Article

Journal of Psychoeducational Assessment
2017, Vol. 35(8) 731–754

DOI: 10.1177/0734282916659208
journals.sagepub.com/home/jpa
SAGE

Revisiting the Relations Between the WJ-IV Measures of Cattell-Horn-Carroll (CHC) Cognitive Abilities and Reading Achievement During the School-Age Years

Damien C. Cormier[1], Kevin S. McGrew[2,3], Okan Bulut[1], and Allyson Funamoto[1]

Abstract

This study examined associations between broad cognitive abilities (Fluid Reasoning [Gf], Short-Term Working Memory [Gwm], Long-Term Storage and Retrieval [Glr], Processing Speed [Gs], Comprehension-Knowledge [Gc], Visual Processing [Gv], and Auditory Processing [Ga]) and reading achievement (Basic Reading Skills, Reading Rate, Reading Fluency, and Reading Comprehension) in a nationally representative school-age sample. Findings indicate that some cognitive abilities were stronger predictors of reading achievement than previously found (e.g., Gf, Ga, and Gs). Most notably, the Woodcock-Johnson–IV Gf cluster was found to be the strongest and most consistent predictor of reading achievement. A secondary analysis suggests that this effect was likely due to the new Number Series test. The results of the study suggest revisions to previous conceptualizations of the associations between the broad Cattell-Horn-Carroll abilities and areas of reading achievement.

Keywords

cognitive abilities, personality/individual differences, CHC theory, academic achievement, reading, disciplines and subjects

Enhancing the development of cognitive and academic skills of students remains a foundational competency of the practice of school psychology (Ysseldyke et al., 2006). This is accomplished through evidence-based practices (Fiorello & Primerano, 2005). The onus is therefore placed on practitioners to continuously reference empirical evidence to guide their practice. It is, however, the responsibility of researchers to provide answers to questions pertinent to practice (Fiorello & Primerano, 2005). Currently, the existing evidence providing practitioners with knowledge of the associations between cognitive abilities and academic achievement (e.g., Evans, Floyd, McGrew,

[1]University of Alberta, Edmonton, Canada
[2]Institute for Applied Psychometrics, Saint Joseph, MN, USA
[3]University of Minnesota, Minneapolis, MN, USA

Corresponding Author:
Damien C. Cormier, Department of Educational Psychology, University of Alberta, 6-107E Education North, Canada, T6G 2G5.
Email: dcormier@ualberta.ca

& Leforgee, 2002; Floyd, Evans, & McGrew, 2003; Floyd, McGrew, & Evans, 2008) are based on measures that are no longer in use, as they have been replaced by revised and re-normed versions. School psychologists would benefit from knowing the associations between current measures of cognitive abilities and academic achievement.

Measuring Cognitive Abilities

The Cattell-Horn-Carroll (CHC) model of intelligence represents the combination of the most empirically valid psychometric taxonomies of human cognitive abilities (McGrew, 2009; Schneider & McGrew, 2012). CHC theory is the most commonly used model to inform contemporary cognitive test development and interpretation (Keith & Reynolds, 2010; Schneider & McGrew, 2012). The CHC taxonomy is considered the most well-validated and comprehensive description of abilities related to cognitive functioning (Ackerman & Lohman, 2006; McGrew, 2009; Newton & McGrew, 2010). A few cognitive batteries have been developed to explicitly operationalize the CHC model, such as the Woodcock-Johnson Tests of Cognitive Abilities, Fourth Edition (WJ-IV COG; Schrank, McGrew, & Mather, 2014b) and the Kaufman Assessment Battery for Children, Second Edition (KABC-II; Kaufman & Kaufman, 2004). Best practices emphasize that measures of cognitive abilities should not be used in isolation to make diagnostic or educational programming decisions (Kamphaus, Winsor, Rowe, & Kim, 2012). It is not surprising that measures of academic achievement are often discussed as being paired with measures of cognitive abilities within the assessment process (e.g., Flanagan, Ortiz, & Alfonso, 2013).

Cognitive Abilities and Academic Achievement

Several broad and narrow CHC abilities have been empirically associated with various components of reading in children and adolescents. In general, each of the broad CHC abilities has a specific association to a particular curricular area of academic achievement (e.g., reading, mathematics, writing), when controlling statistically for other broad CHC abilities (e.g., Evans et al., 2002; Floyd, Clark, & Shadish, 2008; Floyd et al., 2003). There also appear to be distinct associations between broad CHC abilities and specific curricular subskills (e.g., decoding, comprehension; Evans et al., 2002) even when the effect of general intellectual ability (*g*) is included in the analysis (Benson, 2008; Floyd, Keith, Taub, & McGrew, 2007). For a review of 20 years of research on the relationship between CHC cognitive and achievement abilities, see McGrew and Wendling (2010).

Current Study

Evidence regarding the relations between the constructs represented in a measure and other variables, such as variables that the constructs are expected to predict, is an important contribution to the validity evidence gathered for a given measure (AERA, APA, & NCME, 2015). The relationship between cognitive abilities and academic achievement has been of interest to researchers and practitioners for decades (e.g., Hollingworth & Cobb, 1928; Jensen, 1969; Letteri, 1980; Swanson, 1994). Further, one of the primary uses of measures of cognitive abilities continues to be to make decisions regarding educational programming (Kranzler, Benson, & Floyd, 2016). The WJ-IV (Schrank, McGrew, & Mather, 2014a) was developed to measure cognitive abilities, academic achievement, and oral language. The WJ-IV Technical Manual (McGrew, LaForte, & Schrank, 2014) presents correlations among all tests and correlations among all clusters, but it does not report information regarding the relationships (i.e., correlations) between individual tests and clusters. Further, given that the stated purpose of the WJ-IV emphasizes its utility in the

Table 1. Sample Demographics by Age Group.

		Sex		Race					Mother's level of education		
Age	*n*	Male (%)	Female (%)	W (%)	B (%)	I (%)	A/PI (%)	O/M (%)	<HS (%)	HS (%)	C (%)
6	308	49.7	50.3	81.2	13.3	1.6	2.3	1.6	13.7	31.7	54.6
7	310	50.0	50.0	81.0	12.3	0.3	4.5	1.9	12.6	29.7	57.7
8	336	50.3	49.7	78.0	12.5	0.6	6.5	2.4	13.3	28.9	57.8
9	306	49.0	51.0	77.1	14.4	0.7	3.6	4.2	14.0	28.9	57.1
10	314	50.0	50.0	81.2	11.1	0.6	4.8	2.2	14.7	26.9	58.3
11	329	50.5	49.5	75.7	14.0	0.9	6.4	3.0	10.7	30.6	58.7
12	317	50.2	49.8	79.5	12.3	0.9	5.0	2.2	16.2	28.3	55.6
13	307	46.9	53.1	74.9	15.6	1.0	5.9	2.6	11.8	26.9	61.3
14	299	49.8	50.2	81.3	11.4	0.7	5.4	1.3	13.9	31.4	54.7
15	277	52.0	48.0	80.9	13.0	0.4	3.2	2.5	13.0	30.4	56.5
16	284	50.0	50.0	76.1	16.9	0.0	5.6	1.4	10.6	34.6	54.8
17	254	46.5	53.5	78.7	16.9	1.2	1.6	1.6	8.3	33.6	58.1
18	276	46.7	53.3	70.7	22.8	1.1	3.6	1.8	9.4	38.3	52.3
19	295	47.5	52.5	76.6	17.6	0.3	3.7	1.7	1.6	27.0	71.4

Note. W = White; B = Black; I = Indian; A/PI = Asian or Pacific islander; O/M = Other or Mixed; <HS = less than high school graduate; HS = high school graduate; C = some college or more.

assessment of "important abilities" (McGrew et al., 2014, p. 8) in educational and clinical settings for a broad age range, it is important to examine the potential developmental trends for these associations. Fine grained age-specific developmental information is currently unavailable, as much of the WJ-IV technical manual focuses on reporting correlational data for specific age *groupings* (e.g., ages 6 to 8, 9 to 13, 14 to 17). This study aims to address the following research questions:

> **Research Question 1**: What are the associations between the broad WJ-IV CHC clusters and the WJ-IV reading achievement clusters?
> **Research Question 2**: What are the developmental trajectories of the associations between the broad WJ-IV CHC clusters and the WJ-IV reading achievement clusters?

Method

Sample

The normative sample for the WJ-IV COG (Schrank et al., 2014b) and the Woodcock-Johnson Tests of Academic Achievement, Fourth Edition (WJ-IV ACH; Schrank, McGrew, & Mather, 2014c) were used to examine the relationships between broad CHC abilities and reading achievement.[1] The WJ-IV COG and WJ-IV ACH batteries are co-normed. The complete norming sample included data gathered via a matrix sampling plan from 7,416 people ranging from ages 2 to over 90 (McGrew et al., 2014).

The norming sample is representative of the U.S. population across 46 states and the District of Columbia (McGrew et al., 2014). The sample used for this study includes the school-age subsample, which ranges from 6 to 19 years of age, inclusively. Therefore, the total sample size for this study was 4,126. The total sample was divided into individual age groups for ages 6 to 19, inclusively. The sample included scores for all WJ-IV COG CHC ability clusters and WJ-IV ACH reading clusters. Sample demographics, by age group, are presented in Table 1.

Measures

CHC clusters. The WJ-IV COG is comprised of a standard battery of 10 tests and an extended battery of eight additional tests. CHC cluster scores are calculated from pairs or trios of tests included in the standard or extended batteries. The individual tests and their corresponding CHC broad cluster are as follows: Oral Vocabulary and General Information for Comprehension-Knowledge (Gc); Number Series and Concept Formation for Fluid Reasoning (Gf); Verbal Attention and Numbers Reversed for Short-Term Working Memory (Gwm); Letter-Pattern Matching and Pair Cancellation for Processing Speed (Gs); Phonological Processing and Non-word Repetition for Auditory Processing (Ga); Story Recall and Visual-Auditory Learning for Long-Term Storage and Retrieval (Glr); and Visualization and Picture Recognition for Visual Processing (Gv).

