023816
Hoyle, R. H. & Duval, J. L. (2004). Determining the number of factors in exploratory and confirmatory factor analysis. In Kaplan, D. *The SAGE handbook of quantitative methodology for the social sciences* (pp. 302-317). SAGE Publications, Inc. doi: 10.4135/978141298631
Johnson, W., & Bouchard, T. J. J. (2005). Constructive replication of the visual-perceptual-image rotation model in Thurstone's (1941) battery of 60 tests of mental ability. *Intelligence, 33*, 417-430. doi:10.1016/j.intell.2004.12.001
Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*, 183-202.
Kaplan, S.L., & Alfonso, V.C. (1997). Confirmatory factor analysis of the Stanford-Binet Intelligence Scale: Fourth Edition with preschoolers with developmental delays. *Journal of Psychoeducational Assessment, 15*, 226-237.
Kaufman, A.S. (1993). Joint exploratory factor analysis of the Kaufman Assessment Battery for Children and the Kaufman Adolescent and Adult Intelligence Test for 11- and 12-year-olds. *Journal of Clinical Child Psychology*, *22*, 355-364.
Kaufman, A.S., Kaufman, J.C., & McLean, J.E. (1995). Factor structure of the Kaufman Adolescent and Adult Intelligence Test (KAIT) for White, African Americans, and Hispanics. *Educational and Psychological Measurement*, *55*, 365-376.
Kaufman, S.B., Reynolds, M.R., Liu, X., Kaufman, A.S., & McGrew, K.S. (2012). Are cognitive *g* and academic *g* one and the same *g*? An exploration on the Woodcock-Johnson and Kaufman tests. *Intelligence, 40,* 123-138. doi:10.1016/j.intell.2012.01.009

Keith, T.Z. (1990). Confirmatory and hierarchical confirmatory analysis of the Differential Ability Scales. *Journal of Psychoeducational Assessment, 8,* 291-405. https://doi.org/10.1177/073428299000800314

Keith, T.Z., Fugate, M.H., DeGraff, M., Diamond, C.M., Shadrach, E.A., Stevens, M.L. (1995). Using multi-sample confirmatory factor analysis to test for construct bias: An example using the K-ABC. *Journal of Psychoeducational Assessment, 13,* 347-364.

Keith, T. Z., & Kranzler, J. H. (1999). The absence of structural fidelity precludes construct validity: Rejoinder to Naglieri on what the cognitive assessment system does and does not measure. *School Psychology Review, 28*, 303-321.

Keith, T.Z., Kranzler, J.H., & Flanagan, D.P. (2001). What does the Cognitive Assessment System (CAS) measure? Joint confirmatory factor analysis of the CAS and Woodcock-Johnson Test of Cognitive Ability (3rd Edition). *School Psychology Review, 30,* 89-119.

Kenny, D. A. (2016). Testing in CFA and structural equation modeling. Retrieved from http://davidakenny.net/cm/mfactor.htm

Lecerf, T., & Cavinez, G.L. (2018). Complementary exploratory and confirmatory factor analyses of the French WISC-V: Analysis based on the standardization sample. *Psychological Assessment, 30,* 793-808. doi:10.1037/pas0000638

Legree, P. J., Psotka, J., Robbins, J., Roberts, R. D., Putka, D. J., & Mullins, H. M. (2014). Profile similarity metrics as an alternate framework to score rating-based tests: MSCEIT reanalysis. *Intelligence, 47*, 159-174. doi:10.1016/j.intell.2014.09.005

Leohlin, J.C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis.* Mahway, NJ: Lawrence Erlbaum Associated, Inc.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*, 201–226. doi:10.1146/annurev.psych.51.1.201

MacCann, C. (2010). Further examination of emotional intelligence as a standard intelligence: Latent variable analysis of fluid intelligence, crystallized intelligence, and emotional intelligence. *Personality and Individual Differences, 49,* 490-496. doi: 10.1016/j.paid.2010.05.010

MacCann, C., Joseph, D. L., Newman, D. A., & Roberts, R. D. (2014). Emotional intelligence is a second-stratum factor of intelligence: Evidence from hierarchical and bifactor models. *Emotion, 14*, 358-374. doi:10.1037/a0034755

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

Mayer, J.D. (2018). Intelligences about things and intelligences about people. In R.J. Sternberg (Ed.). *The nature of human intelligence (pp.270-286).* Cambridge, England: Cambridge University Press.

Mayer, J. D., Panter, A. T., & Caruso, D. R. (2012). Does personal intelligence exist? Evidence from a new ability-based measure. *Journal of Personality Assessment, 94,* 124-140. doi: 10.1080/00223891.2011.646108.

Mayer, J. D., Panter, A. T., & Caruso, D. R. (2017). A closer look at the Test of Personal Intelligence (TOPI). *Personality and Individual Differences, 111*, 301-311. doi:10.1016/j.paid.2017.02.008

Mayer, J.D., Salovey, P., & Caruso, D.R. (2002). *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) User's Manual.* Toronto, Canada: Multi-Health Systems.

