Brief Article

Journal of Psychoeducational Assessment
2025, Vol. 0(0) 1–16

DOI: 10.1177/07342829251381998
journals.sagepub.com/home/jpa
Sage

Incremental Validity of the Successive Level Approach to Intelligence Test Interpretation

Ryan J. McGill[1], Gary L. Canivez[2], and Stefan C. Dombrowski[3]

Abstract

This study investigated the incremental validity of the successive-level approach to intelligence test interpretation, specifically within the context of school psychology. The successive-level approach assumes that unique information is captured at various levels of intelligence tests (e.g., full scale, index-level, subtest-level). However, previous research has indicated that lower-level scores often fail to explain significant or meaningful variance in achievement outcomes beyond what is accounted for by the global score, suggesting potential redundancy in interpreting lower-level scores. Using the Woodcock-Johnson IV (WJ IV) standardization sample, this study examined the relationship between cognitive ability scores (GIA and CHC clusters) and achievement outcomes. The results indicated that, although the CHC subscale scores contributed some unique variance in achievement, their incremental validity was limited and frequently overshadowed by the GIA composite. These findings align with previous research, suggesting that clinicians may unwittingly commit a duplication fallacy when relying on successive-level score interpretation or related guidance in clinical practice.

Keywords

incremental validity, evidence-based assessment, CHC, intelligence testing

Despite long-standing calls for a “paradigm shift” in clinical service delivery, intelligence testing remains a fixture of school psychology training and practice (Benson et al., 2019). Therefore, it is not surprising that the assessment literature continues to be replenished yearly with articles and materials that promise to shed insight on how intelligence tests should be interpreted from an evidence-based assessment perspective (Dombrowski & McGill, 2024). A confluence of statistical

[1]Department of School Psychology and Counselor Education, William & Mary, Williamsburg, VA, USA
[2]Department of Psychology, Eastern Illinois University Charleston, IL, USA
[3]Department of Graduate Education, Leadership, and Counseling, Rider University Lawrenceville, NJ, USA

Corresponding Author:
Ryan J. McGill, Department of School Psychology and Counselor Education, William & Mary, School of Education, P. O. Box 8795 Williamsburg, VA 23187. USA.
Email: rmcgill@wm.edu

discovery (e.g., factor analysis) and evolution of prevailing theories of the structure of human intellectual functioning has led to an ever increasing array of cognitive test scores, complicating debates on these matters. To deal with this complexity, several prominent interpretive "systems" and heuristics have been developed throughout the history of assessment psychology all which purport to provide clinicians with defensible rationales for engaging in successive interpretation within and between all levels of intelligence tests (e.g., full scale [Step1], index-level [Step 2], subtest level [etc.]).

For example, Rappaport and colleagues (1945) produced one of the first formal successive systems of analyses for an earlier version of the Wechsler Scales providing clinicians with a step-by-step process for analyzing various levels of intelligence test data and was the first system to advocate for the analysis of intra-individual strengths and weaknesses. Similar procedures have been articulated, in some form, in virtually every clinical guidebook and test interpretive manual over the course of the last century (Farmer et al., 2021). Thus, it is not surprising that recent surveys indicate that ∼50% of practitioners report utilizing some variant of successive-level[1] analyses when interpreting scores on contemporary cognitive measures (Kranzler et al., 2020).

Though descendants of this approach to intelligence test interpretation have been refined in recent prominent guides (e.g., Flanagan & Alfonso, 2017; Sattler, 2024), the underlying logic remains the same. In other words, the assumption that unique information is captured at every level of the test, warranting focal consideration to develop a more complete understanding of an examinee's distinct cognitive strengths and weaknesses. As will be demonstrated, this assumption is inherently an incremental validity question that can be empirically tested (Haynes et al., 2019).

Prior Incremental Validity Investigations of Intelligence Tests

First described by Sechrest (1963), incremental validity is a relatively straightforward concept: it examines whether a measure or test adds to the prediction or concurrent validity of a clinical criterion beyond what existing sources of data can explain. If so, a compelling debate can be undertaken regarding the relative value of this increment in terms of its practical utility (VanDerHeyden, 2018). Applying the scientific law of parsimony to intelligence test interpretation, the application of this principle suggests that for lower-levels of the test to be meaningfully interpreted, it must be demonstrated that unique information is captured by these scores beyond that accounted for by scores already interpreted at previous levels (i.e., FSIQ). However, it should be noted that there is presently no agreed upon standard by which incremental validity results are regarded as meaningful as a variance increment beyond zero indicates that new information is provided to the clinician to some degree.