A number of statistical procedures were used to assess and report the reliability of the tests included in the WJ-IV COG. Across the entire norming sample, the median CHC-cluster reliability coefficients for Gc, Gf, Gwm, Gs, Ga, Glr, and Gv are .93, .94, .91, .94, .92, .97, and .86, respectively. The CHC-cluster reliability coefficients for each age level throughout the school years (i.e., ages 6 to 19, inclusively) range from .88 to .98. Extensive evidence of content, predictive, and criterion validity are provided in the WJ-IV COG technical manual (see McGrew et al., 2014). Independent reviews have described the WJ-IV COG as "an excellent measure of psychometric intelligence. The theoretical basis of the test and transparency in test development described in the Technical Manual are exceptional" (p. 389, Reynolds & Niileksela, 2015).

Reading achievement clusters. The WJ-IV ACH is comprised of a standard battery of 11 tests and an extended battery of nine tests. Reading achievement cluster scores are calculated from pairs or trios of tests included in the standard or extended batteries. A three-test "extended" Reading Comprehension cluster score is also available. The individual tests and the corresponding reading achievement clusters relevant to this study are as follows: Letter-Word Identification and Word Attack for Basic Reading Skills; Oral Reading and Sentence Reading Fluency for Reading Fluency; Sentence Reading Fluency and Word Reading Fluency for Reading Rate; and Passage Comprehension and Reading Recall for Reading Comprehension.

The reliability of the WJ-IV ACH clusters was assessed using "Mosier's (1943) unweighted composite" (p. 93, McGrew et al., 2014). When examined by age groups across the school years (i.e., ages 6 to 19, inclusively), the median reliability coefficients for the WJ-IV ACH reading clusters Basic Reading Skills and Reading Comprehension range from r_{cc} = .93 to r_{cc} = .98 and r_{cc} = .91 to r_{cc} = .99, respectively. The CHC-cluster reliability coefficients range from r_{cc} = .96 to r_{cc} = .97 for Reading Fluency and are r_{cc} = .96 for Reading Rate, throughout the school years. The validity evidence for the WJ-IV ACH is also extensive and includes a strong evidence of construct, internal, external, and criterion validity (see McGrew et al., 2014). The WJ-IV ACH battery has received positive independent reviews (Villarreal, 2015).

Data Analysis

Higher order (g) regression models. To demonstrate adequate contributions of the CHC broad clusters in explaining the variation in reading achievement above and beyond the general intelligence factor, a series of regression models were executed. For each of the four reading achievement clusters, there were two nested regression models. The first model was a simple linear regression model which included each of the achievement clusters as the dependent variable and the general intelligence measure (General Intellectual Ability [GIA] cluster) as the sole predictor. The second was a multiple regression model that included each of the achievement clusters as the dependent variable and the *g* cluster (GIA) as well as the seven broad CHC clusters as the predictors.

Because the first model with the *g* cluster was nested within the second model with the *g* cluster and the broad CHC clusters, the R^2 change (i.e., difference in the amount of variance explained) between the two models was used for testing whether the broad CHC clusters are capable of explaining a significant proportion of variance in reading achievement clusters after the variance accounted for by the GIA cluster. This procedure was repeated for each age group in the sample.

Broad CHC abilities regression models. To allow comparisons with prior Woodcock-Johnson Psycho-Educational Battery–Revised (WJ-R; Woodcock & Johnson, 1989) and Woodcock-Johnson, Third edition (WJ-III; Woodcock, McGrew, & Mather, 2001) cognitive–achievement regression research (see Evans et al., 2002; Floyd et al., 2003; McGrew, 1993; McGrew & Hessler, 1995; McGrew & Knopik, 1993), the methods used to present and evaluate the results of the regression models in this study are similar to those used in the prior studies. Therefore, multiple regression was the primary statistical method for examining the associations between cognitive abilities and academic achievement in reading. This method of analysis, in contrast to causal modeling of latent theoretical constructs, is meant to produce practical findings that are meaningful to practitioners.

Four separate regression analyses were conducted using the WJ-IV ACH reading clusters as the dependent variables (i.e., Basic Reading Skills, Reading Fluency, Reading Rate, and Reading Comprehension) and the seven broad CHC cluster scores (i.e., Gc, Gf, Gwm, Gs, Ga, Glr, and Gv) as the predictors. The predictors were entered into the model simultaneously. These four regression models were repeated across all age groups (i.e., ages 6 to 19, inclusively). Age-based standard scores ($M = 100$; $SD = 15$) were used for all analyses. The standardized regression coefficients from each regression model were then interpreted to determine the degree of association between the predictors and the reading clusters. Standardized regression coefficients indicate the proportion of a standard deviation unit of change in the reading clusters as a function of one standard deviation change in the broad CHC cluster scores. Individual data points, by age, are included in the figures, as well as a smoothed curve representing the general developmental trend in the association between individual CHC clusters and broad areas of academic achievement.

Post hoc multiple regression with individual tests. Based on the results from the initial multiple regression analysis, post hoc multiple regression models were completed to better understand some of the novel findings from the first broad CHC cluster level analysis. The relatively high and consistent standardized regression coefficients for the WJ-IV Gf cluster across all four reading achievement clusters were not expected and, in general, were at odds with the extant research literature presenting the cognitive–achievement associations with the WJ-R and WJ-III (see McGrew & Wendling, 2010). It was first hypothesized that Number Series may be accounting for the majority of the variance in the regression models by serving as a proxy for general intelligence (*g*). However, this hypothesis was not supported when individual test *g*-loadings were examined in the WJ-IV Technical Manual (see Table 5-6, McGrew et al., 2014). As indicated in McGrew et al. (2014), it appears that neither Gf test is serving as a proxy for *g* in the multiple regression models, given that the tests Object-Number Sequencing (i.e., a Gwm test), Oral Vocabulary (i.e., a Gc test), and Phonological Processing (i.e., a Ga test) demonstrate *g*-loadings that exceed those of Concept Formation and Number Series for all of the age groups tested (e.g., 6-8, 9-13, 14-19, 20-39, and 40-90+). Therefore, a secondary analysis was conducted to better understand the relationship between individual tests and the results observed at the CHC cluster level.

The post hoc regression models focused on the individual tests, instead of the broad CHC clusters, with each of the WJ-IV ACH clusters (e.g., Basic Reading Skills, Reading Rate, Reading Fluency, and Reading Comprehension) as the dependent variable, again for each of the school-age years (i.e., ages 6 through 19, inclusively). Although multiple regression models could have been used to evaluate all 14 test-level effects, seven-test models were used to avoid the potential influence of multicollinearity that may be introduced due to the various pairs of tests within each of the seven broad CHC clusters. In addition, a seven-test model is more parsimonious, thereby increasing the ease of interpretation of the results. The seven tests used in this study were tests 1 to 7 in the standard battery: Oral Vocabulary (Gc), Numbers Series (Gf), Verbal Attention (Gwm), Letter-Pattern Matching (Gs), Phonological Processing (Ga), Story Recall (Glr), and Visualization (Gv). These tests were selected for inclusion in the GIA cluster score by the WJ-IV authors (from the complete set of WJ-IV cognitive tests) because they were determined to be the best test within each of the broad CHC domain indicators based on multiple criteria specified in the WJ-IV Technical Manual (McGrew et al., 2014).

To examine the unique contribution of the Gf Number Series test above and beyond the GIA tests of the other six CHC domains, a two-step approach was implemented. The first regression model included all seven GIA tests as predictors. This is referred to as the *full model*. The second regression model included the same predictors except for Number Series. This is referred to as the *reduced model*. Because the same six predictors were used in both the full and reduced models, the reduced model was nested within the full model. This nested model structure allows a direct comparison between the models based on the change in R^2 that represents the amount of additional variation explained by the full model (i.e., unique contribution of Number Series) compared with the reduced model. To test the R^2 change between the full model and the reduced model, the following $R^2\Delta F$ test was used:

$$F = \frac{\left(R^2_{full} - R^2_{reduced}\right) / (k_{full} - k_{reduced})}{\left(1 - R^2_{full}\right) / (N - k_{full} - 1)} \quad (1)$$

where R^2_{full} is the R^2 value from the full model, $R^2_{reduced}$ is the R^2 value from the reduced model, k_{full} is the number of predictors in the full model, $k_{reduced}$ is the number of predictors in the reduced model, and N is the sample size. The resulting F-ratio has degrees of freedom of $(k_{full} - k_{reduced})$ and $(N - k_{full} - 1)$. A statistically significant F-ratio from this test suggests that Number Series explains a significant amount of variability in the WJ-IV ACH reading clusters above and beyond the linear combination of the other six CHC individual test scores. Due to the relatively large number of tests to be run (56 tests in total), an alpha level of .001 was set to determine statistical significance of the F-ratio ratio tests.

Results

Higher Order (g) Regression Models

Descriptive statistics for the achievement and cognitive clusters, by age group, can be seen in Table 2. The results from the higher order regression models suggest that the amount of variance explained by the GIA *g*-cluster as the sole predictor ranged from 34% to 58% for ages 6 to 11, from 31% to 51% for ages 12 to 15, and from 29% to 60% for ages 16 to 19. Furthermore, the broad CHC clusters (i.e., Gc, Gf, Gwm, Gs, Ga, Glr, and Gv) as additional predictors explain a significant amount of variance beyond the variation accounted for by the GIA *g*-cluster (see Figures 1 and 2). The R^2 change was not significant only for two age groups (Reading Fluency at age 6 and Basic Reading Skills at age 16). The degree to which the broad CHC clusters account for additional variance in specific areas of reading appears to vary by age and type of reading skill. In general, the broad CHC clusters appear to account for the most additional variance in Reading Rate and Reading Fluency, with the average R^2 change values of .13 and .08, respectively.

