Journal of Educational and Developmental Psychology; Vol. 7, No. 1; 2017
ISSN 1927-0526 E-ISSN 1927-0534
Published by Canadian Center of Science and Education

Re(Examining) Relations between CHC Broad and Narrow Cognitive Abilities and Reading Achievement

Ryan J. McGill[1]

[1] School of Education, The College of William and Mary, Williamsburg, Virginia, USA

Correspondence: Ryan J. McGill, School of Education, The College of William and Mary, P. O. Box 8795, Williamsburg, VA, United States. Tel: 1-757-221-6072. E-mail: rmcgill@wm.edu

Received: January 29, 2017 Accepted: March 20, 2017 Online Published: March 27, 2017

doi:10.5539/jedp.v7n1p265 URL: http://doi.org/10.5539/jedp.v7n1p265

Abstract

Previously, Evans and colleagues (2001) utilized simultaneous multiple regression to examine relations between Cattell-Horn-Carroll (CHC; Schneider & McGrew, 2012) broad and narrow cognitive abilities and reading achievement across the school age span. Although their findings suggest that many broad/narrow abilities had clinically significant effects on reading achievement they failed to account for the potential moderating effects of the general factor. To account for these effects, the current study employed hierarchical multiple regression analysis to reexamine the relationships between CHC dimensions and reading achievement after controlling for the effects of the general factor with 4,722 participants ages 6-18 from the Woodcock Johnson III Psychoeducational Battery (WJ III; Woodcock, McGrew, & Mather, 2001a). Results from the present study indicate that the full scale GIA composite (as a proxy for *g*) consistently accounted for large effects across the school age span for all of the reading achievement variables that were assessed. Among the broad and narrow abilities, only Gc consistently accounted for meaningful proportions of reading scores beyond *g*. As a consequence, researchers are encouraged to give greater consideration to the dimensionality of broad and narrow CHC measures when examining cognitive-achievement relationships or they may risk over-interpreting the predictive effects associated with these indices. Potential implications for clinical application of CHC theory are also discussed.

Keywords: incremental validity, CHC, general intelligence

1. Introduction

In psychology and education, the Cattell-Horn-Carroll (CHC) model of human cognitive abilities has emerged as the consensus psychometric-based model for understanding the structure of human intelligence (McGrew, 2009). CHC theory was developed as a synthesis of the Gf-Gc theory (Horn & Cattell, 1966) and Carroll's (1993) three-stratum model. It conceptualizes cognitive abilities within a hierarchical taxonomy in which elements are stratified according to breadth. The most general ability resides at the apex of the model at Stratum III and is referred to as a general factor of intelligence or *g*. The next level (Stratum II) includes seven to nine *broad* abilities (e.g., Fluid Reasoning [Gf], Crystallized Ability [Gc]). At the bottom of the model are over 70 *narrow* abilities (Stratum I) which are organized according to their mapping onto the Stratum II dimensions.

Intelligence testing research has been significantly impacted over the last 15 years as a result of the rise of the CHC model. Additionally, CHC theory has been used to provide a theoretical and empirical blueprint for understanding cognitive abilities and how they may relate to academic learning (Keith & Reynolds, 2010; McGrew & Wendling, 2010). Although several contemporary cognitive tests reference CHC theory within their technical and interpretive manuals, the Woodcock-Johnson III Tests of Cognitive Abilities (WJ III COG; Woodcock, McGrew, & Mather, 2001c) was the first commercial test to utilize CHC specifically for its structural foundation. Additionally, the WJ III COG was the first test reported to measure all of the proposed broad cognitive abilities in the most recent iteration of the CHC model (Schneider & McGrew, 2012). Accordingly, the WJ III COG has been utilized as the *de facto* reference instrument for making refinements to the CHC model and validating many of the inferences made about potential relationships between broad cognitive abilities and academic achievement in the technical and professional literature over the last 15 years (Dombrowski, McGill, & Canivez, 2016).

With regard to CHC cognitive-achievement relationships, a series of predictive validity investigations (Evans et al., 2001; Floyd, Evans, & McGrew, 2003; Floyd, McGrew, & Evans, 2008) completed shortly after the publication of the WJ III COG are particularly noteworthy. In sum, these studies provide evidence of differential predictive effects for CHC-related broad abilities for reading, mathematics, and written language achievement across the age span. In addition to being highly cited (~450 citations) across the school, developmental, and educational psychology literatures, these studies have been instrumental in supporting the dissemination of numerous CHC-related technologies in clinical practice (e.g., Fiorello & Primerano, 2005; Flanagan, Ortiz, & Alfonso, 2013; Flanagan, Ortiz, Alfonso, & Dynda, 2006; Hale et al., 2010; McGrew & Wendling, 2010).

With regard to reading achievement, Evans and colleagues (2001) utilized simultaneous multiple regression to examine relations between CHC broad and narrow abilities and various reading abilities across 14 age groups. Their findings suggest that several abilities (i.e., Phonological Awareness, Processing Speed [Gs], and Long-Term Retrieval [Glr]) consistently accounted for significant effects in reading achievement across the age span. As a result, they encouraged primary consideration of these and other related CHC broad and narrow abilities when investigating reading skill development. However, Evans et al. (2001) did not to include an estimate of general intelligence in their prediction models, a variable that has a rich history of accounting for meaningful levels of academic achievement variance (Buckhalt, 2002; Canivez, 2013b; McGhee, 2002; Naglieri & Bornstein, 2003). To wit, Thorndike (1986) noted that 85% to 90% of predictable variance in measures of achievement may be accounted for by a single general score (i.e., FSIQ), that is thought to estimate general intellectual ability. As a consequence of this omission, the unique contributions of CHC broad and narrow abilities in predicting reading abilities above and beyond a more parsimonious general intelligence dimension was unclear.

According to Pedhazur (1997), leaving out variables that are known to have strong predictive effects on dependent variables is a type of specification error that can lead to misleading regression estimates. Additionally, the results produced from structural validity and latent variable modeling studies raise additional concern about the potential impact of this omission. Several of these studies specifically analyzed the WJ III COG (e.g., Dombrowski, 2013, 2014; Dombrowski & Watkins, 2013; Strickland, Watkins, & Caterino, 2015). In these structural validity studies, it was concluded that the WJ III COG was potentially overfactored (i.e., too many factors extracted), was a solid measure of general intellectual ability, but that caution should be heeded when interpreting the broad and narrow ability scores as singular entities as they were saturated with non-trivial proportions of variance attributable to *g* and lacked enough target construct variance for confident clinical interpretation.

Floyd, Keith, Taub, and McGrew (2007) employed Structural Equation Modeling (SEM) with the WJ III COG to examine the latent predictive effects of CHC-related abilities on reading decoding. In contrast, to Evans et al. (2001), the authors chose to model a general intelligence factor and found that its influence on reading was mediated through the Stratum II broad abilities. Utilizing the same methodological approach, Benson (2007) later concluded that *g* had *direct* and significant effects on reading abilities and that the effects associated with the broad and narrow abilities was mostly small. Alternatively, Floyd, Meisinger, Gregg, and Keith (2012) suggested that an integrative model with both direct and indirect effects from *g* best predicted reading comprehension across development. Although researchers may disagree as to whether *g* has a direct or indirect latent influence on reading, the potential influence of this dimension has been accounted for in a host of psychometric studies examining cognitive-achievement relations in educational and developmental psychology (Canivez, 2013a; Carroll, 1993, 1997; Gottfredson, 1997; McDermott, Goldberg, Watkins, Stanley, & Glutting, 2006; McGill, 2016; Watkins, Lei, & Canivez, 2007).

It is also important to note that researchers examining CHC-achievement relationships have increasingly relied upon SEM and other related latent variable modeling techniques to support the primacy of broad ability interpretation in clinical practice. However, Glutting, Watkins, Konold, and McDermott (2006) argued that reliance on SEM can lead to over-interpretation of spurious constructs. Furthermore, they suggested that "psychologists cannot directly apply results from SEM" (p. 111) because latent scores differ from the observed scores used in practice. As a result, Canivez (2013a) suggested that it is necessary to consider the degree to which observed and latent variable investigations converge, with a preference for the former as those results can be more readily applied to clinical decision-making.