Published incremental validity investigations of previous versions of prominent intelligence tests such as the Differential Ability Scales (Youngstrom et al., 1999), Kaufman Assessment Battery for Children-Second Edition (KABC-II; McGill, 2015), Reynolds Intellectual Assessment Scales (RIAS; Nelson & Canivez, 2012), Wechsler Adult Intelligence Scale-Fourth Edition (Canivez, 2013), Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Glutting et al., 2006), and the Woodcock Johnson Tests of Cognitive Abilities-Third Edition (WJ III Cognitive; McGill & Busse, 2015) have been informative on the matter but thus far do not appear to have had much substantive impact on intelligence test interpretation training and practice (Lockwood et al., 2022).

The results of these investigations suggest that, except for a few cases involving unique clinical samples or predictor-criterion overlap, index-level scores whether jointly or in isolation rarely capture unique portions of achievement variance beyond that already accounted for by the omnibus full scale score (e.g., psychometric g). That is, if a practitioner chooses to interpret the full-scale composite and assumes that there is meaningful variance accounted for at Step 2, they would likely be surprised to know how small that variance increment is (e.g., 2% to 5%) in terms of predicting norm-referenced achievement outcomes (Schneider & Kaufman, 2019). Yet, practice guides and clinical training continue to emphasize primary interpretation at the index level, particularly, in the field of school psychology where the use of intelligence tests is largely justified based on the *assumed* concurrent validity of unique cognitive-achievement relations and supposed pedagogical outcomes (Zaboski et al., 2018) that emanate from those linkages.[2] Nevertheless, there are some who detract from this approach to evaluating empirical outcomes. Particularly, those most furtively invested in the broader profile analysis paradigm (e.g., Hale & Fiorello, 2004). To wit, a special issue of *Applied Neuropsychology* was devoted to addressing the *supposed* methodological issues raised by Hale et al. (2007) that called into question previous incremental validity results. Their contentions focused mainly on two issues: (a) reversed order of entry results were equivalent thus resulting in conflicting interpretive guidance in which the FSIQ yielded little to know incremental validity depending on when it was entered, (b) use of an alternative methodology (commonality analysis) which seemed to find that an analog to the FSIQ did not account for meaningful variance when it was put into the regression equation at the end of the analyses, and (c) collinearity among the indicators biases results obtained via hierarchical regression. The order of entry argument, in general, was thoroughly rejected by several scholars throughout the issue. Put simply, one must have a compelling theoretical reason for determining when to enter variables in a hierarchical regression (Pedhazur, 1997). The successive-level approach dictates that the FSIQ is given primacy at Step 1, and this should be entered first in the regression accordingly. The collinearity argument is sound in theory if one chooses to interpret beta weights which are de-stabilized in the presence of collinearity though there is no effect on $R^2/\triangle R^2$, which is the basis of incremental validity interpretation (Tabachnick & Fidell, 2019). Thus, if interpretation focuses on that metric, as is the case here, then collinearity is *not* an issue. Moreover, the threat itself would seem to serve as warning of sorts for the very enterprise of the successive-level approach to test interpretation via interpretive redundancy.

Interestingly, Schneider (2008) thoroughly rebuked the use commonality analyses and the conclusions stemming from that approach with respect to assessing incremental validity outcomes. While debates on these matters produced considerable acrimony in various professional forums during the aughts, all the principal arguments levied against the use of hierarchical regression that were invoked at that time have been found methodologically or conceptually wanting.

Purpose of the Current Study

Despite the consistency of previous results suggesting a general lack of incremental validity for most commercial ability measures, it could be argued that our understanding of the matter is not settled given the adaptive changes of modern tests and the augmentation of the construction of both group-level scores and omnibus composites which promise to augment incremental validity on the basis that they are no longer linear derivatives of the subtests in most circumstances. Accordingly, the purpose of the current study was to address these gaps in the literature. A comprehensive incremental validity study of a modernly constructed ability measure has yet to be conducted to determine whether extant results could be replicated or generalized to our present understanding regarding effective test interpretive practices. This is especially relevant to the field of school psychology where concurrent validity between intelligence tests and norm-referenced achievement test scores is considered a fundamental cornerstone.[3]