Table 2. Means and Standard Deviations for Standard Scores for the Reading Achievement Clusters and the CHC Broad Ability Clusters for the 14 Age Groups.

	Cluster										
Age	BRS	RF	RR	RC	Gc	Gf	Gwm	Gs	Ga	Glr	Gv
6	102.5	101.7	100.8	102.5	101.0	102.5	104.0	103.5	102.3	102.3	103.2
	(13.3)	(14.7)	(14.2)	(14.4)	(14.2)	(14.2)	(12.8)	(14.4)	(13.6)	(15.6)	(15.5)
7	101.6	100.0	100.8	102.4	100.7	101.5	101.4	100.4	99.1	100.0	101.4
	(14.2)	(15.4)	(14.8)	(14.6)	(15.2)	(15.6)	(14.5)	(13.4)	(16)	(15)	(15.9)
8	101.8	101.1	101.2	101.3	100.0	100.9	100.2	100.0	100.1	100.5	101.2
	(15.2)	(15.3)	(15.4)	(15.4)	(15.1)	(14.9)	(16.5)	(15.6)	(17.1)	(15.2)	(16.4)
9	99.2	99.6	100.1	99.7	100.0	99.9	100.8	100.4	99.3	99.8	101.4
	(14.9)	(15.3)	(14.7)	(14.8)	(14.6)	(14.6)	(14.6)	(15.1)	(14.8)	(15.7)	(14.4)
10	99.5	98.6	99.7	100.2	99.8	99.2	101.1	100.4	100.0	100.4	99.6
	(16.4)	(17.1)	(15.7)	(17.6)	(16.4)	(16.1)	(15.8)	(15.9)	(15.9)	(16.4)	(16.2)
11	100.8	100.7	101.2	102.2	101.0	101.0	99.7	100.6	100.3	100.8	100.0
	(14.7)	(15.5)	(14.8)	(16.1)	(14.6)	(14.7)	(14.8)	(15.4)	(15.5)	(15.4)	(15.1)
12	99.8	99.2	99.5	99.1	99.5	98.3	100.6	99.8	99.1	98.3	99.4
	(15.4)	(15.6)	(16.1)	(15.3)	(15.1)	(14.8)	(14.4)	(15.1)	(15.8)	(14.7)	(16.2)
13	99.2	98.8	98.1	97.8	99.7	97.7	100.8	98.8	98.6	100.1	100.2
	(16.3)	(15.7)	(15.6)	(17.6)	(16.4)	(15.7)	(15.5)	(15.4)	(15.6)	(15.3)	(16.5)
14	100.8	101.3	100.4	100.5	101.1	100.9	101.0	98.7	100.7	100.9	101.7
	(16.9)	(16.5)	(16.8)	(17.1)	(16.3)	(16.5)	(16)	(15.6)	(16.2)	(16.6)	(15.6)
15	100.0	99.0	98.2	98.4	99.3	99.5	100.0	99.1	99.9	99.9	100.6
	(16.5)	(15.1)	(16)	(16)	(15.9)	(15.2)	(15.2)	(14.3)	(15.6)	(14.8)	(14.7)
16	100.6	101.0	101.0	100.1	101.2	99.9	101.0	100.4	101.9	102.1	102.2
	(15.8)	(16.5)	(17.3)	(17.6)	(15.9)	(16.5)	(16.2)	(16.5)	(16.5)	(16.1)	(15.7)
17	101.3	101.9	101.1	99.5	100.8	99.8	101.0	99.8	101.1	100.7	100.4
	(16.6)	(15.6)	(15.8)	(16.5)	(16.7)	(15.6)	(16.7)	(15.9)	(15.1)	(15.9)	(15.2)
18	98.3	97.5	96.9	96.7	98.3	96.6	99.4	96.4	98.6	97.8	98.8
	(16.7)	(16.2)	(16.7)	(16.6)	(15.9)	(15.8)	(16.3)	(17.3)	(15.2)	(16.7)	(16.9)
19	101.2	101.7	101.9	101.2	102.2	101.1	101.3	99.3	100.8	99.6	100.6
	(16.9)	(16.5)	(16.9)	(17.9)	(15.5)	(16.7)	(16.8)	(17.3)	(15.4)	(15.9)	(16.2)

Note. Standard deviation values are in parentheses. CHC = Cattell-Horn-Carroll; BRS = Basic Reading Skills; RF = Reading Fluency; RR = Reading Rate; RC = Reading Comprehension; Gc = Comprehension-Knowledge; Gf = Fluid Reasoning; Gwm = Short-Term Working Memory; Gs = Processing Speed; Ga = Auditory Processing; Glr = Long-Term Storage and Retrieval; Gv = Visual Processing.

Broad CHC Abilities Regression Models

The individual standardized regression coefficients from the regression models with the reading cluster scores as the dependent variable and the seven broad CHC clusters as the predictors are summarized by age groups in Figures 3 to 7 (see Appendix for a complete list of all standardized regression coefficients for Broad CHC abilities regression models). A distance weighted least squares (DWLS) smoother with a tension value of .50 was used to produce the smoothed curves. The smoothed curves are considered the best approximation of the population parameters since the age-differentiated point values contain an unknown degree of sampling error (see McGrew & Wrightson, 1997). Only models with standardized regression coefficients consistently at or above .10 are presented, due to values below .10 representing no practical significance (Evans et al., 2002; Floyd et al., 2003; McGrew, 1993; McGrew & Hessler, 1995).

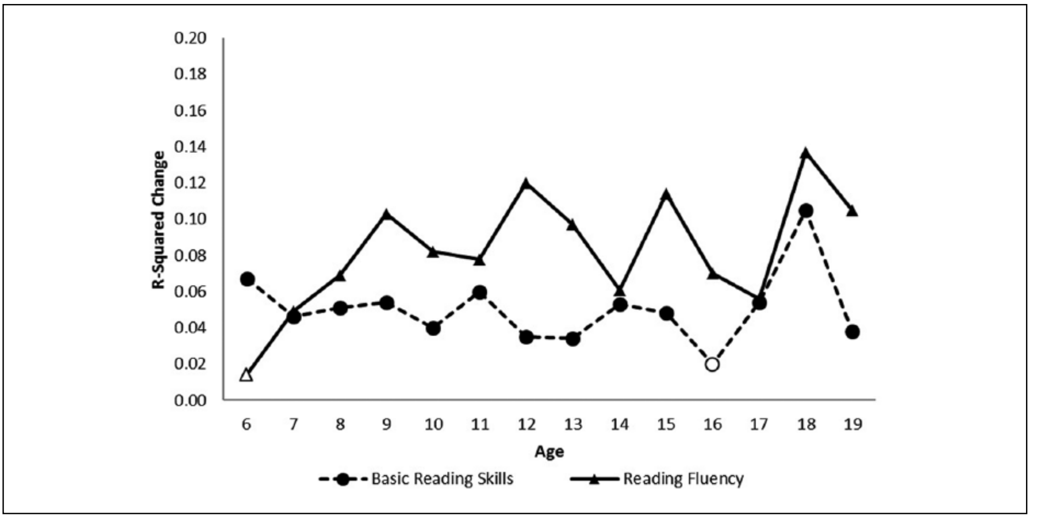

Figure 1. R-squared change by age group for nested broad CHC models compared to the full g-model for Basic Reading Skills and Reading Fluency.
Note. R^2 change was not significant for white-filled points.

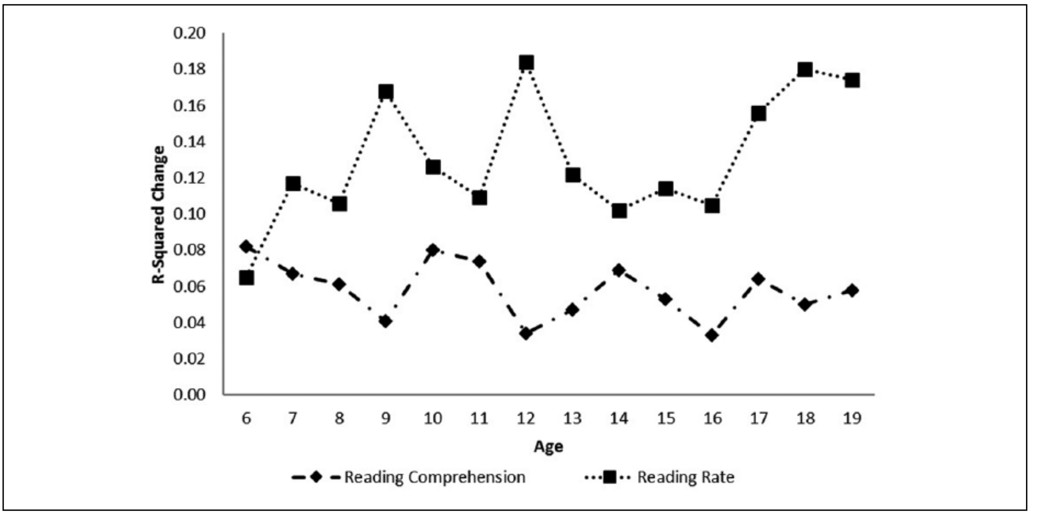

Figure 2. R-squared change by age group for nested broad CHC models compared to the full g-model for Reading Comprehension and Reading Rate.