Mayer, J. D., & Skimmyhorn, W. (2017). Personality attributes that predict cadet performance at west point. *Journal of Research in Personality, 66,* 14-26. doi:10.1016/j.jrp.2016.10.012

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. doi:10.1016/j.intell.2008.08.004

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

McPherson, J., & Burns, N.R. (2007). Gs invaders: Assessing a computer game-like test of processing speed. *Behavior Research Methods, 39,* 876-883.

Merz, Z.C., Van Patten, R., Hurless, N., Grant, A., McGrath, A.B. (2019). Furthering the understanding of the Wechsler Adult Intelligence Scale – Fourth Edition factor structure in a clinical sample. *Applied Neuropsychology: Adult.* doi:10.1080/23279095.2019.1585351

Morgan, K.E., Rothlisberg, B.A., McIntosh, D.E., & Hunt, M.S. (2009). Confirmatory factor analysis of the KABC-II in preschool children. *Psychology in the Schools, 46,* 515-525. doi:10.1002/pits.20394

Muthén, L.K. (2005). *Negative residual variance.* Retrieved from http://www.statmodel.com/discussion/messages/11/555.html?1358188287

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén

Nakano, S., & Watkins, M.W. (2013). Factor structure of the Wechsler Intelligence Scales for Children-Fourth Edition among referred Native American students. *Psychology in the Schools, 50,* 957-968. doi:10.1002/pits.21724

Ng, T. W. H., Sorensen, K. L., Zhang, Y. & Yim, F. H. K. (2019). Anger, anxiety, depression, and negative affect: Convergent or divergent? *Journal of Vocational Behavior 110,* 186-202.

Niileksela, C. R., Reynolds, M. R., & Kaufman, A. S. (2013). An alternative Cattell–Horn–Carroll (CHC) factor structure of the WAIS-IV: Age invariance of an alternative model for ages 70–90. *Psychological Assessment, 25,* 391-404. doi:10.1037/a0031175

Ogata, K. (2015). Joint factor structure of the WISC-III and K-ABC for Japanese maltreated children: A preliminary investigation. *TPM, 22,* 477-483.

Phelps, L., McGrew, K. S., Knopik, S. N., & Ford, L. (2005). The general (g), broad, and narrow CHC Stratum characteristics of the WJ III and WISC-III tests: A confirmatory cross-battery investigation. *School Psychology Quarterly, 20,* 66-88. doi:10.1521/scpq.20.1.66.64191.

Pezzuit, L., & Orsini, A. (2016). Are there sex differences in the Wechsler Intelligence Scale for Children – Fourth Edition? *Learning and Individual Differences, 45,* 307-312. doi:10.1027/1614-0001/a000003

Ree, M. J., Carretta, T. R., & Teachout, M. S. (2015). Pervasiveness of dominant general factors in organizational measurement. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 8,* 409–427. doi:10.1017/iop.2015.16

Reverte, I., Golay, P., Favez, N., Rossier, J., Lecerf, T. (2014). Structural validity of the Wechsler Intelligence Scale for Children (WISC-IV) in a French-speaking Swiss sample. *Learning and Individual Differences, 29,* 114-119. doi:10.1016/j.lindif.2013.10.013

Reynolds, M.R., Hajovsky, D.B., Pace, J.R., & Niileksela, C.R. (2016). What does the Shipley-2 measure for children and adolescents? Integrated and conjoin confirmatory factor analysis with the WISC-IV. *Assessment, 23,* 23-41. doi:10.1177/1073191115572695

Reynolds, M.R., & Keith, T.Z. (2017). Multi-group and hierarchical confirmatory factor analysis of the Wechsler Intelligence Scale for Children – Fifth Edition: What does is measure? *Intelligence, 62,* 31-47. doi:10.1016/j.intell.2017.02.005

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

Rowe, E.W., Dandridge, J., Pawlush, A., Thompson, D.F., & Ferrier, D.E. (2014). Exploratory and confirmatory factor analyses of the WISC-IV with gifted students. *School Psychology Quarterly, 29,* 536-552. doi: 10.1037/spq0000009

Sanders, S., McIntosh, D. E., Dunham, M., Rothlisberg, B. A., & Finch, H. (2007). Joint confirmatory factor analysis of the differential ability scales and the Woodcock-Johnson Tests of Cognitive Abilities-third edition. *Psychology in the Schools, 44*, 119-138. doi:10.1002/pits.20211

Sechrest, L. (1963). Incremental validity: A recommendation. *Educational and Psychological Measurement*, *23,*153-158. doi:10.1177/001316446302300113

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* (p. 73–163). The Guilford Press.