Given the multidimensionality inherent in contemporary measures of CHC-related abilities, researchers (e.g., Beaujean, 2015; Brunner, Nagy, & Wilhelm, 2012; Gignac, 2007, 2016; Gustafsson & Aberg-Bengtsson, 2010; Mansolf & Reise, 2016; Reise, Moore, & Haviland, 2010) have consistently recommended that the effects of the higher-order *g*-factor should be partialed out or controlled for prior to making inferences regarding the relative importance of lower-order cognitive variables. Failure to do so, may risk overestimating the effects of lower-order variables at the expense of the higher-order dimension (Carretta & Ree, 2001; Carroll, 1993, 1995; Chen, Hayes, Carver, Laurenceau, & Zhang, 2012; Gignac, 2007). One of the ways in which this can be achieved at the observed level of measurement is to examine the incremental predictive validity provided by broad ability measures after controlling for the effects of variance already accounted for by the full scale IQ score (as a proxy for *g*).

Hierarchical Multiple Regression Analysis (HMRA) is a well-established statistical procedure for assessing incremental validity in the social sciences. In this procedure, the full scale score is entered first into a regression equation followed by the lower order scores (i.e., factor-level scores, subtests) to predict a criterion achievement variable. This entry technique allows for the predictive effects of the lower-level scores to be assessed while controlling for the effects of the full scale score and operates conceptually in very much the same way as the Schmid and Leiman technique (1957) for residualizing variance in exploratory factor analysis.

Incremental validity studies using HMRA have been conducted previously on the WJ III COG (McGill, 2015a; McGill & Busse, 2015) and other related intelligence tests (e.g., Canivez, 2013a; McGill, 2015b; McGill & Spurgin, 2016). Across these studies, it was consistently demonstrated that the omnibus full scale IQ score accounted for most of the reliable achievement variance that could be predicted in the regression models and that little additional incremental variance was accounted for by the lower-order broad and narrow ability scores after controlling for the predictive effects of the general factor. However, these studies failed to examine predictive effects across a relevant age span. As latent variable modeling studies (McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Tucker-Drob, 2009) have furnished evidence to suggest differential patterns of development for CHC-related broad abilities across the age span, it is possible that some CHC abilities may emerge to account for meaningful portions of variance beyond g at specific age points. In fact, it is this evidence that appeared to guide the analytical strategy employed by Evans and colleagues (2001) in their WJ III COG reading analyses across the school age. Unfortunately, such an investigation has yet to be conducted.

Accordingly, the current study sought to examine the incremental validity of narrow and broad CHC scores on the WJ III COG in predicting variance in reading measures from the Woodcock-Johnson III Tests of Achievement (WJ III ACH; Woodcock, McGrew, & Mather, 2001b) across the school age span (ages 6-18). Given the results of previous incremental validity research, it is believed that a reexamination of the data utilized by Evans et al. (2002) using alternative methods (i.e., HMRA) that account for the effects of the General Intellectual Ability composite score (GIA) may provide a different perspective on CHC cognitive-achievement relations specifically as they relate to reading achievement. Additionally, the present study is the first to examine the incremental validity of cognitive-achievement relationships across the school age span. If the results furnished previously by Evans and colleagues (2001) cannot be replicated in the present study, it may suggest a more circumspect appraisal of the relative importance of CHC broad and narrow abilities in predicting achievement in general and specific reading abilities.

2. Method

2.1 Participants

The participants were children and adolescents ages 6-0 to 18-11 (N = 4,722) drawn from the standardization sample for the WJ III (Woodcock et al., 2001a). Demographic characteristics are provided in detail in the WJ-IV Technical Manuals (McGrew, Schrank, & Woodcock, 2007; McGrew & Woodcock, 2001). The standardization sample was obtained using stratified proportional sampling across demographic variables of sex, race, ethnicity, geographic region, community type, and parent educational level. Examination of the tables in the Technical Manual revealed a close correspondence to the 2005 U. S. census estimates across the stratification variables. The present sample was selected on the basis that it permitted direct comparison to the results furnished by Evans et al. (2001).

2.2 Measurement Instruments

The WJ III COG is a multidimensional test of general intelligence for ages 2 to 90 years. The measure is comprised of 20 subtests, 14 of which contribute to the measurement of seven CHC-based Stratum II broad cluster scores: Comprehension-Knowledge (Gc), Fluid Reasoning (Gf), Auditory Processing (Ga), Visual-Spatial Thinking (Gv), Short-Term Memory (Gsm), Long-Term Retrieval (Glr), and Processing Speed (Gs). Additionally, six *clinical* cluster scores (Phonemic Awareness, Working Memory, Broad Attention, Cognitive Fluency, Executive Processes, and Delayed Recall) thought to reflect more narrow CHC dimensions are also available through different configurations of the subtests. All of the CHC clusters are differentially weighted according to their relative *g* loadings and then combined to form the Stratum III GIA composite. All variables on the WJ III COG are expressed as standard scores with a mean of 100 and a standard deviation of 15. Extensive normative and psychometric data can be found in the WJ III technical manuals (McGrew et al., 2007; McGrew & Woodcock, 2001).

The WJ III-ACH is a comprehensive academic assessment battery designed to measure five academic domains: Reading, Written Language, Mathematics, Oral Language, and Academic Knowledge. The WJ III ACH is comprised of 22 subtests that combine to provide 17 broad clusters and a total achievement composite score. Broad clusters are expressed as standard scores with a mean of 100 and a standard deviation of 15. Mean internal consistency estimates for the included ages in this study ranged from .82 to .96 for the composite and broad scores that were assessed. Additional technical information for the WJ III ACH can be found in the WJ III technical manuals (McGrew et al., 2007; McGrew & Woodcock, 2001).

Table 1. Univariate descriptive statistics for WJ-III cognitive-achievement variables

Variables	M	SD	Skewness	Kurtosis
GIA	100.13	14.27	-0.19	0.46
Comprehension-Knowledge	101.12	13.89	-0.38	0.73
Fluid Reasoning	100.38	14.69	-0.29	0.44
Auditory Processing	100.06	14.43	0.6	0.73
Visual Spatial Thinking	100.13	11.92	-0.13	1.57
Short-Term Memory	100.75	14.35	-0.09	0.82
Long-Term Retrieval	100.31	13.2	-0.04	0.84
Processing Speed	100.33	13.58	-0.14	0.96
Phonemic Awareness	99.05	15.3	-0.11	0.9
Working Memory	99.84	13.9	-0.16	0.99
Broad Attention	99.02	14.41	-0.18	0.85
Cognitive Fluency	100.75	12.94	0.03	1.09
Executive Processes	99.59	13.13	-0.3	0.74
Delayed Recall	99.89	11.09	-0.22	0.74
Broad Reading	100.7	14.89	-0.34	1
Basic Reading Skills	100.66	14.49	-0.33	0.68
Reading Comprehension	100.81	14.79	-0.35	1.11

2.3 Data Analyses

Data analyses for the current study proceeded in several steps. First, participants from the WJ III normative sample ages 6-18 were divided into 13 age brackets. Next, HMRAs were conducted for each age bracket to assess the proportions of WJ III ACH reading cluster score variance accounted for by the observed WJ III COG GIA and CHC broad and narrow cluster scores across the school age. Using, SPSS version 22, the WJ III COG GIA was entered into the first block, and the seven CHC broad and six narrow cluster scores were entered jointly into the second block of the regression equations. Broad and narrow cluster effects also were individually assessed by entering each cluster alone into the second block. The WJ III ACH Broad Reading, Basic Reading Skills, and Reading Comprehension scores were used as criterion variables in the present analyses. The change in the WJ III ACH achievement variance predicted by the broad and narrow cluster scores in the second block of the regression model provided an estimate of the incremental prediction beyond the GIA in the first block of the model. According to Pedhazur (1997), these variance partitioning procedures are appropriate for the goals of the current study.

The results were interpreted using the resulting R^2 statistic as an effect size. Guidelines for interpreting R^2 as an effect size are found in Cohen (1988); they are "small," .01; "medium," .09; and "large," .25. The critical coefficient in HMRA analyses is the incremental squared multiple correlation coefficient (ΔR^2). The ΔR^2 represents the amount of variance that is explained by an Independent Variable (IV) after controlling for the effects of IVs previously entered into a regression equation. At present, there are no conventional guidelines for interpreting the ΔR^2 coefficient, thus Cohen's interpretive framework for R^2 was applied.

2.4 Power Analysis

Given the fact that HMRA analyses was conducted across 13 different age brackets rather than the normative sample as a whole, an *a priori* power analysis was conducted to determine the minimum number of cases needed to exhibit adequate power for this study. Analyses were conducted utilizing G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009), a software tool for general power analysis. Using a linear fixed model multiple regression design for assessing incremental R^2 increase with seven tested predictors (CHC factors) and eight total predictors (inclusion of the GIA in block one of the regression equation), power equal to .80, and an alpha level of .05, it was determined that a sample size of 153 cases was needed to detect a medium effect size ($f^2 = .09$; Cohen, 1988). It is expected that this study will yield small to medium effect sizes based upon the results of previous empirical investigations of the incremental prediction beyond the full-scale score provided by various intelligence test part-scores.