Method and Analyses

Participants were drawn from the nationally representative Woodcock-Johnson IV (WJ IV; Schrank et al., 2014a) standardization sample. The sample for the current study contained the age groups of 6-8, 9-13 and 14-19 years to comport with school-age as well as prior incremental studies conducted on previous versions of the measurement instrument (e.g., McGill & Busse, 2015). The focus on school-age was determined because it best comports with previous WJ III/IV cognitive-achievement relations studies and the age ranges at which those linkages are most useful for making diagnostic inferences pertaining to the presence of learning disabilities in children and adolescents. The WJ IV was normed on a nationally representative sample of 7,416 participants from age 2 to 90 plus controlling for census region, gender, country of birth, race, community type, parent education, and occupational level. The age groups contained an average of 885, 1,572, and 1,685 participants, respectively. The total normative sample (N = 7,416) reported should be viewed with caution as a "planned missingness" design was employed during standardization; thus, selected cohorts only received a portion of the tests described and missing data were mathematically imputed (Canivez, 2017).

Measurement Instrument

The WJ IV features 50 subtests that are distributed across three co-normed but distinct batteries: Tests of Cognitive Ability (WJ IV COG; Schrank et al., 2014b), Tests of Achievement (WJ IV ACH; Schrank et al., 2014a), and Tests of Oral Language (WJ IV OL; Schrank et al., 2014b).

The WJ IV COG is a multidimensional test of cognitive abilities for ages 2 to 90 plus years, composed of 18 subtests, 14 of which contribute to the measurement of seven hypothesized Cattell-Horn-Carroll (CHC; Schneider & McGrew, 2018) cluster scores: Comprehension-Knowledge, Fluid Reasoning, Auditory Processing, Visual-Spatial Thinking, Long-Term Retrieval, Short-Term Working Memory, and Cognitive Processing Speed. Additionally, a full scale differentially weighted general intellectual ability (GIA) composite score thought to serve as proxy for general intelligence is provided.[4] All variables on the WJ IV COG are expressed as standard scores with a mean of 100 and a standard deviation of 15. Median reliabilities for the cognitive scores expressed as independent variables in the present study ranged from .86 to .97. Extensive normative and psychometric data can be found in the WJ IV technical manual (McGrew et al., 2014).

The WJ IV ACH is a comprehensive academic assessment battery designed primarily to measure three academic domains: Reading, Mathematics, and Writing. The WJ IV ACH is composed of 20 subtests that combine to provide 20 cluster-level scores as well as a Broad Achievement composite score. Achievement clusters are expressed as standard scores with a mean of 100 and a standard deviation of 15.[5] Median reliabilities for the achievement scores expressed as dependent variables in the present study ranged from .92 to .99. Additional technical information for the WJ IV ACH can be found in the WJ IV Technical Manual (McGrew et al., 2014).

The WJ IV OL represents a new battery for the WJ IV, featuring many subtests that were previously located in the achievement battery for the WJ III. According to the Technical Manual (McGrew et al., 2014) it contains 12 additional oral language, expression, and listening comprehension tasks designed to serve as a diagnostic supplement to the remainder of the battery. Two clusters, expressed as conventional standard scores, contained within that battery are of importance (i.e., Oral Expression and Listening Comprehension) to the present investigation as they were featured in previous incremental validity studies (e.g., McGill & Busse, 2015) and represent focal points of learning disability identification featured in federal regulations (e.g. IDEIA). As such, they were included as additional dependent variables in the current design. Median reliabilities for Oral Expression and Listening Comprehension were .89 and .90, respectively. The WJ IV organizational framework for the IV and DV variables analyzed is provided in Table 1.

Data Analysis

Correlations between the WJ IV COG scores (GIA and seven CHC 7 scores) and the 7 WJ IV ACH scores (Basic Reading, Reading Comprehension, Reading Fluency, Math Calculation, Math Problem

Solving, Written Expression, and Broad Achievement) plus 2 Oral Language scores (Oral Expression and Listening Comprehension) aligning with IDEA were extracted from the WJ IV manual (McGrew et al., 2014). The GIA and seven CHC cluster scores were treated as independent variables (IVs), and the WJ IV ACH/OL achievement scores were treated as dependent variables (DVs).