Each graph includes two horizontally parallel lines corresponding to standardized regression coefficients of .10 and .30. These lines serve as guides for interpreting the significance of the smoothed regression coefficient values and correspond to the rules-of-thumb used in prior WJ-R and WJ-III studies (Evans et al., 2002; Floyd et al., 2003; McGrew, 1993; McGrew & Hessler, 1995; McGrew & Knopik, 1993). As summarized by Evans et al. (2002), "These rules operationally define practical significance to be associated with standardized regression coefficients of .10 or above. Coefficients ranging from .10 to .29 are classified as representing *moderate* effects, whereas those .30 or above are classified as strong effects" (p. 251).

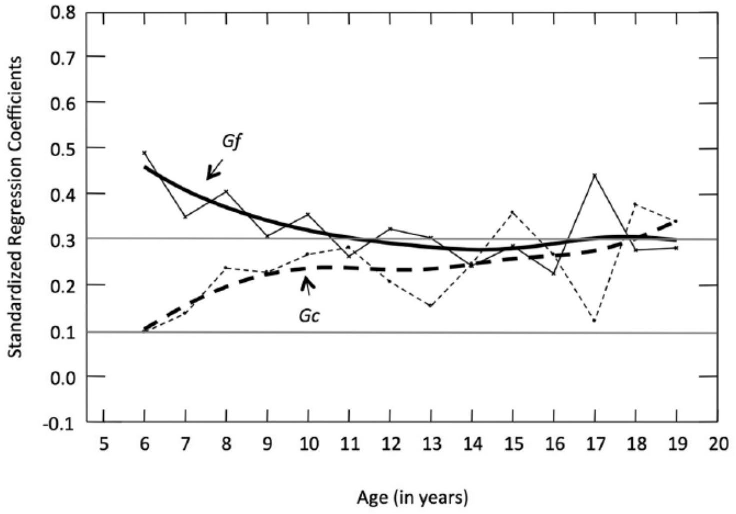

Figure 3. Basic Reading Skills and Gf and Gc clusters.
Note. Gf = Fluid Reasoning; Gc = Comprehension-Knowledge.

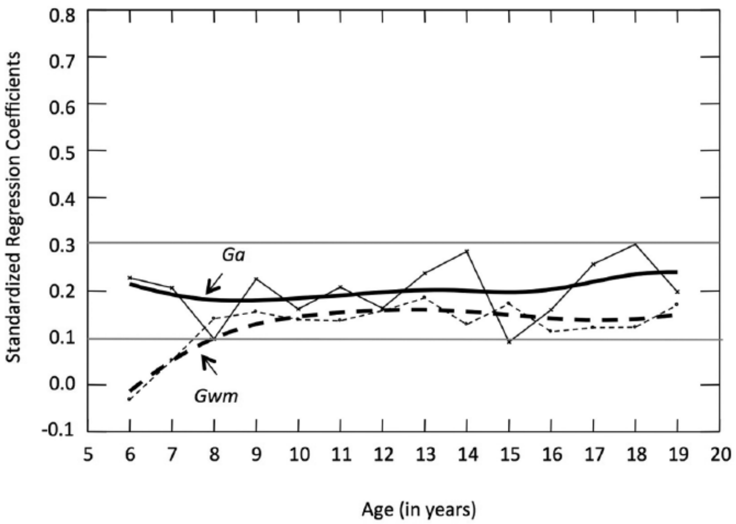

Figure 4. Basic Reading Skills and Ga and Gwm clusters.
Note. Ga = Auditory Processing; Gwm = Working Memory.

Basic Reading Skills. The CHC clusters with the most consistent association with Basic Reading Skills are Gf, Gc, Gwm, and Ga. The predictive values of Glr, Gv, and Gs, across all age groups, were not practically significant (coefficients ≤ .10). It should be noted that Glr, Gv, and Gs do display significant correlations with basic reading skills when considered in isolation (see correlation matrices in Appendix F of the WJ-IV Technical manual). Their classification as not

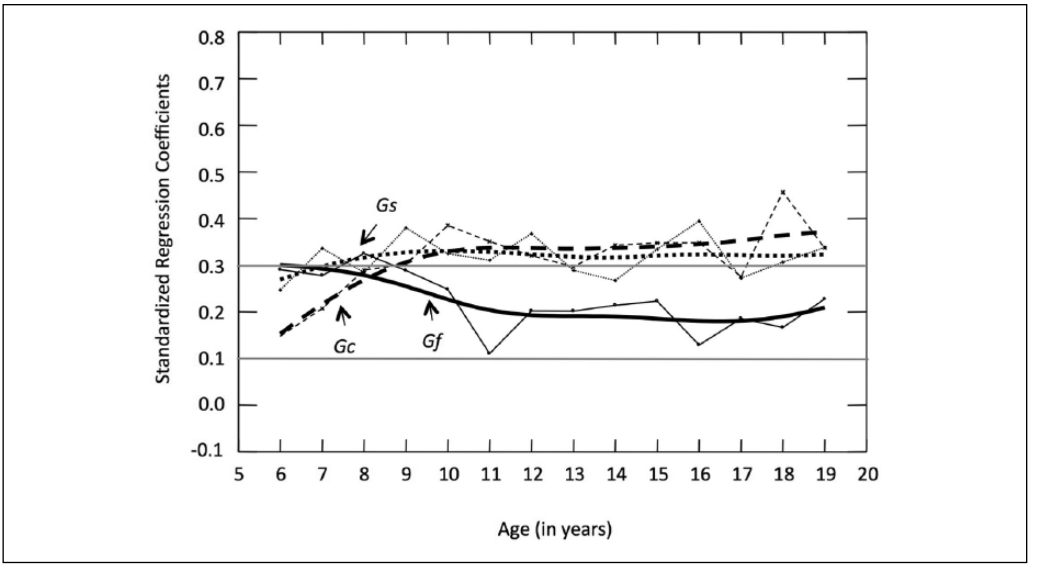

Figure 5. Reading Fluency and Gc, Gf and Gs clusters.
Note. Gs = Processing Speed; Gc = Comprehension-Knowledge; Gf = Fluid Reasoning.

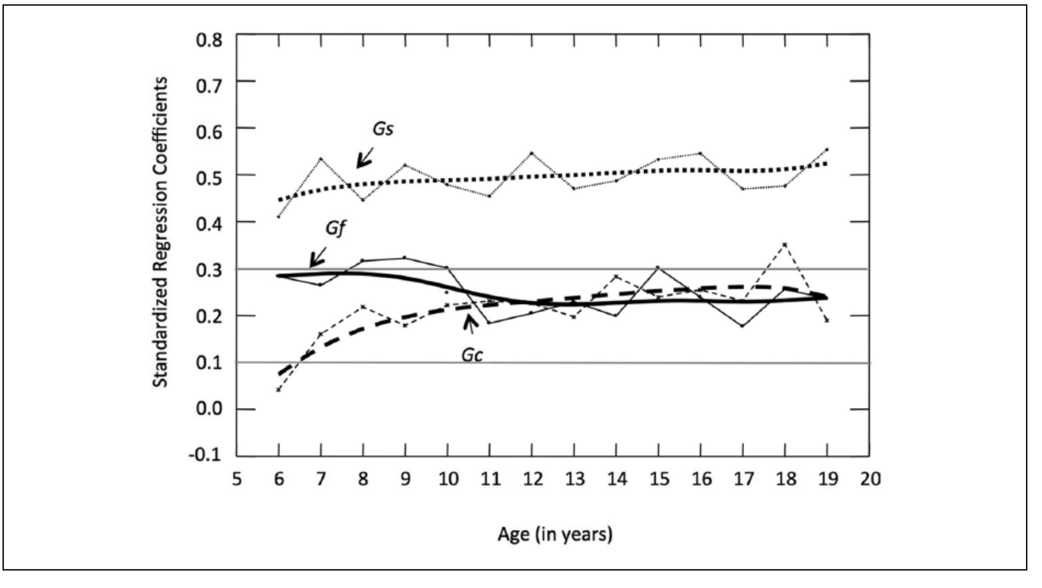

Figure 6. Reading Rate and Gf, Gc and Gs clusters.
Note. Gs = Processing Speed; Gf = Fluid Reasoning; Gc = Comprehension-Knowledge.

practically significant in that the current study are based on models that control statistically for the concurrent predictive power of all other CHC clusters.

Reading Rate and Reading Fluency. The general trend of the associations between broad CHC clusters and Reading Rate and Reading Fluency are moderate for Gf and Gc at most all ages, and strong for Gs. The only notable difference between these reading skills is that Gc appears to

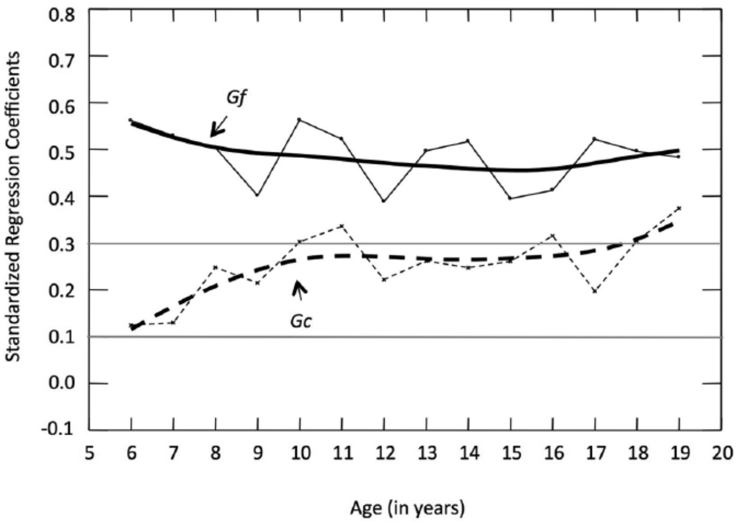

Figure 7. Reading Comprehension and Gf and Gc clusters.
Note. Gf = Fluid Reasoning; Gc = Comprehension-Knowledge.

consistently have a stronger relationship with Reading Fluency (.40 to .50) than it does with Reading Rate (.30). No practically significant association was found between Reading Fluency or Reading Rate and the Ga, Gv, Gsm, and Glr clusters.