3. Results

The means, standard deviations, skewness, and kurtosis statistics for all of the WJ III cognitive and achievement variables are listed in Table 1. The mean (99.05 to 101.12) and standard deviation ranges (11.09 to 14.89) for the cognitive-achievement variables generally reflect values that would be expected for normally distributed standard score variables. Skewness values ranged from -0.38 to 0.60. Additionally, inspection of the residual plots of the data indicated that the regression models utilized in this study met the assumptions for homoscedasticity of the residuals. The distribution of cases across the 13 age brackets are reported in Table 2. HMRA sample sizes ranged from 249 (age 17) to 579 (age 10).

Table 2. Distribution of cases across age level in the school age sample (7-18)

Age	n
6	308
7	335
8	431
9	533
10	579
11	428
12	352
13	324
14	292
15	302
16	308
17	249
18	281

3.1 Broad Reading

Table A.1 presents the results from hierarchical multiple regression analyses for Broad Reading. The GIA accounted for statistically significant ($p < .05$) portions of the Broad Reading scores in all of the age brackets that were assessed. Across the 13 regression models utilized to predict Broad Reading, the GIA accounted for 48% (age 6) to 70% (age 17; M = 59%) of the criterion variance. The R^2 values that corresponded to those variance increments all reflect large effects using Cohen's (1988) interpretive guidelines. As illustrated in Figure 1, the amount of reliable criterion variance accounted for by the GIA was consistently large whereas, the contributions made by the broad and narrow ability clusters were more modest.

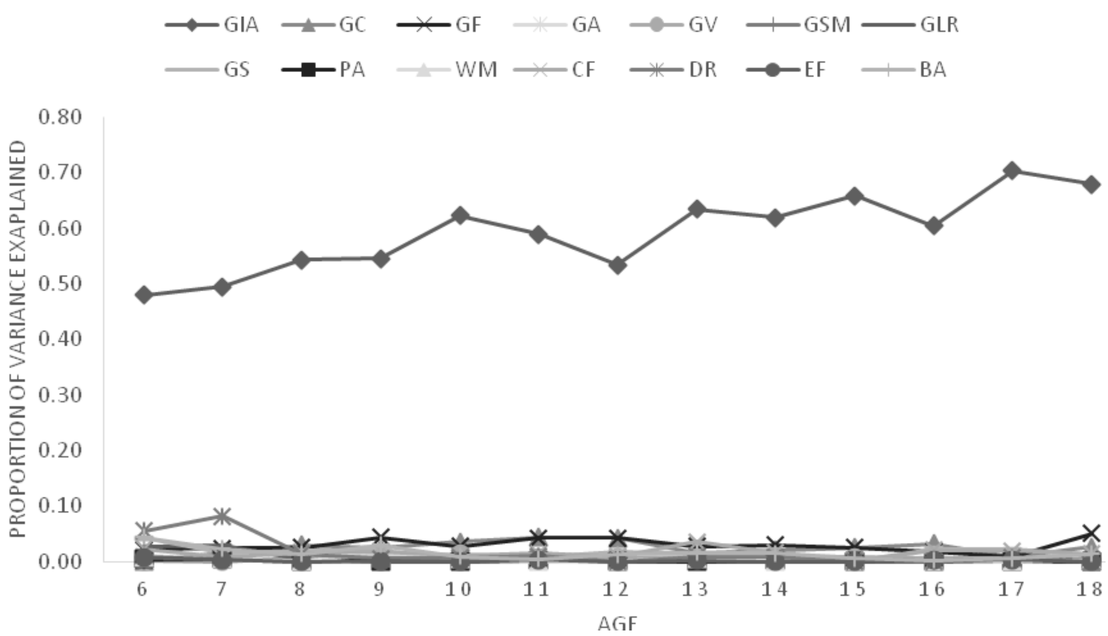

Figure 1. Incremental predictive effects of general and broad/narrow cognitive abilities on Broad Reading

GIA = General Intellectual Ability Composite; GC = Comprehension-Knowledge; GF = Fluid Reasoning; GA = Auditory Processing; GV = Visual-Spatial Thinking; GSM = Short-Term Memory; GLR = Long-Term Retrieval; GS = Processing Speed; PA = Phonemic Awareness; WM = Working Memory; CF = Cognitive Fluency; DR = Delayed Recall; EF = Executive Processes; BA = Broad Attention. Squared multiple correlation coefficient values represents proportion of variance accounted for by variables at their entry point into regression equation after controlling for the effects of the general factor (e.g., $R^2/\Delta R^2$ values multiplied by 100).

Broad clusters entered jointly into the second block of the regression equations accounted for 4% (age 17) to 8% (ages 9, 11, 16; M = 7%) additional variance beyond *g*. The ΔR^2 values that corresponded to those variance increments reflect small effects. The incremental variance coefficients attributed to individual WJ III COG broad clusters ranged from 0% to 5%. Although ANOVA-based tests of significance indicated that the broad clusters on the WJ III COG accounted for significant portions of incremental achievement variance beyond the effects of the GIA, the effect size estimates associated with these indicators were more circumspect.

Narrow clusters entered jointly into the second block of the regression equations accounted for 1% (age 15) to 11% (ages 9, 7; M = 4%) additional variance beyond *g*. The ΔR^2 values that corresponded to those variance increments reflect small to moderate effects. The incremental variance coefficients attributed to individual WJ III COG narrow clusters ranged from 0% to 8% (Delayed Recall, age 7). Although ANOVA-based tests of significance indicated that the narrow clusters on the WJ III COG contributed significant portions of incremental achievement variance beyond the effects of the GIA, effect size estimates were more circumspect.

3.2 Basic Reading Skills

Table A.2 presents the results from hierarchical multiple regression analyses for Basic Reading Skills. The GIA accounted for statistically significant ($p < .05$) portions of the Basic Reading scores in all of the age brackets that were assessed. Across the 13 regression models utilized to predict Basic Reading, the GIA accounted for 40% (age 12) to 63% (age 17; M = 49%) of the criterion variance. The R^2 values that corresponded to those variance increments all reflect large effects using Cohen's interpretive guidelines. As illustrated in Figure 2, the amount of reliable criterion variance accounted for by the GIA was consistently large whereas, the contributions made by the broad and narrow ability clusters were more modest.

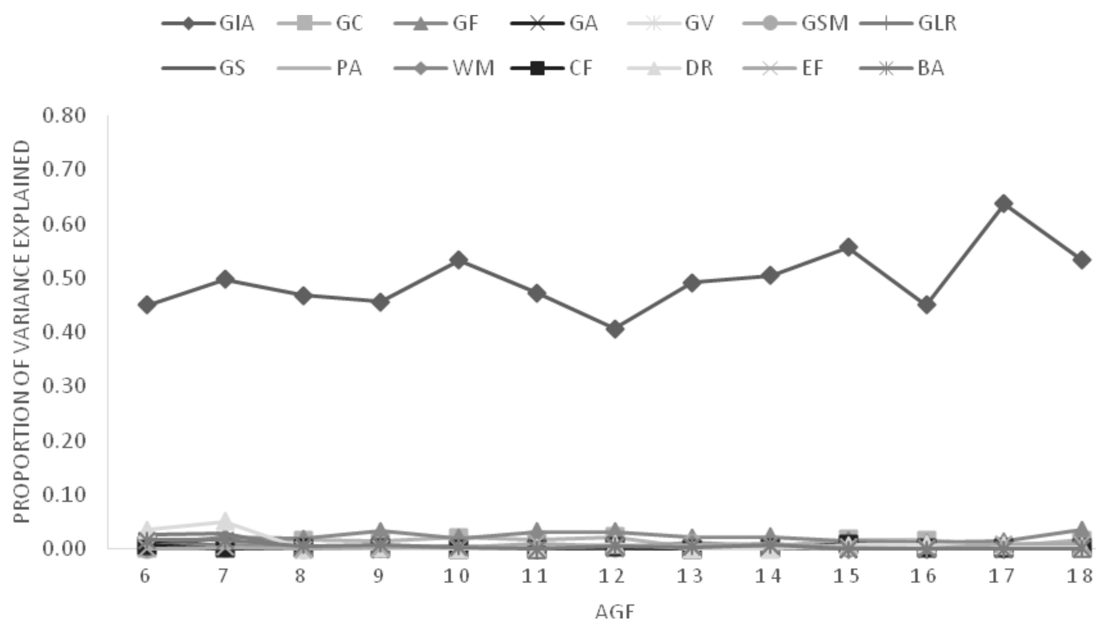

Figure 2. Incremental predictive effects of general and broad/narrow cognitive abilities on Basic Reading

GIA = General Intellectual Ability Composite; GC = Comprehension-Knowledge; GF = Fluid Reasoning; GA = Auditory Processing; GV = Visual-Spatial Thinking; GSM = Short-Term Memory; GLR = Long-Term Retrieval; GS = Processing Speed; PA = Phonemic Awareness; WM = Working Memory; CF = Cognitive Fluency; DR = Delayed Recall; EF = Executive Processes; BA = Broad Attention. Squared multiple correlation coefficient values represents proportion of variance accounted for by variables at their entry point into regression equation after controlling for the effects of the general factor (e.g., $R^2/\Delta R^2$ values multiplied by 100).