Schneider, W. J., & Newman, D. A. (2015). Intelligence is multidimensional: Theoretical review and implications of specific cognitive abilities. *Human Resource Management Review, 25,* 12-27. doi:10.1016/j.hrmr.2014.09.004

Spearman, C. (1904). 'General intelligence,' objectively determined and measured. *The American Journal of Psychology, 15*, 201-293. doi:10.2307/1412107

Stankov, L. (1978). Fluid and crystallized intelligence and broad perceptual factors among 11 to 12-year olds. *Journal of Educational Psychology, 70*, 324-334. doi:10.1037/0022-0663.70.3.324

Sterne, J.A., Sutton, A.J., Ioannidis, J.P., Terrin, N., Jones, D.R., Lau, J., Carpenter, J., Rücker, G., Harbord, R.M., Schmid, C.H., Tetzlaff, J., Deeks, J.J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D.G., Moher, D., Higgins, J.P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomized controlled trials. *BMJ, 343.* doi: 10.1136/bmj.d4002.

Strickland, T., Watkins, M.W., & Caterino, L.C. (2015). Structure of the Woodcock-Johnson III Cognitive Tests in a referral sample of elementary school students. *Psychological Assessment, 27,* 689-697. doi: 10.1037/pas0000052.

Styck, K.M., & Watkins, M.W. (2017). Structural validity of the WISC-IV for students with ADHD. *Journal of Attention Disorders, 21,* 921-928. doi: 10.1177/1087054714553052

Taub, G.E., & Benson, N. (2013). Matters of consequence: An empirical investigation of the WAIS-III and WAIS-IV and implications for addressing the Atkins Intelligence Criterion. *Journal of Forensic Psychology Practice, 13,* 27-48. doi:10.1080/15228932.2013.746913

Taub, G.E., & McGrew, K.S. (2004). A confirmatory factor analysis of the Cattell-Horn-Carroll theory and cross-age invariance of the Woodcock-Johnson Tests of Cognitive Abilities III. *School Psychology Quarterly, 19,* 72-87. doi:10.1521/scpq.19.1.72.29409

Thaler, N.S., Barchard, K.A., Parke, E., Jones, W.P., Etcoff, L.M., & Allen, D.N. (2015). Factor structure of the Wechsler Intelligence Scale for Children: Fourth Edition in children with ADHD. *Journal of Attention Disorders, 19,* 1013-1021. doi: 10.1177/1087054712459952

Thurstone, L.L. (1937). Psychology as a quantitative rational science. *Science, 85,* 228-232.

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

Undheim, J. O. (1976). Ability structure in 10-11-year-old children and the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 68,* 411-423. https://doi.org/10.1037/0022-0663.68.4.411

Undheim, J. O. (1978). Broad ability factors in 12- to 13-year-old children, the theory of fluid and crystallized intelligence, and the differentiation hypothesis. *Journal of Educational Psychology, 70*, 433-443. doi:10.1037/0022-0663.70.3.433

Undheim, J. O. (1981). On intelligence III: Examining developmental implications of Cattell's broad ability theory and of an alternative neo-Spearman model. *Scandinavian Journal of Psychology, 22*, 243-249. doi:10.1111/j.1467-9450.1981.tb00400.x

Undheim, J. O., & Gustafsson, J. (1987). The hierarchical organization of cognitive abilities: Restoring general intelligence through the use of linear structural relations (LISREL). *Multivariate Behavioral Research, 22,* 149-171. doi: 10.1207/s15327906mbr2202_2.

van Aert, R.C., Witcherts, J.M., van Assen, M.A. (2019). Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis. *PLoSONE, 14,* 1-32. doi:10.1371/journal.pone.0215052

van der Maas, H.L., Dolan, C.V., Grasman, R.P., Wicherts, J.M., Huizenga, H.M., & Raijmakers, M.E. (2006). A dynamic model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological Review, 113*, 842-861. https://doi.org/10.1037/0033-295X.113.4.842

van Driel, O. P. (1978). On various causes of improper solutions in maximum likelihood factor analysis. *Psychometrika*, *43*, 225-243.

Velicer, W. F., & Jackson, D. N. (1990a). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research, 25*, 1-28. doi: 10.1207/s15327906mbr2501_1.

Visser, B.A., Ashton, M.C., & Vernon, P.A. (2006). *g* and the measurement of multiple intelligences: A response to Gardner. *Intelligence, 34,* 507-510. doi:10.1016/j.intell.2006.04.006

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

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*, 817-835. doi:10.1037/a0016127

Waller, N.G., & Waldman, I.D. (1990). A reexamination of the WAIS-R factor structure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 139-144.

Weiss, L.G., Keith, T.Z., Zhu, J., & Chen, H. (2013). WISC-IV and clinical validation of the four- and five- factor interpretative approaches. *Journal of Psychoeducational Assessment, 31,* 114-131. https://doi.org/10.1177/0734282913478032

Williams, T.H., McIntosh, D.E., Dixon, F., Newton, J.H., & Youman, E. (2010) A confirmatory factor analysis of the Stanford-Binet Intelligence Scales, fifth edition, with a high achieving sample. *Psychology in Schools, 47,* 1071-1083. doi:10.1002/pits.20525