Reading Comprehension. As summarized in Figure 7, the only broad CHC clusters demonstrating a consistently significant association with Reading Comprehension over the course of the school years are Gf and Gc. Gf demonstrates a consistently strong association with Reading Comprehension throughout the school years. The association between Gc and Reading Comprehension, however, is in the lower half of the moderate effect window during the early school years (e.g., approximately ages 6 through 7). Gc then steadily increases to the top half of the moderate range from ages 8 to 19.

Post hoc Multiple Regression Analyses

The results of post hoc analyses focusing on the additional contribution of Number Series beyond the other tests (i.e., Oral Vocabulary, Verbal Attention, Letter-Pattern Matching, Phonological Processing, Story Recall, and Visualization) in predicting the reading cluster scores (i.e., Basic Reading Skills, Reading Fluency, Reading Rate, and Reading Comprehension) suggest that Number Series is a very strong predictor of all four areas of reading achievement, particularly Basic Reading Skills and Reading Comprehension. This finding is consistent across all age groups (see Table 3). The removal of Number Series from the analysis results in a median R^2 change of .10 for Basic Reading Skills and Reading Comprehension. This means that, on average, the Number Series test accounts for an additional 10% of the Basic Reading Skills and Reading Comprehension score variance *above and beyond the combined effect of the six other GIA tests*. The effect appears to be smaller for Reading Rate and Reading Fluency, with a median R^2 change of .02 and .03, respectively, across all school-age groups.

Table 3. Comparison of Test-Level Multiple Regression Models Including and Excluding Number Series.

			R^2			F values	
WJ-IV cluster	Age	n	With NS	Without NS	Change	Test for R^2 change	Critical value (α = .001)
Basic Reading Skills	**6**	244	0.59	0.42	**0.17**	97.85	11.09
	7	292	0.56	0.47	**0.09**	58.09	11.05
	8	332	0.61	0.48	**0.13**	108.00	11.02
	9	306	0.57	0.47	**0.10**	69.30	11.04
	10	314	0.61	0.47	**0.14**	109.85	11.03
	11	329	0.52	0.45	**0.07**	46.81	11.02
	12	317	0.53	0.42	**0.11**	72.32	11.03
	13	307	0.51	0.41	**0.10**	61.02	11.04
	14	299	0.53	0.46	**0.07**	43.34	11.04
	15	277	0.53	0.48	**0.05**	28.62	11.06
	16	284	0.52	0.47	**0.05**	28.75	11.06
	17	254	0.62	0.47	**0.15**	97.11	11.08
	18	276	0.56	0.50	**0.06**	36.55	11.06
	19	295	0.62	0.55	**0.07**	52.87	11.05
Reading Rate	**6**	244	0.41	0.38	**0.03**	12.00	11.09
	7	292	0.49	0.48	0.01	5.57	11.05
	8	332	0.57	0.52	**0.05**	37.67	11.02
	9	306	0.6	0.54	**0.06**	44.70	11.04
	10	314	0.54	0.51	**0.03**	19.96	11.03
	11	329	0.47	0.45	**0.02**	12.11	11.02
	12	317	0.45	0.43	**0.02**	11.24	11.03
	13	307	0.56	0.54	**0.02**	13.59	11.04
	14	299	0.48	0.46	**0.02**	11.19	11.04
	15	277	0.53	0.51	**0.02**	11.45	11.06
	16	284	0.55	0.53	**0.02**	12.27	11.06
	17	254	0.4	0.39	0.01	4.10	11.08
	18	276	0.48	0.46	0.02	10.31	11.06
	19	295	0.57	0.56	0.01	6.67	11.05
Reading Fluency	**6**	244	0.41	0.36	**0.05**	20.00	11.09
	7	292	0.49	0.46	**0.03**	16.71	11.05
	8	332	0.55	0.5	**0.05**	36.00	11.02
	9	306	0.55	0.49	**0.06**	39.73	11.04
	10	314	0.5	0.47	**0.03**	18.36	11.03
	11	329	0.42	0.41	0.01	5.53	11.02
	12	317	0.43	0.4	**0.03**	16.26	11.03
	13	307	0.49	0.46	**0.03**	17.59	11.04
	14	299	0.45	0.43	0.02	10.58	11.04
	15	277	0.48	0.47	0.01	5.17	11.06
	16	284	0.54	0.53	0.01	6.00	11.06
	17	254	0.43	0.41	0.02	8.63	11.08
	18	276	0.49	0.48	0.01	5.25	11.06
	19	295	0.53	0.52	0.01	6.11	11.05
Reading Comprehension	**6**	244	0.58	0.45	**0.13**	73.05	11.09
	7	292	0.59	0.48	**0.11**	76.20	11.05
	8	332	0.62	0.52	**0.10**	85.26	11.02
	9	306	0.54	0.44	**0.10**	64.78	11.04

(continued)

Table 3. (continued)

WJ-IV cluster	Age	*n*	R^2 With NS	Without NS	Change	*F* values Test for R^2 change	Critical value (α = .001)
	10	314	0.59	0.46	**0.13**	97.02	11.03
	11	329	0.5	0.39	**0.11**	70.62	11.02
	12	317	0.49	0.42	**0.07**	42.41	11.03
	13	307	0.57	0.45	**0.12**	83.44	11.04
	14	299	0.54	0.44	**0.10**	63.26	11.04
	15	277	0.51	0.47	**0.04**	21.96	11.06
	16	284	0.6	0.56	**0.04**	27.60	11.06
	17	254	0.57	0.42	**0.15**	85.81	11.08
	18	276	0.58	0.51	**0.07**	44.67	11.06
	19	295	0.65	0.58	**0.07**	57.40	11.05

Note. Boldface font indicates that the test of R^2 change yielded a result that exceeded the critical value, indicating a statistically significant change in the R^2 between the two models. WJ-IV = Woodcock-Johnson, Fourth Edition. NS = number series.

Discussion

The current results are, in many respects, notably different from those reported in the previous WJ-R and WJ-III studies. The evolving nature of CHC theory and cognitive test batteries, as well as advancements in our understanding of human cognitive abilities from other fields, such as cognitive neuroscience, provide the impetus for continuously revisiting questions related to the associations between measures of CHC abilities and areas of academic achievement. Relying on previous WJ-R and WJ-III-based research alone could result in erroneous assumptions about the associations between certain WJ-IV CHC measures and areas of reading achievement. Thus, these findings have significant practical implications for the appropriate use of the WJ-IV tests and clusters.

Gf and Number Series

The most intriguing and unexpected result in the current study is the finding that the WJ-IV Gf cluster is the strongest and most consistent predictor of all forms of reading achievement across all ages. The Gf cluster is a strong predictor of basic reading skills and reading comprehension and a moderate predictor of the reading rate and fluency. This finding is at odds with the previous WJ-III Gf cluster multiple regression research (Evans et al., 2002). The results of the post hoc analyses suggest that the Number Series test is contributing significantly to the strong predictive power of the Gf cluster. Number Series tasks have been referenced in the literature for decades (e.g., Carroll, 1993; LeFevre & Bisanz, 1986; Quereshi & Smith, 1998) and have demonstrated the potential to predict work performance (Bertling, 2012). The current findings, however, appear to be the first to directly demonstrate the strong relative predictive power of number series tasks for reading achievement. The unexpected finding that the WJ-IV Number Series test had a much higher correlation with reading achievement than it did in the WJ-III is perplexing and warrants further investigation.[2] A number of possible hypotheses are offered.

First, it is possible that changes have occurred in the population's exposure to tasks similar to number series items. The WJ-III and WJ-IV were published in 2001 and 2014, respectively,[3] which is over a decade between editions. As noted by the National Council of Teachers of Mathematics (NCTM; 2010), "the past two decades have seen an era of unprecedented mathematics curriculum development across grades K-12" (p. x). Also, mathematical reasoning games for entertainment (e.g., Sudoku) have become more prevalent and accessible via web page tutorials (http://www.funwithpuzzles.com/2015/02/easy-mathematical-brain-teasers-with.html), on-line videos (https://www.youtube.com/watch?v=utmf0pSOgk0), and mobile phone or tablet apps (https://itunes.apple.com/us/app/find-next-in-number-series/id1067642974?mt=8). However, it is not possible to evaluate, in this article, whether any systematic math curriculum changes or the impact of increased game-like instruction on quantitative reasoning may have produced differences in the school-age population that would change the underlying constructs being measured by number series items.