Broad clusters entered jointly into the second block of the regression equations accounted for 2% (age 17) to 6% (ages 6; M = 4%) variance beyond g. The ΔR^2 values that corresponded to those variance increments reflect small effects. The incremental variance coefficients attributed to individual WJ III COG broad clusters ranged from 0% to 3%. Although ANOVA-based tests of significance indicated that the broad clusters on the WJ III COG contributed significant portions of incremental achievement variance beyond the effects of the GIA, effect size estimates were not clinically significant.

Narrow clusters entered jointly into the second block of the regression equations accounted for 1% (ages 8-10, 16, 17) to 7% (age 7; M = 2%) additional variance beyond g. The ΔR^2 values that corresponded to those variance increments reflect small effects. The incremental variance coefficients attributed to individual WJ III COG narrow clusters ranged from 0% to 5% (Delayed Recall, age 7). Although ANOVA-based tests of significance indicated that the narrow clusters on the WJ III COG contributed significant portions of incremental achievement variance beyond the effects of the GIA, effect size estimates were more circumspect.

3.3 Reading Comprehension

Table A.3 presents the results from hierarchical multiple regression analyses for Reading Comprehension. The GIA accounted for statistically significant ($p < .05$) portions of the Reading Comprehension scores in all of the age brackets that were assessed. Across the 13 regression models utilized to predict Reading Comprehension, the GIA accounted for 46% (age 6) to 67% (age 17; M = 61%) of the criterion variance. The R^2 values that corresponded to those variance increments all reflect large effects using Cohen's interpretive guidelines. As illustrated in Figure 3, the amount of reliable criterion variance accounted for by the GIA was consistently large whereas, the contributions made by the broad and narrow ability clusters were more modest.

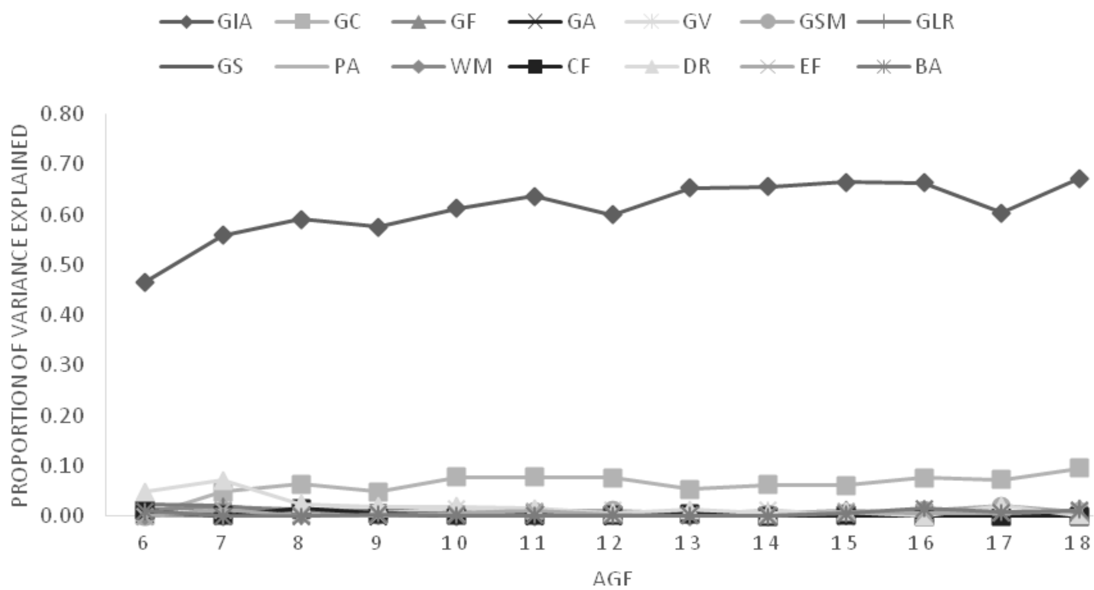

Figure 3. Incremental predictive effects of general and broad/narrow cognitive abilities on Reading Comprehension

GIA = General Intellectual Ability Composite; GC = Comprehension-Knowledge; GF = Fluid Reasoning; GA = Auditory Processing; GV = Visual-Spatial Thinking; GSM = Short-Term Memory; GLR = Long-Term Retrieval; GS = Processing Speed; PA = Phonemic Awareness; WM = Working Memory; CF = Cognitive Fluency; DR = Delayed Recall; EF = Executive Processes; BA = Broad Attention. Squared multiple correlation coefficient values represents proportion of variance accounted for by variables at their entry point into regression equation after controlling for the effects of the general factor (e.g., $R^2/\Delta R^2$ values multiplied by 100).

Broad clusters entered jointly into the second block of the regression equations accounted for 3% (age 6) to 10% (ages 17; *M* = 8%) additional variance beyond *g*. The ΔR^2 values that corresponded to those variance increments reflect small to moderate effects. The incremental variance coefficients attributed to individual WJ III COG broad clusters ranged from 0% to 9%. With only the variance coefficient associated with the Comprehension-Knowledge cluster at age 17 (ΔR^2_{Gc} = .09) accounting for meaningful amounts of achievement variance on its own. Although ANOVA-based tests of significance indicated that the broad clusters on the WJ III COG contributed significant portions of incremental achievement variance beyond the effects of the GIA, most of corresponding effect size estimates were mostly small.

Narrow clusters entered jointly into the second block of the regression equations accounted for 0% (age 14) to 9% (age 7; *M* = 4%) additional variance beyond *g*. The ΔR^2 values that corresponded to those variance increments reflect small to moderate effects. The incremental variance coefficients attributed to individual WJ III COG narrow clusters ranged from 0% to 7% (Delayed Recall, age 7). Although ANOVA-based tests of significance indicated that the narrow clusters on the WJ III COG contributed significant portions of incremental achievement variance beyond the effects of the GIA, effect size estimates were more circumspect.

3.4 Post-Hoc Power Analysis

A post hoc power analysis revealed that for each of the IVs, moderate $R^2/\Delta R^2$ effect sizes (e.g., ≥ .09) could be reliably detected with α set at .05, at a power of greater than .94 in all of the regression models that were estimated in the current study. As a result, the failure to consistently locate meaningful incremental predictive effects for the Stratum I/II variables (as per Evans et al., 2001) in the present study is not likely the result of sampling bias (Nakagawa, 2004).

4. Discussion

The present study reexamined the predictive effects of broad and narrow CHC constructs for reading achievement across the school age span. Although a previous investigation by Evans et al. (2001), posited that several CHC-related abilities on the WJ III COG were clinically significant predictors of reading measures across the same age span, the potential moderating effects of the general factor were not accounted for in their study. As noted by Rodriguez, Reise, and Haviland (2016), “unless all relevant variables are in the predictor space, one cannot know the true unique relation between predictors and the criterion” (p. 233).

Subsequent structural and latent variable modeling studies conducted on the WJ III COG suggest that (a) many of the broad and narrow ability scores contain large portions of variance attributable to *g* (Dombrowski, 2013, 2014; Dombrowski & Watkins, 2013), (b) *g* exerts strong direct or indirect influence on reading achievement (Benson, 2007; Floyd et al., 2007; Floyd et al., 2012; Vanderwood, McGrew, Flanagan, & Keith, 2002), and (c) when the predictive effects *g* are controlled for, the aggregate incremental contributions of broad and narrow abilities may be small (McGill, 2015; McGill & Busse, 2015). As a consequence, the present study sought to reexamine the findings produced by Evans and colleagues (2001) utilizing an alternative analytical scheme (HMRA) to determine the extent to which WJ III COG broad and narrow cluster scores provided meaningful improvements in the prediction of WJ III ACH reading scores *beyond* the GIA composite across the same age range.