Hierarchical multiple regression (HMR) using SPSS version 27 using the sample syntax reported in the accompanying Appendix was conducted. HMR is a statistical technique used to assess the relationship between variables while controlling for the influence of other variables. It allows researchers to examine the unique contribution of predictor variables to the outcome, especially when those predictors are entered into the model in a specific order based on theoretical or practical considerations (Tabachnick & Fidell, 2019). Consistent with best practice (i.e., Pedhazur, 1997), the GIA was entered into the first block and the CHC cluster scores were entered both as a concerted CHC block as well as each individual CHC score individually into the second block of the regression models predicting the nine achievement DVs. This way, the incremental combined effects of CHC assessment in total could be assessed as well as uniquely in terms of individual broad abilities. As noted by McGill and Busse (2015), order of entry in HMR is not arbitrary and should be based on consideration of relevant theory, as the WJ IV COG is designed to cohere with CHC theory which depicted in the literature as a hierarchical model with psychometric *g* at the apex, it stands to reason that the corresponding GIA composite should be entered first followed by the CHC clusters which are ordered lower in the model. This is also in keeping with the successive interpretive procedures articulated specifically for the instrument by Sattler (2025) which is the primary focus of the study. The resulting $R^2/\triangle R^2$ estimates were interpreted as an effect size using Cohen's (1988) conventional guidelines which are as follows: “small,” .01; “medium,” .09; and “large,” .25.

Results

Across ages 6–19, in the 27 HMR models that were examined (Tables 2–4), the CHC cluster scores, both jointly and in isolation consistently failed to account for meaningful effects for unique variance of WJ IV achievement and oral language dimensions ($\triangle R^2$ = .01–.23, *Mdn* = .06) beyond the variance accounted for by the GIA composite score (R^2 = .27−.71, *Mdn* = .49). However, there were notable instances (i.e., Oral Expression and Listening Comprehension) where various combinations of CHC cluster scores accounted for large effects beyond the GIA, though raises concern about the jangle fallacy given these variables. For example, the broad ability (Comprehension-Knowledge) variable that consistently explains the most unique variance in the case of Oral Expression also contains a measure of “Vocabulary” that is differentiated based only on an apparent facet in terms of the presentation of that content to the examinee. One could easily make the argument that Oral Expression is a derivative of that attribute which has simultaneously also been defined as “Verbal Ability” via the WJ III Woodcock Performance Model that was thought to oversee the CHC aspects of that test's theoretical structure (Taub & McGrew, 2014).

While these results were largely consistent with incremental validity results from the WJ III (McGill & Busse, 2015), some notable departures were observed. Namely, the predictive effects of the GIA were relatively weaker overall and the unique effects of the CHC cluster scores were stronger across some areas of achievement and oral language outcomes approaching or equaling what can be interpreted as moderate-large effects on the WJ IV ACH/OL. For example, out of the 27 regression models that were estimated, moderate to large effects were consistently observed across school-age for Oral Expression and Listening Comprehension. Of note, moderate effects were observed for Reading Fluency at ages 9–13 (CHC total) and 14–19 (CHC total), and for Math Problem Solving at ages 9–13 (CHC total, Fluid Reasoning). The finding of moderate effects for the latter two outcomes are a notable development for the measure and a positive addition to the broader incremental validity literature for commercial ability measures. Though it should be noted that, in isolation, CHC cluster scores mostly accounted for negligible effects even when the combined

Table 1. Assignment of WJ IV Tests to Cognitive-Achievement Interpretive Clusters

Battery	Cluster																
Test of cognitive abilities	GIA	Gc	Gf	Gwm	Gs	Ga	Glr	Gv	OE	LC	BR	RC	RF	MC	MP	WE	BA
Oral vocabulary	*	*															
Number series	*		*														
Verbal attention	*			*													
Letter-pattern matching	*				*												
Phonological processing	*					*											
Story recall	*						*										
Visualization								*									
General information																	
Concept formation			*														
Numbers reversed				*													
Nonword repetition						*											
Visual-auditory learning							*										
Picture recognition								*									
Pair cancellation					*												
Tests of achievement																	
Letter-word identification											*						*
Word attack											*						
Passage comprehension												*					*
Reading recall												*					
Oral reading													*				
Sentence reading fluency													*				*
Calculation														*			*
Math facts fluency														*			*
Applied problems															*		
Number matrices															*		*
Writing samples																*	*
Spelling																	*
Sentence writing fluency																*	*
Tests of oral language																	
Picture vocabulary									*								
Oral comprehension										*							
Sentence repetition									*								
Understanding directions										*							

Note. GIA = General Intellectual Ability, Gc = Comprehension-Knowledge, Gf = Fluid Reasoning, GSM = Short-Term Working Memory, Gs = Cognitive Processing Speed, Ga = Auditory Processing, Glr = Long-Term Retrieval, Gv = Visual-Spatial Thinking, OE = Oral Expression, LC = Listening Comprehension, BR = Basic Reading Skills, RC = Reading Comprehension, RF = Reading Fluency, MC = Math Calculation Skills, MP = Math Problem Solving, WE = Written Expression, BA = Broad Achievement.