Second, the WJ-IV authors indicate that the concept of cognitive complexity was used to increase the cognitive processing demands on certain tests. Task analysis of the Number Series test indicates that many of the items require the successful completion of numerous procedural steps: relation detection, detection of periodicity, completion of pattern description, and extrapolation (Holzman, Pellegrino, & Glaser, 1983). The cognitive complexity associated with Number Series could be explained by the task's demands placed on memory load and its relational complexity (Bertling, 2012). For example, the ability of an examinee to evaluate serial and relational hypotheses when attempting each item may require the executive function of "placekeeping" ability (Hambrick & Altmann, 2015)—An ability that increases the load on working memory capacity. Working memory load is the amount of information that needs to be held in memory, to be used and possibly manipulated within seconds or minutes. Relational complexity has been defined by "the number of relationships between elements that define the right solution" (Bertling, 2012, p. 96). Bertling (2012) provided a description of the interplay between these two aspects of cognitive complexity during a number series task:

> The test-taker has to hold active in mind several possible rule combinations while storing intermediate result(s) in working memory as well. This does not only make the solution of such a number series very hard; it also allows for different strategic approaches to reduce complexity. (p. 95)

Support for this hypotheses would require demonstrating that the WJ-IV Number Series test item content changed to elicit more complex cognitive processing. Surface-level comparisons of the WJ-III and WJ-IV technical manuals suggest no apparent major changes. The tests had 47 and 42 items respectively, and a range of similar *W*-scores (approximately 111 and 108), as reflected by the mean reported *W*-score from ages 5 to the asymptote of the growth curve in the respective technical manuals (McGrew et al., 2014; McGrew, Schrank, & Woodcock, 2007). However, an inspection of the items in each test reveals that many of the WJ-III Number Series items were replaced with new items in the WJ-IV. Of the 42 WJ-IV Number Series items, 18 (42.9%) were not in the WJ-III.

The number series literature has indicated that a number of variables can change the item difficulty and cognitive processing demands of items. For example, Holzman et al.'s (1983) classic number series research indicated that number series items can vary in difficulty or complexity based on a number of empirically classifiable characteristics of the items: "the influence of working-memory placekeeper demands, period length, pattern description length, relational complexity, category of arithmetic operation, string length, and directional conflicts in the relations governing the series" (p. 609). Bertling's (2012) review of the number series item generation literature indicates that item complexity can be varied as a function of such characteristics as (a) the type of task (identify rule-discrepant element; continue series; fill out missing element), (b) rule combination required (sequential; all in one step; hierarchical overlap), (c) the number of rules in a series, (d) arithmetic rules (basic vs. complex operations), (e) the magnitude of the numbers, and (f) length of the series (number of elements). Informal analysis of the items not shared between the WJ-III and WJ-IV Number Series tests suggests that a shift in the content between the two tests may be a plausible hypothesis that warrants further study. For example, in terms of number of elements (i.e., the number of integers presented plus missing element blank spaces), 100% of the WJ-IV-unique set had four or five element items (four elements = 16 items; 88.9%; five elements = 2; 11.1%) whereas 79.2% of the WJ-III-unique items had such four or five element items (four elements = 15; 62.5%; five elements = 4; 16.7%)—a difference of 20.8%. In contrast, the WJ-III-unique set had 20.8% items with six or more (6, 7, 9) elements. The change in a large portion of items may have produced changes in the level of cognitive processing required, or the cognitive construct(s) measured from the WJ-III to the WJ-IV—changes that increased the tests correlation with reading achievement. Is the WJ-IV Number Series now a more mixed measure of reasoning (Gf) and acquired knowledge (Gq), or did the item changes increase the degree of relational cognitive complexity required? This question moves beyond the scope of this article, and may be a promising direction for future research.

Finally, another hypothesis is that the change in the Number Series test association with school achievement might reflect an unknown methodological artifact. To reduce participant response burden, the WJ-IV norming data were gathered via a complex matrix sampling plan—subsets of norm subjects were administered one of three different blocks of tests and portions of a common core linking block. This required the use of multiple data imputation plausible values methodology to produce complete records for the construction of norms and technical analysis (McGrew et al., 2014). Given the complexity of this design and data imputation procedures, as well as no information provided in the technical manual on the test composition of the four different norming blocks of tests, it is not possible to determine if this design and the amount of imputed data may have introduced some form of methodological artifact into the WJ-IV Number Series data. The WJ-III norm sample was comprised of over 8,700 subjects, of which over 7,000 (approximately 80%) were administered the Number Series test.[4] The final reported sample size for the WJ-IV norm sample is 7,416. In the WJ-IV technical manual, summary statistics for Numbers Series are based on over 6,700 subjects. Depending on whether the WJ-IV Number Series test was in the core linking block (n = 3,400 to 3,800 subject) or one of the other three norming blocks (n = 1,500 to 2,200 subjects), the portion of imputed norm data for Number Series could range from approximately 59% to 77%. This amount of plausible value imputation, combined with the authors' reporting of indications of violation of the assumption of multivariate normality and issues with some multicollinearity in this large collection of tests (many tests that are highly related—e.g., all the reading and writing tests), it is plausible that the increased association of the WJ-IV Number Series test with school achievement may be a methodologically based artifact caused by some unknown degree of bias in the imputation of the WJ-IV Number Series data (e.g., a "norming block effect"). This hypothesis can only be examined by accessing the complete WJ-IV norming data and details regarding the matrix sampling design.

Additional Moderate to Strong Predictors of Reading Achievement

The results suggest that the WJ-IV Gc cluster generally has moderate relationship to reading throughout the school-age years. However, the WJ-III Gc cluster demonstrated moderate effects at the youngest ages (6 to approximately 8 years of age) and a monotonically increasing strong effect from approximately ages 9 through 19 for the WJ-III basic reading skills and reading comprehension clusters (Evans et al., 2002). It is possible that these more moderate findings for the WJ-IV Gc cluster, when compared with the prior WJ-R and WJ-III Gc findings, are due to the increased strength of association for the new WJ-IV Gf cluster. That is, the WJ-IV Gf cluster accounts for more of the reading achievement variance, leaving less reading achievement variance to be accounted for by Gc and the other WJ-IV broad CHC clusters.

Similarly, Gs was previously reported to be a low to moderate predictor of basic reading skills and reading comprehension from age 6 to approximately age 9 (Evans et al., 2002). In the current investigation, Gs was not a significant predictor of basic reading skills or reading comprehension. The Reading Rate and Reading Fluency clusters were not available in the WJ-III battery and therefore were not evaluated in previous studies. The current results suggest that Gs is a strong predictor of Reading Rate and Reading Fluency across all of the school-age years. These findings, however, are not surprising considering the speeded nature of the tasks involved in the Gs cluster tests and those used to measure Reading Rate and Reading Fluency.

The Ga cluster demonstrated a consistent moderate association with basic reading skills at all ages (6 through 19 years of age), a finding at variance with the WJ-III Ga cluster research. Evans et al. (2002) previously reported that the WJ-III Ga cluster was only a moderate predictor of basic reading skills and reading comprehension during the early school years (e.g., ages 6-8). In Evans et al., the smoothed curve for the WJ-III Ga cluster and reading comprehension was in the lower portion of the moderate effect size windows for approximately ages 6 through 8. Given the relative weakness of these prior WJ-III findings, plus no practical significance at any other ages, we adopt the conservative interpretation that these previous limited and weak findings most likely reflect sampling error. The 100% change of the WJ-IV Ga cluster (see Table 4) appears to have increased the association and importance of the WJ-IV Ga cluster for understanding basic reading skills and reading comprehension across all school years. As outlined in the WJ-IV technical manual (McGrew et al., 2014), the new WJ-IV Ga cluster is comprised of measures of much more cognitively complex auditory processes (PC-phonetic coding; LA-speed of lexical access [sound-based lexical access]; UM-memory for sound patterns) than those measured by the WJ-III Ga cluster (PC-phonetic coding; US/U9-sound discrimination and resistance to auditory stimulus distortion). The increase in the cognitive complexity of the auditory processes measured by the WJ-IV Ga cluster may be contributing to these tests being a better measure of reading-related skills across the school-age population.

Weaker Predictors of Reading Achievement

The WJ-III memory clusters were previously reported as demonstrating consistently moderate associations with basic reading skills for all school ages (Gsm) or moderate associations for ages 6 through approximately 9 to 10 years (Glr; Evans et al., 2002). More importantly, the WJ-III also included a two-test working memory cluster that is more comparable with the WJ-IV Gwm cluster (than the WJ-III Gsm cluster). The WJ-III working memory (Numbers Reversed; Auditory Working Memory) and WJ-IV Gwm clusters (Numbers Reversed; Verbal Attention) are both comprised of two tests of aspects of working memory. Yet, the WJ-III Working Memory cluster was consistently more related to both basic reading skills and reading comprehension at most all school ages (Evans et al., 2002) whereas the WJ-IV Gwm cluster was similarly moderate in association, but only for basic reading skills. This finding reflects either the change in the composition of the WJ-III to WJ-IV Reading Comprehension clusters noted previously, or the possibility that other WJ-IV clusters, notably Gf and Gc, are accounting for more reading achievement variance—leaving less variance to be explained by the WJ-IV Gwm cluster.

The WJ-IV Glr cluster demonstrated no significant association with any of the four reading achievement clusters, a finding that is inconsistent with the Glr cluster relationship seen with reading comprehension for the WJ-III (Evans et al., 2002). The reduction in Glr association with reading achievement may reflect the different composition of the WJ-III Glr (Visual-Auditory Learning; Retrieval Fluency) and WJ-IV Glr clusters (Story Recall; Visual-Auditory Learning) or the previously mentioned hypothesis that other revised WJ-IV CHC clusters are accounting for more reading achievement variance. The results herein, however, suggest that the WJ-IV Glr cluster has little value in predicting components of reading achievement, when controlling statistically for other all other broad CHC abilities.