Results from the present study indicate that the GIA score consistently accounted for statistically significant and large effects across the school age for all of the reading achievement variables that were assessed (R^2 coefficients ranged from .49 to .61). While the broad and narrow abilities as a whole accounted for moderate increments in prediction beyond *g* for Broad Reading (ages 6-7) and Reading Comprehension (Ages 7, 10, 18), the contributions of individual indicators were mostly trivial. Whereas several broad and narrow abilities accounted for statistically significant effects, the magnitude of these effects was consistently small (.00 to .06) after controlling for *g*. However, meaningful incremental prediction was accounted for by several Stratum I/II variables (e.g., Crystallized Ability, Delayed Recall) for some reading measures (Basic Reading Skills and Reading Comprehension). Additionally, despite evidence suggesting that latent CHC broad and narrow constructs differentiate across the age span, the manifestation of these effects in the current study was less consistent. Whereas, linear increases in prediction was observed for Gc in the Reading Comprehension model, this same effect on other areas of reading was not observed. In contrast, the predictive effects of the GIA score (as a proxy for *g*) increased linearly across school age in all of the regression models that were assessed suggesting that the accumulative effects of general ability may render the GIA a more robust predictor of reading achievement as individual’s progress through school.

The present results diverge from those produced from Evans et al. (2001) in several ways. First, although Evans and colleagues suggested that Phonological Awareness, Gs, and Glr were clinically significant predictors of WJ III ACH reading abilities, the unique contributions of those indicators was more modest in the present study. More importantly, the predictive effects associated with broad and narrow dimensions as a whole were manifestly weaker once the effects of the GIA were accounted for. Given the WJ III structural validity findings furnished by Dombrowski (2013, 2014) and Dombrowski and Watkins (2013) indicating that these measures mostly sample general intelligence, the latter finding was not surprising.

The present study is the first to describe the incremental validity of cognitive-achievement relationships across a relevant age span. Whereas, related predictive validity studies with cognitive measures (e.g., Benson, Kranzler, & Floyd, 2016; Canivez, 2013a; McGill, 2015a, 2015b; McGill & Spurgin, 2016) have largely suggested that broad and narrow cognitive abilities account for trivial portions of achievement after controlling for general ability, the present results suggest that this conclusion as a general rule may be overstated. As previously mentioned, although the GIA consistently accounted for large reading effects, lower-order scores accounted for meaningful incremental prediction in several circumstances.

Specifically, the present analyses indicate well that Gc may account for meaningful reading variance beyond the GIA. This finding is consistent with the corpus of latent variable modeling research examining the predictive effects of CHC dimensions on reading achievement (Beaujean, Parkin, & Parker, 2014; Benson, 2007; Floyd et al., 2007; Floyd et al., 2012; Vanderwood et al., 2002) and is consonant with the *investment theory* first proposed by Cattell (1971). Cattell argued that cognitive resources are invested selectively in the environment, resulting in the development of specific broad abilities over others. As a result, when predicting reading achievement, it may be beneficial to go beyond *g*. Nevertheless, the failure to replicate the broader results produced by Evans et al. (2001), suggest that a more circumspect appraisal of the importance of CHC dimensions in relationship to the development of reading skills may be needed in the professional literature.

4.1 Study Limitations and Future Directions

This study is not without limitations that should be considered when interpreting the results. Most notably, the WJ III has recently been revised and is now currently in its fourth edition (WJ IV; Schrank, McGrew, & Mather, 2014). Nevertheless, the current study utilized data from the previous measurement instrument for several reasons. First, as the stated goal of the study was ostensibly to reexamine the results furnished by Evans et al. (2001), multiple regression analyses with the same sample and measurement instrument were necessary for a direct comparison to these results given the CHC content and structural changes implemented in the WJ IV (see McGrew, LaForte, & Schrank, 2014). Second, over the last 15 years the WJ III has served as the preeminent reference instrument for making refinements to the CHC model (McGrew, 2009; McGrew & Wendling, 2010; Schneider & McGrew, 2012) as well as for understanding broader cognitive-achievement relationships in educational and developmental psychology. While there is no doubt that the WJ IV is poised to take its place, the so-called "reproducibility crisis" (Pashler & Wagenmakers, 2012) in scientific psychology illustrates well that it is sometimes beneficial to reevaluate the evidence-base for widely accepted theories (or recommended application of those theories) in light of new developments by researchers. Relatedly, given the critical role that the WJ battery has played in the development of CHC theory, it is believed that the present results will be important for linking and establishing an *evidentiary chain* with related analyses on the WJ IV and other CHC-related measurement instruments.

As previously noted, structural validity studies on the WJ III COG suggest that the broad and narrow dimensions may not be measured well, if at all, apart from general intelligence across the school-age. Nevertheless, structural validity is necessary but not singularly sufficient for establishing construct validity. As a consequence, additional examinations of concurrent and predictive relationships with external measures are also important elements of scale validation (Canivez, 2013b; Cronbach & Meehl, 1955). According to Schneider, Mayer, and Newman (2016) it is not enough that a factor analysis supports the existence of a latent dimension, that factor should also "predict something that matters, above and beyond the other facets of intelligence" (p. 12). The current results add to a growing literature base suggesting that most of the predictive variance accounted for by broad and narrow CHC dimensions on the WJ and related measures can be sourced to a more global general intelligence dimension. As a consequence, Cucina and Howardson (2016) suggest that even if one accepts the legitimacy of all posited CHC vectors, these findings indicate that these dimensions may not be sampled well by existing measures. As noted long ago by McGhee (2002), "Deconstruction of *g* into more narrowly defined primary abilities does not allow for greater diagnostic interpretation of strengths and weaknesses, but steers away from how these abilities are integrated" (p. 201).

Related research on the WJ IV is presently starting to accumulate however, a recent structural validity study of the WJ IV COG (Dombrowski et al., 2016) failed to locate several posited CHC broad abilities and, like its predecessor, that general intelligence accounted for most of the reliable variance in the lower-order measures. Accordingly, researchers have also begun to update the previous series of cognitive-achievement relations studies (Evans et al., 2002; Floyd et al., 2003, 2008) with the WJ IV. Whereas, Cormier, McGrew, Bulut, and Funamoto (2016) included the general factor in their reading achievement analyses, this variable was not included in a related analyses of writing achievement produced by the same research team (Cormier et al., 2016). As noted long ago by Meehl (1990), one can only have confidence in a theory when it has been subjected to a "risky" empirical test, with due consideration to the specification of appropriate parameters (i.e., *g*) which may obviate the importance of the target variables in question (lower-order CHC variables). Additionally, both studies failed to cite any of the previous research questioning the structural or predictive validity of CHC-related scores on the WJ III which is vital for informing future studies with the WJ IV and other related CHC instruments (Dombrowski et al., 2016).

It is also important to point out that the present study relied upon observed standard scores. In contrast, Evans et al. (2001) assessed the predictive effects of broad and narrow abilities using *W*-scores which are derived from a 1-parameter Rasch measurement model. Nevertheless, it should be noted that WJ III standard scores used in the present study are derived utilizing the same *W* scale as a reference anchor (Jaffe, 2009).

Although conventional guidelines (Cohen, 1988) for interpreting the R^2 statistic as an effect size were employed in the present study, methodologists (Dawes, 1999; Keith, 2015) have cautioned against ridged application of these criteria to make inferences about the relative importance of variables. Schneider and Newman (2015) noted that while broad abilities have historically accounted for modest levels of incremental prediction after accounting for general intelligence, this should not automatically preclude their consideration in practical applications of intelligence testing. That is, the relative importance of cognitive variables as it relates to prediction remains very much in the eye of the beholder. Whereas some clinicians may find an additional 4% of predicted variance to be beneficial, others may balk at the cost in time and assessment resources needed to obtain these modest increments.