Table 2. Incremental Contribution of Woodcock–Johnson IV Tests of Cognitive Abilities CHC Scores in Predicting Woodcock–Johnson IV Tests of Achievement Scores Beyond the GIA for Normative Participants Ages 6-8 (N = 885)

	Basic reading skills			Reading comprehension			Reading fluency		
Predictor	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.49*	-	49	.53*	-	53	.44*	-	44
CHC cluster scores (*df* = 7)	.54	.05*	5	.59	.06*	6	.49	.05*	5
Comprehension-knowledge	.49	.00*	0	.53	.00	0	.44	.01*	1
Fluid reasoning	.51	.02*	2	.56	.03*	3	.44	.01*	1
Auditory processing	.50	.01*	1	.53	.00	0	.44	.00	0
Visual-spatial thinking	.49	.00*	0	.54	.00*	0	.44	.00	0
Long-term retrieval	.49	.00	0	.54	.00	0	.44	.00	0
Short-term working memory	.49	.00	0	.54	.01*	1	.44	.00	0
Cognitive processing speed	.51	.02*	2	.56	.03*	3	.45	.02*	2
	Math calculation skills			Math problem solving			Written expression		
	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.53*	-	53	.55*	-	55	.49*	-	49
CHC cluster scores (*df* = 7)	.61	.07*	7	.63	.08*	8	.52	.03*	3
Comprehension-knowledge	.54	.00*	0	.55	.01*	1	.51	.02*	2
Fluid reasoning	.54	.01*	1	.62	.07*	7	.49	.00*	0
Auditory processing	.58	.05*	5	.55	.00*	0	.49	.00	0
Visual-spatial thinking	.54	.00*	0	.55	.00	0	.49	.00	0
Long-term retrieval	.54	.01*	1	.55	.00	0	.49	.00	0
Short-term working memory	.54	.00*	0	.55	.00	0	.49	.00	0
Cognitive processing speed	.56	.02*	2	.56	.01*	1	.49	.00	0
	Oral expression			Listening comprehension			Broad achievement		
	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.27*	-	27	.41*	-	41	.66*	-	66
CHC cluster scores (*df* = 7)	.51	.23*	23	.63	.22*	22	.68	.03*	3
Comprehension-knowledge	.42	.15*	15	.46	.05*	5	.66	.00	0
Fluid reasoning	.28	.01*	1	.47	.06*	6	.67	.01*	1
Auditory processing	.37	.10*	10	.44	.03*	3	.66	.00*	0
Visual-spatial thinking	.29	.02*	2	.46	.05*	5	.66	.00	0
Long-term retrieval	.28	.01*	1	.44	.03*	3	.66	.00*	0
Short-term working memory	.27	.00	0	.42	.01*	1	.66	.00*	0
Cognitive processing speed	.28	.01*	1	.44	.03*	3	.66	.00*	0

Note. GIA = General Intellectual Ability composite score. CHC = Cattell–Horn–Carroll scores. All coefficients rounded to nearest hundredth, may not equate due to rounding.

[a]Represents proportion of variance accounted for by variables at their entry point into regression equation. *R2/ΔR2* values multiplied by 100.

*p < .05.

model indicated moderate to large effects overall. Whether this is the outcome of the fact that the GIA is differentially weighted using only half of the CHC narrow dimension scores from the WJ III or a function of improved measurement, awaits further determination. Nevertheless, there are prescient examples of note; particularly the contribution of Comprehension-Knowledge and Fluid Reasoning for Oral Expression and Math Problem Solving, respectively, that users must consider.