Table 4. The Changing Composition of the WJ-R/WJ-III *g* and CHC Broad Clusters.

	g-cluster			CHC factors		
	WJ-R	WJ-III	WJ-IV			
CHC broad factor and individual tests	BCA-Std	GIA-Std	GIA	WJ-R	WJ-III	WJ-IV
Glr						
Memory for names	X			X		
Visual-auditory learning		X		X	X	X
Retrieval fluency					X	
Story recall			X			X
Gsm/Gwm						
Memory for sentences	X			X		
Memory for words				X	X	
Numbers reversed		X			X	X
Object-number sequencing						
Verbal attention			X			X
Gs						
Visual matching (number-pattern matching)	X	X		X	X	
Cross out				X		
Decision speed					X	
Pair cancellation						X
Letter-pattern matching			X			X
Ga						
Incomplete words	X			X		
Sound blending		X		X	X	
Auditory attention					X	
Phonological processing			X			X
Nonword repetition						X
Gv						
Visual closure	X			X		
Picture recognition				X	X	X
Spatial relations		X	•		X	
Block rotation (DS)			•			
Visualization (SR + BR)			X			X
Gf						
Analysis-synthesis	X			X	X	
Concept formation		X		X	X	X
Number series (DS)			X			X
Gc						
Picture vocabulary	X	•		X		
Oral vocabulary		•	X	X		X
Verbal analogies		•				
Verbal comprehension (PV + OV + VA)		X			X	
General information					X	X

Note. WJ-R = Woodcock-Johnson Psycho-Educational Battery–Revised; WJ-III = Woodcock-Johnson, Third edition; CHC = Cattell-Horn-Carroll; WJ-IV = Woodcock-Johnson Tests of Cognitive Abilities, Fourth Edition; BCA = Broad Cognitive Ability; GIA = General Intellectual Ability cluster; Glr = Long-Term Retrieval; X = test; Gwm = Working Memory; Gs = Processing Speed; Ga = Auditory Processing; Gv = Visual Processing; • = subtest; Gf = Fluid Reasoning; Gc = Comprehension-Knowledge; DS = Diagnostic Supplement; SR = Spatial Relations; BR = Block Rotation; PV = Picture Vocabulary; OV = Oral Vocabulary; and VA = Verbal Analogies. An X below a series of • represents a test that is the combination of these subtests.

Similar to all prior WJ-R and WJ-III cognitive–achievement relations regression studies, the WJ-IV Gv cluster failed to demonstrate any statistical or practical association with any reading achievement cluster at any age. This lack of significance continues to perpetuate the “Gv Mystery” (McGrew & Wendling, 2010, p. 665). Almost all WJ-R, WJ-III, or other CHC designed research has failed to demonstrate significant associations between measures of Gv and reading achievement. However, significant non-CHC designed research has reported more positive findings for Gv and reading and math achievement. As suggested by McGrew and Wendling (2010),

> lack of significance does not mean that Gv abilities are not involved in reading and math. Obviously, individuals use their eyes when reading and when processing diagrams during reading and math. Gv measures, as currently designed in intelligence batteries, simply may have no achievement variance to account for because the more powerful predictors (e.g., Gc, Gsm, Ga) account for the lion’s share of the reliable variance in the achievement variables. (p. 666)

Generalizability of Findings

Many contemporary cognitive batteries have been based on the CHC theory of intelligence. However, the majority of the research conducted to date which has examined the associations between CHC abilities and areas of academic achievement has been completed with the WJ-R and WJ-III COG and WJ-R and WJ-III ACH batteries (McGrew & Wendling, 2010). Therefore, a healthy degree of caution is advised when attempting to generalize the current results across other WJ batteries and to other tests of cognitive abilities. Given the extent to which the relationship of the broad abilities with reading have changed from the WJ-III to the WJ-IV, caution should be applied in generalizing these findings to other tests without clear and convincing evidence that the non-WJ-IV tests operationalize the broad and narrow abilities in the same way as operationalized in the WJ-IV. In summary, the question regarding the strength of the associations between other measures of cognitive abilities and reading achievement is an empirical one that needs to be answered in future research.

Limitations

This study deliberately focused on examining the linear relations between the seven manifest WJ-IV broad CHC and four reading achievement clusters. The extant CHC cognitive–achievement research literature consists of similar linear regression studies with manifest test battery composites, as well as studies that use structural equation modeling (SEM) to examine the direct and indirect effects of the latent general intelligence (*g*) factor concurrently with latent CHC broad and narrow factors (McGrew & Wendling, 2010). The results of the current study need to be integrated with recent WJ-IV SEM *g* and specific abilities research (Niileksia, Reynolds, Keith, & McGrew, 2016) to better understand the relations between the WJ-IV manifest cognitive and achievement measures and the latent cognitive and achievement factors they represent. Also, to date, we are unaware of any CHC studies that have attempted to explore cognitive–achievement relations with nonlinear models. For particular age groups it is possible that a nonlinear relationship might be anticipated between the WJ-IV CHC cluster scores and the WJ-IV ACH reading scores. Future research studies should investigate the extent to which the CHC cluster scores (WJ-IV and other intelligence batteries) may demonstrate a nonlinear relationship with students’ reading achievement. Finally, research is needed to explore the relations between the WJ-IV cognitive measures and reading achievement as operationalized by other tests (e.g., Wechsler Individual Achievement Test, Third Edition; Kaufman Test of Educational Achievement, Third Edition).

Appendix

Standardized Regression Coefficients for Broad CHC Abilities Regression Models

		Achievement cluster			
		Basic reading skills	Reading rate	Reading fluency	Reading comprehension
Age group	predictor	b^*	b^*	b^*	b^*
6	Gc	0.10	0.04	0.14	0.12
	Gf	0.52	0.28	0.28	0.56
	Gwm	−0.03	0.04	0.05	−0.09
	Gs	0.03	0.42	0.24	0.04
	Ga	0.23	−0.05	0.06	0.17
	Glr	0.02	0.03	0.05	0.02
	Gv	0.02	0.03	0.04	0.08
	R^2	0.52	0.38	0.38	0.53
7	Gc	0.15	0.16	0.2	0.14
	Gf	0.38	0.28	0.28	0.56
	Gwm	0.05	−0.03	0.05	0.05
	Gs	0.02	0.48	0.29	−0.02
	Ga	0.23	0.04	0.12	0.09
	Glr	0	−0.05	−0.01	0.01
	Gv	0.03	0.03	0.02	0.02
	R^2	0.50	0.53	0.51	0.55
8	Gc	0.24	0.21	0.29	0.24
	Gf	0.40	0.30	0.32	0.49
	Gwm	0.15	0.01	0.09	0.07
	Gs	−0.01	0.45	0.29	0.03
	Ga	0.11	0	0.04	0.04
	Glr	−0.14	−0.11	−0.1	−0.12
	Gv	0.14	0.08	0.06	0.19
	R^2	0.54	0.57	0.57	0.59
9	Gc	0.22	0.18	0.29	0.21
	Gf	0.30	0.32	0.28	0.4
	Gwm	0.15	0.01	0.06	0.1
	Gs	0.10	0.53	0.38	0.07
	Ga	0.22	0.02	0.07	0.12
	Glr	−0.14	−0.05	−0.03	−0.05
	Gv	0.07	−0.01	0	0.07
	R^2	0.51	0.60	0.56	0.48
10	Gc	0.27	0.23	0.37	0.28
	Gf	0.35	0.31	0.23	0.51
	Gwm	0.13	−0.05	0.05	−0.01
	Gs	0.05	0.48	0.3	−0.02
	Ga	0.16	−0.01	−0.01	0.11
	Glr	−0.11	−0.01	0.04	−0.08
	Gv	0	−0.03	−0.05	0.03
	R^2	0.51	0.58	0.55	0.55

(continued)

Appendix (continued)

		Achievement cluster			
		Basic reading skills	Reading rate	Reading fluency	Reading comprehension
Age group	predictor	b^*	b^*	b^*	b^*
11	Gc	0.28	0.23	0.33	0.3
	Gf	0.26	0.18	0.1	0.48
	Gwm	0.14	0.01	0.1	0.09
	Gs	0.01	0.47	0.31	−0.05
	Ga	0.22	0.12	0.16	0
	Glr	−0.11	0.01	0	−0.03
	Gv	0.05	−0.11	−0.09	0.03
	R^2	0.45	0.49	0.45	0.47
12	Gc	0.20	0.21	0.31	0.22
	Gf	0.31	0.19	0.19	0.37
	Gwm	0.15	−0.05	0.04	0
	Gs	0.10	0.51	0.36	0.07
	Ga	0.17	−0.04	0.03	0.11
	Glr	−0.17	0.12	0.05	0.03
	Gv	0.10	−0.08	−0.04	0.11
	R^2	0.46	0.49	0.5	0.46
13	Gc	0.16	0.21	0.31	0.24
	Gf	0.29	0.23	0.2	0.44
	Gwm	0.18	0.05	0.05	0.08
	Gs	0.09	0.46	0.28	0.03
	Ga	0.23	0.08	0.15	0.07
	Glr	−0.12	0.03	0.05	0.03
	Gv	−0.01	−0.13	−0.11	−0.02
	R^2	0.45	0.57	0.53	0.51
14	Gc	0.24	0.27	0.34	0.23
	Gf	0.24	0.19	0.21	0.5
	Gwm	0.12	0.05	0.09	0.04
	Gs	0.03	0.45	0.25	−0.02
	Ga	0.27	−0.04	0.02	0.1
	Glr	−0.17	0	0	−0.06
	Gv	0.13	−0.04	−0.01	0.09
	R^2	0.47	0.51	0.49	0.53
15	Gc	0.35	0.24	0.37	0.26
	Gf	0.26	0.29	0.22	0.37
	Gwm	0.16	−0.02	0.06	0.09
	Gs	0.12	0.48	0.32	0.16
	Ga	0.09	−0.03	0	−0.13
	Glr	−0.11	−0.05	0.03	0.13
	Gv	0.02	−0.04	−0.06	0.07
	R^2	0.51	0.53	0.54	0.52
16	Gc	0.27	0.23	0.34	0.28
	Gf	0.24	0.23	0.13	0.39
	Gwm	0.12	−0.02	0.04	0.02