Finally, as in other incremental predictive validity investigations, multicollinearity of the GIA and the broad and narrow cluster scores in the hierarchical multiple regression analyses was observed in the present study due to the linear combination of subtests to produce cluster scores and the GIA. However, it should be noted that multicollinearity is not a threat to validity in regression studies that are limited to interpreting the R^2 statistic (Cohen, Cohen, West, & Aiken, 2003; Pedhazur, 1997; Tabachnick & Fidell, 2013), nor does it invalidate the use of HMRA to detect improvements in R^2 such as those provided by the broad and narrow scores beyond the GIA (Schneider, 2008). Although, it has long been suggested (e.g., Hale, Fiorello, Kavanagh, Holdnack, & Aloe, 2007; Keith, 2015) that the predictive effects of lower-order scores are suppressed because they lack the freedom to vary from global scores such as the GIA, this argument fails to take into consideration that the correlations between broad/narrow dimensions and *g* are far from unitary. As previously mentioned, the GIA composite is differentially weighted based upon the *g* loadings from its constituent measures. To buttress this position, it is worth noting that in the present study the Gs cluster consistently produced some of the weakest residual predictive effects despite having a relatively low *g* loading. In contrast, the Gc cluster consistently produced the largest residual predictive estimates in the despite its relatively large contribution to the overall GIA score.

Relatedly, due to the hierarchical structure of the measurement instrument, the importance of order of entry when utilizing HMRA to assess the incremental effects of IVs must also be considered. Hale, Fiorello, Kavanaugh, Holdnak, and Aloe (2007) demonstrated that by entering the first-order factor scores from a previous iteration of the Wechsler Intelligence Scale prior to entering the FSIQ score, the predictive effects of FSIQ were diminished to the point of being inconsequential. As a result, Hale and colleagues argued that order of entry arbitrarily determines whether scores such as the GIA mean *everything* or *nothing* due to the long established fact that variables entered first into a regression equation capture greater criterion variance than variables entered later (Cohen et al., 2003). However, order of entry is not an arbitrary process and must be determined *a priori* according to the expected theoretical relationships between variables (Pedhazur, 1997). The proposed indirect hierarchical structural model for the WJ III COG (as per CHC theory) support entering the GIA score prior to the broad and narrow clusters due to the fact that these scores are subordinate to the GIA. Further, reverse entry conflicts with CHC theory and constitutes a violation of the scientific law of parsimony.

5. Conclusion

In sum, the current reexamination of previous results for reading achievement, coupled with independent structural and incremental predictive validity investigations of the WJ battery and other CHC-related instruments (e.g., Benson et al., 2016; Canivez, 2008; Canivez & McGill, 2016; Canivez, Watkins, & Dombrowski, 2016a, 2016b; DiStefano & Dombrowski, 2006; McGill & Spurgin, 2015, 2016) coalesce to suggest a more circumspect appraisal of the empirical and practical importance of many broad and narrow abilities within the CHC lexicon as measured by commercial ability tests. Nevertheless, clinical applications of the CHC model (e.g., Fiorelllo & Primano, 2015; Flanagan, Ortiz, & Alfonso, 2013) in education and psychology continue to focus disproportionally on the importance of broad and narrow abilities in isolation, with little regard for *g* despite the influence of Carroll's (1993) treatise on the broader CHC marriage. Although well intentioned, a countering body of scientific literature is presently accumulating to suggest that these practices may not be psychometrically defensible.

Acknowledgments

The WJ-III standard score data utilized in this study was used with permission from the Woodcock-Muñoz Foundation. A preliminary version of this research was presented at the 2015 annual meeting of the International Society for Intelligence Research.

References

Beaujean, A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher-order models. *Journal of Intelligence*, *3*, 121-136. https://doi.org/10.3390/jintelligence3040121

Beaujean, A. A., Parkin, J., & Parker, S. (2014). Comparing Cattell-Horn-Carroll factor models: Differences between bifactor and higher order factor models in predicting language achievement. *Psychological Assessment*, *26*, 789-805. https://doi.org/10.1037/a0036745

Benson, N. (2007). Cattell-Horn-Carroll cognitive abilities and reading achievement. *Journal of Psychoeducational Assessment*, *26*, 27-41. https://doi.org/10.1177/0734282907301424

Benson, N. F., Kranzler, J. H., & Floyd, R. G. (2016). Examining the integrity of measurement of cognitive abilities in the prediction of achievement: Comparisons and contrasts across variables from higher-order and bifactor models. *Journal of School Psychology*, *58*, 1-19. https://doi.org/10.1016/j.jsp.2016.06.001

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, *80*, 796-846. https://doi.org/10.1111/j.1467-6494.2011.00749.x

Buckhalt, J. A. (2002). A short history of g: Psychometrics' most enduring and controversial construct. *Learning and Individual Differences*, *13*, 101-114. https://doi.org/10.1016/S1041-6080(02)00074-2

Canivez, G. L. (2008). Orthogonal higher order factor structure of the Stanford-Binet Intelligence Scales-fifth edition for children and adolescents. *School Psychology Quarterly*, *23*, 533-541. https://doi.org/10.1037/a0012884

Canivez, G. L. (2013a). Incremental criterion validity of WAIS-IV factor index scores: Relationships with WIAT-II and WIAT-III subtest and composite scores. *Psychological Assessment*, *25*, 484-495. https://doi.org/10.1037/a0032092

Canivez, G. L. (2013b). Psychometric versus actuarial interpretation of intelligence and relates aptitude batteries. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.), *The Oxford handbook of child psychological assessment* (pp. 84-112). New York: Oxford University Press.

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

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

Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2016b). Structural validity of the Wechsler Intelligence Scale for Children-Fifth Edition: Confirmatory factor analyses with the 16 primary and secondary Subtests. *Psychological Assessment*. https://doi.org/10.1037/pas0000358

Carretta, T. R., & Ree, M. J. (2001). Pitfalls of ability research. *International Journal of Selection and Assessment*, *9*, 325-335. https://doi.org/10.1111/1468-2389.00184

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

Carroll, J. B. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research*, *30*, 429-452. https://doi.org/10.1207/s15327906mbr3003_6

Carroll, J. B. (1997). Psychometrics, intelligence, and public perception. *Intelligence*, *24*, 25-52. https://doi.org/10.1016/S0160-2896(97)90012-X

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

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*, 219-251. https://doi.org/10.1111/j.1467-6494.2011.00739.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Psychology Press.

Cohen, J., Cohen, P., West, S. G., & Aiken, L., S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, N.J: L. Erlbaum Associates.

Cormier, D. C., Bulut, O., McGrew, K. S., & Frison, J. (2016). The role of Cattell-Horn-Carroll (CHC) cognitive abilities in predicting writing achievement during the school-age years. *Psychology in the Schools*, *53*, 787-803. https://doi.org/10.1002/pits.21945

Cormier, D. C., McGrew, K. S., Bulut, O., & Funamoto, A. (2016). Revisiting the relations between the WJ-IV measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during the school-age years. *Journal of Psychoeducational Assessment*. https://doi.org/10.1177/0734282916659208

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302. https://doi.org/10.1037/h0040957

Cucina, J. M., & Howardson, G. N. (2016). Woodcock-Johnson-III, Kaufman Adolescent and Adult Intelligence Test (KAIT), Kaufman Assessment Battery for Children (KABC), and Differential Ability Scales (DAS) Support Carroll but not Cattell-Horn. *Psychological Assessment*. https://doi.org/10.1037/pas0000389

Dawes, R. M. (1999). Two methods for studying the incremental validity of a Rorschach variable. *Psychological Assessment*, *11*, 297-302. https://doi.org/10.1037/1040-3590.11.3.297

DiStefano, C., & Dombrowski, S. C. (2006). Investigating the theoretical structure of the Stanford-Binet-Fifth Edition. *Journal of Psychoeducational Assessment*, *24*, 123-136. https://doi.org/10.1177/0734282905285244

Dombrowski, S. C. (2013). Investigating the structure of the WJ-III Cognitive at school age. *School Psychology Quarterly*, *28*, 154-169. https://doi.org/10.1037/spq0000010

Dombrowski, S. C. (2014). Investigating the structure of the WJ-III Cognitive in early school age through two exploratory bifactor analysis procedures. *Journal of Psychoeducational Assessment*, *32*, 483-494. https://doi.org/10.1177/0734282914530838

Dombrowski, S. C., McGill, R. J., & Canivez, G. L. (2016). Exploratory and hierarchical factor analysis of the WJ-IV Cognitive at school age. *Psychological Assessment*. https://doi.org/10.1037/pas0000350

Dombrowski, S. C., & Watkins, M. W. (2013). Exploratory and higher order factor analysis of the WJ-III full test battery: A school-aged analysis. *Psychological Assessment*, *25*, 442-455. https://doi.org/10.1037/a0031335

Evans, J. J., Floyd, R. G., McGrew, K. S., & LeForgee, M. H. (2001). The relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review*, *31*, 246-262. Retrieved from http://www.nasponline.org

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Fiorello, C. A., & Primerano, D. (2005). Research into practice: Cattell-Horn-Carroll cognitive assessment in practice: Eligibility and program development issues. *Psychology in the Schools*, *42*, 525-536. https://doi.org/10.1002/pits.20089

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

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Dynda, A. M. (2006). Integration of response to intervention and norm-referenced tests in learning disability identification: Learning from the Tower of Babel. *Psychology in the Schools*, *43*, 807-825. https://doi.org/10.1002/pits.20190

Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and mathematics achievement across the school-age years. *Psychology in the Schools*, *40*, 155-171. https://doi.org/10.1002/pits.10083

Floyd, R. G., Keith, T. Z., Taub, G. E., & McGrew, K. S. (2007). Cattell-Horn-Carroll cognitive abilities and their effects on reading decoding skills: G has indirect effects, more specific abilities have direct effects. *School Psychology Quarterly*, *22*, 200-233. https://doi.org/10.1037/1045-3830.22.2.200

Floyd, R. G., McGrew, K. S., & Evans, J. J. (2008). The relative contributions of the Cattell-Horn-Carroll cognitive abilities in explaining writing achievement during childhood and adolescence. *Psychology in the Schools*, *45*, 132-144. https://doi.org/10.1002/pits.20284

Floyd, R., Meisinger, E., Gregg, N., & Keith, T. (2012). An explanation of reading comprehension across development using models from Cattell-Horn-Carroll theory: Support for integrative models of reading. *Psychology in the Schools*, *49*, 725-743. https://doi.org/10.1002/pits.21633

Gignac, G. E. (2007). Multi-factor modeling in individual differences research: Some recommendations and suggestions. *Personality and Individual Differences*, *42*, 37-48. https://doi.org/10.1016/j.paid.2006.06.019

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence*, *55*, 57-68. https://doi.org/10.1016/j.intell.2016.01.006

Glutting, J. J., Watkins, M. W., Konold, T. R., & McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC-IV in estimating reading and math achievement on the WIAT-II. *Journal of Special Education*, *40*, 103-114. https://doi.org/10.1177/00224669060400020101

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, *24*, 79-132. https://doi.org/10.1016/S0160-2896(97)90014-3

Gustafsson, J.-E., & Aberg-Bengtsson, L. (2010). Unidimensionality and interpretability of psychological instruments. In S. E. Embretson (Ed.), *Measuring psychological constructs: Advances in model-based approches* (pp. 97-122). Washington, DC: American Psychological Association. https://doi.org/10.1037/12074-005

Hale, J., Alfonso, V., Berninger, V., Bracken, B., Christo, C., Clark, E., … Yalof, J. (2010). Critical issues in response-to-intervention, comprehensive evaluation, and specific learning disabilities identification and intervention: An expert white paper consensus. *Learning Disability Quarterly*, *33*, 223-236. https://doi.org/10.1177/073194871003300310

Hale, J. B., Fiorello, C. A., Kavanagh, J. A., Holdnack, J. A., & Aloe, A. M. (2007). Is the demise of IQ interpretation justified? A response to special issue authors. *Applied Neuropsychology*, *14*, 37-51. https://doi.org/10.1080/09084280701280445

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. https://doi.org/10.1037/h0023816

Jaffe, L. E. (2009). *Development, interpretation, and application of the W score and the relative proficiency index (Woodcock-Johnson III Assessment Service Bulletin No. 11)*. Rolling Meadows, IL: Riverside Publishing.

Keith, T. Z. (2015). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling* (2nd ed.). New York: Routledge.

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*, 635-650. https://doi.org/10.1002/pits.20496

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. *Multivariate Behavioral Research*, *51*, 698-717. https://doi.org/10.1080/00273171.2016.1215898

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology*, *38*, 115-142. https://doi.org/10.1037/0012-1649.38.1.115

McDermott, P. A., Goldberg, M. M., Watkins, M. W., Stanley, J. L., & Glutting, J. J. (2006). A nationwide epidemiologic modeling study of LD: Risk, protection, and unintended impact. *Journal of Learning Disabilities*, *39*, 230-251. https://doi.org/10.1177/00222194060390030401

McGhee, R. L. (2002). The McGhee prophecies: Commentary on "Is g a viable construct for school psychology"? *Learning and Individual Differences*, *13*, 197-203. https://doi.org/10.1016/S1041-6080(02)00079-1

McGill, R. J. (2015a). Incremental criterion validity of the WJ-III COG clinical clusters: Marginal predictive effects beyond the general factor. *Canadian Journal of School Psychology*, *30*, 51-63. https://doi.org/10.1177/0829573514553926

McGill, R. J. (2015b). Interpretation of KABC-II scores: An evaluation of the incremental validity of Cattell-Horn-Carroll (CHC) factor scores in predicting achievement. *Psychological Assessment*, *27*, 1417-1426. https://doi.org/10.1037/pas0000127

McGill, R. J. (2016). Invalidating the full scale IQ score in the presence of significant factor score variability: Clinical acumen or clinical illusion? *Archives of Assessment Psychology*, *6*(1), 49-79. Retrieved from http://www. assessmentpsychologyboard.org/journal/index.php/AAP

McGill, R. J., & Busse, R. T. (2015). Incremental validity of the WJ III COG: Limited predictive effects beyond the GIA-E. *School Psychology Quarterly*, *30*, 353-365. https://doi.org/10.1037/spq0000094

McGill, R. J., & Spurgin, A. R. (2015). Exploratory higher order analysis of the Luria interpretive model on the Kaufman Assessment Battery for Children-Second Edition (KABC-II) school-age battery. *Assessment*. https://doi.org/10.1177/1073191115614081

McGill, R. J., & Spurgin, A. R. (2016). Assessing the incremental value of KABC-II Luria model scores in predicting achievement: What do they tell us beyond the MPI? *Psychology in the Schools*, *53*, 677-689. https://doi.org/10.1002/pits.21940

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

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

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). *Woodcock-Johnson III Normative Update technical manual*. Rolling Meadows, IL: Riverside Publishing.

McGrew, K. S., & Wendling, B. J. (2010). Cattell-Horn-Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools*, *47*, 651-675. https://doi.org/10.1002/pits.20497

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

Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry*, *1*, 108-141. https://doi.org/10.1207/s15327965pli0102_1

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment*, *21*, 244-260. https://doi.org/10.1177/073428290302100302

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology*, *15*, 1044-1045. https://doi.org/10.1093/beheco/arh107

Pashler, H., & Wagenmakers, E.-J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, *7*, 528-530. https://doi.org/10.1177/1745691612465253

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). New York: Holt, Rinehart, & Winston.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*, 544-559. https://doi.org/10.1080/00223891.2010.496477

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, *98*, 223-237. https://doi.org/10.1080/00223891.2015.1089249

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*, 53-61. https://doi.org/10.1007/BF02289209

Schneider, W. J. (2008). Playing statistical ouija board with commonality analysis: Good questions, wrong assumptions. *Applied Neuropsychology*, *15*, 44-53. https://doi.org/10.1080/09084280801917566

Schneider, W. J. (2013). What if we took our models seriously? Estimating latent scores in individuals. *Journal of Psychoeducational Assessment*, *31*, 186-201. https://doi.org/10.1177/0734282913478046

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99-114). New York: 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. https://doi.org/10.1016/j.hrmr.2014.09.004

Schneider, W., Mayer, J., & Newman, D. (2016). Integrating hot and cool intelligences: Thinking broadly about broad abilities. *Journal of Intelligence*, *4*, 1-25.

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

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. https://doi.org/10.1037/pas0000052

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson Education.