(continued)

Appendix (continued)

		Achievement cluster			
		Basic reading skills	Reading rate	Reading fluency	Reading comprehension
Age group	predictor	b*	b*	b*	b*
	Gs	0.11	0.52	0.39	0.07
	Ga	0.17	−0.07	0.01	0.05
	Glr	−0.06	0.01	0.05	0.06
	Gv	0.02	−0.06	−0.03	0.05
	R^2	0.48	0.54	0.55	0.56
17	Gc	0.12	0.24	0.29	0.2
	Gf	0.41	0.17	0.19	0.49
	Gwm	0.12	−0.03	0.05	0.08
	Gs	−0.02	0.47	0.28	−0.06
	Ga	0.23	−0.01	0.07	0.08
	Glr	−0.10	−0.04	0.02	−0.01
	Gv	0.11	0	−0.01	0.03
	R^2	0.52	0.44	0.44	0.49
18	Gc	0.36	0.34	0.45	0.29
	Gf	0.26	0.24	0.16	0.47
	Gwm	0.12	−0.13	−0.01	0.08
	Gs	0.03	0.49	0.33	0.01
	Ga	0.27	0.02	0.08	0
	Glr	−0.25	−0.09	−0.07	−0.01
	Gv	0.03	−0.1	−0.05	0.04
	R^2	0.55	0.52	0.55	0.56
19	Gc	0.31	0.17	0.32	0.32
	Gf	0.28	0.23	0.23	0.45
	Gwm	0.17	−0.06	0.06	0.14
	Gs	0	0.57	0.35	−0.03
	Ga	0.18	0.04	0.05	−0.06
	Glr	−0.07	−0.02	−0.02	−0.01
	Gv	−0.01	−0.04	−0.07	0.13
	R^2	0.53	0.58	0.56	0.62

Note. Gc = Comprehension-Knowledge; Gf = Fluid Reasoning; Gwm = Working Memory; Gs = Processing Speed; Ga = Auditory Processing; Glr = Long-Term Retrieval; Gv = Visual Processing.
*b** represents standardized regression coefficients.

Acknowledgments

The authors thank Riverside Publishing Company/Houghton Mifflin Harcourt for providing access to the WJ-IV norm data.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Dr. Kevin McGrew is a coauthor of the Woodcock-Johnson, Fourth Edition (WJ-IV) battery and discloses that he has a financial interest in the WJ-IV.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Standardization data from the Woodcock-Johnson™ IV (WJ-IV™). Copyright © 2014 by The Riverside Publishing Company. All rights reserved. Used with permission of the publisher.
2. We wish to thank an anonymous reviewer who provided feedback in the form of comparisons of correlations between respective sets of Woodcock-Johnson, Third Edition (WJ-III) cognitive and oral language tests and school achievement and the similar WJ-IV cognitive and oral language tests correlations with achievement. These observations required the authors to engage in a closer examination of possible content differences between the WJ-III and WJ-IV Number Series tests.
3. The WJ-III norms were refreshed in 2007 with a normative update, but the sample was comprised of the same subjects as the original published in 2001.
4. These WJ-III numbers are calculated from the table of summary statistics in the WJ-III Normative Update Technical Manual (McGrew et al., 2007).

References

Ackerman, P. L., & Lohman, D. F. (2006). Individual differences in cognitive functions. In P. A. Alexander, & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 139-161). Mahwah, NJ: Lawrence Erlbaum.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2015). Standards for educational and psychological testing (2nd ed.). Washington, DC: American Psychological Association.

Benson, N. (2008). Cattell–Horn–Carroll cognitive abilities and reading achievement. *Journal of Psychoeducational Assessment*, *26*, 27-41. doi:10.1177/0734282907301424

Bertling, J. P. (2012). *Measuring reasoning ability: Applications of rule-based item generation* (Unpublished doctoral dissertation). Münster, Germany: University of Münster.

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

Evans, J. J., Floyd, R. G., McGrew, K. S., & Leforgee, M. H. (2002). The relations between measures of Cattell–Horn–Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review*, *31*, 246-262.

Fiorello, C. A., & Primerano, D. P. (2005). Research into practice: Cattell–Horn–Carroll cognitive assessment in practice: Eligibility and program development issues. *Psychology in the Schools*, *42*, 525-536.

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

Floyd, R. G., Clark, M. H., & Shadish, W. R. (2008). The exchangeability of IQs: Implications for professional psychology. *Professional Psychology: Research and Practice*, *39*, 414-423.

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.

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. doi: 10.1002/pits.20284

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. doi:10.1037/1045-3830.22.2.200

Hambrick, D. Z., & Altmann, E. M. (2015). The role of placekeeping ability in fluid intelligence. *Psychonomic bulletin & review*, *22*, 1104-1110.

Hollingworth, L. S., & Cobb, M. V. (1928). Children clustering at 165 IQ and children clustering at 146 compared for three years in academic achievement. In G. M. Whipple (Ed.), *The twenty-seventh yearbook of the National Society for the Study of Education: Nature and Nurture, Part II—Their influence upon achievement* (pp. 3-33). Bloomington, IL: Public School Publishing Company.

Holzman, T. G., Pellegrino, J. W., & Glaser, R. (1983). Cognitive variables in series completion. *Journal of Educational Psychology*, *75*, 603-618.

Jensen, A. (1969). How much can we boost IQ and scholastic achievement. *Harvard Educational Review*, *39*, 1-123.

Kamphaus, R. W., Winsor, A. P., Rowe, E. W., & Kim, S. (2012). A history of intelligence test interpretation. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 99-144). New York, NY: Guilford Press.

Kaufman, A. S., & Kaufman, N. L. (2004). Kaufman assessment battery for children (2nd ed.). Circle Pines, MN: American Guidance Service.

Keith, T. Z., & Reynolds, M. R. (2010). CHC theory and cognitive abilities: What we've learned from 20 years of research. *Psychology in the Schools*, *47*, 635-650.

Kranzler, J. H., Benson, N., & Floyd, R. G. (2016). Intellectual assessment of children and youth in the United States of America: Past, present, and future. *International Journal of School & Educational Psychology*. Advance online publication. doi:10.1080/21683603.2016.1166759

LeFevre, J.-A., & Bisanz, J. (1986). A cognitive analysis of number-series problems: Sources of individual differences in performance. *Memory & Cognition*, *14*, 287-298.

Letteri, C. A. (1980). Cognitive profile: Basic determinant of academic achievement. *The Journal of Educational Research*, *73*, 195-199.

McGrew, K. S. (1993). The relationship between the Woodcock-Johnson Psychoeducational Battery-Revised Gf-Gc cognitive clusters and reading achievement across the life-span. *Journal of Psychoeducational Assessment Monograph Series: Woodcock-Johnson Psycho-Educational Assessment Battery—Revised,* 39-53.

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.

McGrew, K. S., & Hessler, G. (1995). The relationship between the WJ-R Gf-Gc cognitive clusters and mathematics achievement across the life-span. *Journal of Psychoeducational Assessment*, *13*, 21-38.

McGrew, K. S., & Knopik, S. (1993). The relationship between the WJ-R Gf-Gc cognitive clusters and writing achievement across the life-span. *School Psychology Review*, *22*, 687-695.

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

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

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. doi:10.1002/pits.20497

McGrew, K. S., & Wrightson, W. (1997). The calculation of new and improved WISC-III subtest reliability, uniqueness, and general factor characteristic information through the use of data smoothing procedures. *Psychology in the Schools*, *34*, 181-195.

Mosier, C. I. (1943). On the reliability of a weighted composite. *Psychometrika*, *8*, 161-168.

National Council of Teachers of Mathematics. (2010). Preface. In B. J. Reys, R. E. Reys, & R. Rubenstein (Eds.), *Mathematics curriculum: Issues, trends, and future directions* (Vol. 72, pp. ix-xiv). Reston, VA: Author.

Newton, J. H., & McGrew, K. S. (2010). Introduction to the special issue: Current research in Cattell–Horn–Carroll–based assessment. *Psychology in the Schools*, *47*, 621-634.

Niilekesia, C. R., Reynolds, M. R., Keith, T. Z., & McGrew, K. S. (2016). A special validity study of the Woodcock-Johnson IV: Acting on evidence for specific abilities. In D. P. Flanagan & V. C. Alfonso (Eds.), *WJ IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 65-106). Boston, MA: Elsevier.

Quereshi, M. Y., & Smith, H. (1998). Reasoning ability in older adults measured through letter and number series. *Current Psychology*, *17*, 20-27.

Reynolds, M. R., & Niileksela, C. R. (2015). Test review: Woodcock-Johnson IV Tests of Cognitive Abilities. *Journal of Psychoeducational Assessment*, *33*, 381-390.

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* (pp. 99-144). New York, NY: Guilford Press.

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

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

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

Swanson, H. L. (1994). Short-term memory and working memory do both contribute to our understanding of academic achievement in children and adults with learning disabilities? *Journal of Learning Disabilities*, *27*, 34-50.

Villarreal, V. (2015). Test Review: Woodcock-Johnson IV Tests of Achievement. *Journal of Psychoeducational Assessment*, *33*, 391-398.

Woodcock, R. R., McGrew, K. S., & Schrank, F. A. (2001). *Woodcock-Johnson III Technical Manual*. Itasca, IL: Riverside Publishing.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock–Johnson Psycho–Educational Battery–Revised*. Chicago, IL: Riverside Publishing.

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

Ysseldyke, J., Burns, M., Dawson, P., Kelley, B., Morrison, D., Ortiz, S., . . . Telzrow, C. (2006). *School psychology: A blueprint for training and practice III*. Bethesda, MD: National Association of School Psychologists.