Thorndike, R. L. (1986). The role of general ability in prediction. *Journal of Vocational Behavior*, *29*, 332-339. https://doi.org/10.1016/0001-8791(86)90012-6

Tucker-Drob, E. M. (2009). Differentiation of cognitive abilities across the life span. *Developmental Psychology*, *45*, 1097-1118. https://doi.org/10.1037/a0015864

Vanderwood, M. L., McGrew, K. S., Flanagan, D. P., & Keith, T. Z. (2002). The contribution of general and specific cognitive abilities to reading achievement. *Learning and Individual Differences*, *13*, 159-188. https://doi.org/10.1016/S1041-6080(02)00077-8

Watkins, M. W., Lei, P.-W., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence*, *35*, 59-68. https://doi.org/10.1016/j.intell.2006.04.005

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

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001b). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001c). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Appendix A

Table A.1 Incremental contribution of CHC cognitive abilities in predicting broad reading beyond the general factor across the school age

	6	7	8	9	10	11	12	13	14	15	16	17	18	*M*
GIA	0.48*	0.49*	0.54*	0.54*	0.62*	0.59*	0.53*	0.63*	0.61*	0.65*	0.60*	0.70*	0.68*	0.59
CHC (df = 7)[a]	0.07*	0.06*	0.07*	0.08*	0.07*	0.08*	0.10*	0.06*	0.06*	0.05*	0.08*	0.04*	0.07*	0.07
Gc	0.00	0.01*	0.03*	0.02*	0.03*	0.04*	0.04*	0.01*	0.02*	0.02*	0.03*	0.00*	0.02*	0.02
Gf	0.02*	0.02*	0.02*	0.04*	0.02*	0.04*	0.04*	0.02*	0.03*	0.02*	0.01*	0.00*	0.05*	0.02
Ga	0.01*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00*	0.00	0.00	0.00
Gv	0.00	0.00	0.00*	0.00	0.01*	0.00*	0.00	0.00*	0.01*	0.00*	0.00	0.00	0.00	0.00
Gsm	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Glr	0.02*	0.02*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gs	0.04*	0.01*	0.01*	0.02	0.00	0.00*	0.01*	0.01*	0.01*	0.00	0.02*	0.02*	0.01*	0.01
Clinical (df = 6)[a]	0.10*	0.11*	0.04*	0.05*	0.02*	0.03*	0.03*	0.04*	0.04*	0.01*	0.02*	0.03*	0.03*	0.04
PA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
WM	0.00*	0.00	0.00	0.01*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CF	0.02*	0.00*	0.01*	0.01*	0.01*	0.01*	0.00*	0.03*	0.01*	0.00	0.02*	0.01*	0.01*	0.01
DR	0.05*	0.08*	0.01*	0.00	0.00	0.01*	0.00	0.00*	0.00*	0.00	0.00	0.00	0.00	0.01
EF	0.00*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BA	0.04*	0.02*	0.01*	0.02*	0.00*	0.00	0.01*	0.01*	0.01*	0.00	0.00	0.00	0.01*	0.01

Note. GIA = General Intellectual Ability; CHC = Cattell-Horn-Carroll; Gc = Comprehension-Knowledge; Gf = Fluid Reasoning; Ga = Auditory Processing; Gv = Visual-Spatial Thinking; Gsm = Short-Term Memory; Glr = Long-Term Retrieval; Gs = Processing Speed; PA = Phonemic Awareness; WM = Working Memory; CF = Cognitive Fluency; DR = Delayed Recall; EF = Executive Processes; BA = Broad Attention. Coefficients represent the proportion of variance accounted for by variables at their entry point into regression equation. [a]Degrees of freedom reflects controlling for the effects of the GIA. *$p < .05$.

Table A.2 Incremental contribution of CHC cognitive abilities in predicting basic reading skills beyond the general factor across the school age

	6	7	8	9	10	11	12	13	14	15	16	17	18	*M*
GIA	0.45*	0.49*	0.46*	0.45*	0.53*	0.47*	0.40*	0.49*	0.50*	0.55*	0.45*	0.63*	0.53*	0.49
CHC (df = 7)[a]	0.06*	0.05*	0.03*	0.04*	0.03*	0.03*	0.05*	0.03*	0.03*	0.03*	0.04*	0.02*	0.05*	0.04
Gc	0.00	0.01*	0.01*	0.01*	0.02*	0.01*	0.02*	0.00	0.00	0.01*	0.01*	0.00	0.01*	0.01
Gf	0.01*	0.02*	0.01*	0.03*	0.02*	0.03*	0.03*	0.02*	0.02*	0.01*	0.01*	0.01*	0.03*	0.02
Ga	0.01*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00*	0.00	0.00	0.00	0.00	0.00
Gv	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gsm	0.00	0.00	0.00	0.00*	0.00	0.00	0.00	0.01*	0.00	0.00	0.00	0.00	0.00	0.00
Glr	0.02*	0.02*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00*	0.00	0.00	0.00	0.00
Gs	0.01*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Clinical (df = 6)[a]	0.06*	0.07*	0.01*	0.01*	0.01*	0.02*	0.04*	0.02*	0.03*	0.02*	0.01*	0.02*	0.01*	0.02
PA	0.00*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00*	0.00	0.00	0.00	0.00	0.00
WM	0.00	0.00*	0.00	0.00	0.00	0.00	0.00	0.01*	0.00	0.00	0.00	0.00	0.00	0.00
CF	0.00*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00*	0.00	0.00	0.00	0.00
DR	0.03*	0.05*	0.00	0.00	0.00	0.00	0.01*	0.00	0.00	0.00*	0.00*	0.00	0.00	0.00
EF	0.00	0.00	0.00	0.00	0.00	0.01*	0.00	0.00*	0.00*	0.00	0.00	0.00*	0.00*	0.00
BA	0.01*	0.01*	0.00	0.00*	0.00	0.00	0.00	0.00	0.00*	0.00	0.00	0.00	0.00	0.00

Note. GIA = General Intellectual Ability; CHC = Cattell-Horn-Carroll; Gc = Comprehension-Knowledge; Gf = Fluid Reasoning; Ga = Auditory Processing; Gv = Visual-Spatial Thinking; Gsm = Short-Term Memory; Glr = Long-Term Retrieval; Gs = Processing Speed; PA = Phonemic Awareness; WM = Working Memory; CF = Cognitive Fluency; DR = Delayed Recall; EF = Executive Processes; BA = Broad Attention. Coefficients represent the proportion of variance accounted for by variables at their entry point into regression equation. [a]Degrees of freedom reflects controlling for the effects of the GIA. *$p < .05$.

Table A.3 Incremental contribution of CHC cognitive abilities in predicting reading comprehension beyond the general factor across the school age

	6	7	8	9	10	11	12	13	14	15	16	17	18	*M*
GIA	0.46*	0.55*	0.59*	0.57*	0.61*	0.63*	0.60*	0.65*	0.65*	0.66*	0.66*	0.60*	0.67*	0.61
CHC (df = 7)[a]	0.03*	0.07*	0.08*	0.07*	0.09*	0.08*	0.09*	0.07*	0.08*	0.07*	0.08*	0.08*	0.10*	0.08
Gc	0.00	0.04*	0.06*	0.04*	0.07*	0.07*	0.07*	0.05*	0.06*	0.06*	0.07*	0.07*	0.09*	0.06
Gf	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*	0.00*	0.00	0.00	0.01*	0.00*	0.00	0.01*	0.01
Ga	0.00	0.00	0.00	0.00	0.00	0.00	0.00	00.00	0.00	0.00	0.00	0.00	0.00	0.00
Gv	0.00	0.00	0.00	0.00	0.01*	0.00	0.01*	0.00	0.01*	0.00	0.00	0.00	0.00	0.00
Gsm	0.00	0.01*	0.00*	0.00	0.00	0.01*	0.01*	0.00	0.00	0.00*	0.01*	0.02*	0.00*	0.00
Glr	0.02*	0.02*	0.00	0.00	0.00*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Gs	0.00*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00*	0.00	0.00	0.01*	0.00
Clinical (df = 6)[a]	0.05*	0.09*	0.03*	0.02*	0.03*	0.03*	0.01*	0.03*	0.00*	0.02*	0.02*	0.04*	0.02*	0.03
PA	0.00	0.00*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
WM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01*	0.00	0.00	0.00
CF	0.01*	0.00	0.01*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DR	0.04*	0.07*	0.02*	0.01*	0.01*	0.01*	0.00*	0.01*	0.00	0.00*	0.00	0.02*	0.00	0.02
EF	0.00	0.01*	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00*	0.00*	0.00*	0.01*	0.00

BA	0.01*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01*	0.00*	0.01*	0.00

Note. GIA = General Intellectual Ability; CHC = Cattell-Horn-Carroll; Gc = Comprehension-Knowledge; Gf = Fluid Reasoning; Ga = Auditory Processing; Gv = Visual-Spatial Thinking; Gsm = Short-Term Memory; Glr = Long-Term Retrieval; Gs = Processing Speed; PA = Phonemic Awareness; WM = Working Memory; CF = Cognitive Fluency; DR = Delayed Recall; EF = Executive Processes; BA = Broad Attention. Coefficients represent the proportion of variance accounted for by variables at their entry point into regression equation. [a]Degrees of freedom reflects controlling for the effects of the GIA. *$p < .05$.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).