Discussion

The joint test standards (AERA, APA, & NCME, 2014) stress the importance of demonstrating that subscale scores on published tests have sufficient differentiation from composite scores in their ability to convey unique information. Incremental validity is one aspect of broader construct validity that provides a useful vantage for evaluating those claims. The present results suggest that while the WJ IV does provide improvements in the increments of unique achievement and Oral Language variance captured by the CHC subscales beyond that already accounted for by the GIA composite, issues with interpretive redundancy likely remain for most of the prediction models examined. This is particularly true when following successive-level interpretive systems which presume that unique information is captured at all levels of the test as a matter of course. To wit, Cognitive Processing Speed, Long-Term Storage and Retrieval, and Short-Term Working Memory consistently failed to account for incremental variance percentages beyond ~0-1.

The results from this study, coupled with previous incremental validity investigations of other intelligence tests, suggest that the assumption undergirding the successive-level approach may lead clinicians to unwittingly commit a duplication fallacy when engaging in the successive-level approach, where redundant information is treated as informative when it is likely illusory (Garb et al., 2008). While previous CHC cognitive-achievement relations studies featuring the WJ IV (e.g., Cormier et al., 2016, 2017) suggest there is a wide array of unique patterns and configurations of cognitive scores predicting[6] specific areas of achievement outcomes over the course of the lifespan, regardless of one's theoretical vantage point, the existence of such patterns cannot be dismissed and suggest that there is proof of concept for Step 2 guided interpretations. Further complicating the matter are structural equation modeling[7] studies which (e.g., Caemmerer et al., 2018) report significant direct effects for various CHC dimensions on areas of achievement in concert with moderate to large indirect effects for *g*.

However, extrapolating from this body of literature to inform focal clinical decisions may be more complicated than is often portrayed in a majority of the WJ IV interpretive literature. For example, a meta-analysis by Zaboski and colleagues (2018) concluded that, once the effects of general intelligence are considered, the influence of those patterns largely disappears, apart from unique effects associated with Crystallized Ability (i.e., Comprehension-Knowledge). While the uniqueness of these patterns should not be discounted, their relative specificity for focal diagnostic and treatment decisions awaits further adjudication given the potential limitations noted above.

Limitations

Despite the present results, the present study must be de-limited in terms of its contribution to the literature in concert with its limitations. First and foremost, as previously discussed what is regarded as "meaningful" within the context of incremental validity is largely subjective or in the proverbial eye of the beholder. Of course, there are Cohen's (1988) guidelines for interpreting R^2 but it is debatable whether those thresholds should be applied to $\triangle R^2$ and, if so, in what contexts. For example, one may find a variance increment of as low as 5% to be meaningful in a different predictive context. For example, if the total variance explained and that uniquely by a general composite are both low then there are circumstances in which 5% of additional variance cannot be dismissed. While that is generally not

Table 3. Incremental Contribution of Woodcock-Johnson IV Tests of Cognitive Abilities CHC Scores in Predicting Woodcock-Johnson IV Tests of Achievement Scores Beyond the GIA for Normative Participants Ages 9-13 (*N* = 1,572)

Predictor	Basic reading skills			Reading comprehension			Reading fluency		
	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.48*	-	48	.49*	-	49	.42*	-	42
CHC cluster scores (df = 7)	.48	.01*	1	.54	.05*	5	.51	.09*	9
Comprehension-knowledge	.48	.01*	1	.50	.01*	1	.46	.03*	3
Fluid reasoning	.48	.00	0	.51	.02*	2	.42	.00	0
Auditory processing	.48	.01*	1	.49	.00	0	.42	.00	0
Visual-spatial thinking	.48	.00	0	.49	.00	0	.42	.00*	0
Long-term retrieval	.49	.02*	2	.49	.00	0	.42	.00	0
Short-term working memory	.48	.00	0	.50	.01*	1	.42	.00	0
Cognitive processing speed	.48	.01*	1	.51	.02*	2	.45	.03*	3
	Math calculation skills			Math problem solving			Written expression		
	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.53*	-	53	.58*	-	58	.45*	-	45
CHC cluster scores (df = 7)	.61	.08*	8	.67	.09*	9	.47	.02*	2
Comprehension-knowledge	.53	.00	0	.58	.00*	0	.47	.02*	2
Fluid reasoning	.54	.01*	1	.65	.08*	8	.45	.00	0
Auditory processing	.56	.03*	3	.58	.00*	0	.45	.00	0
Visual-spatial thinking	.53	.00*	0	.58	.00	0	.45	.00	0
Long-term retrieval	.54	.01*	1	.58	.00	0	.45	.00	0
Short-term working memory	.54	.00*	0	.58	.00*	0	.45	.00*	0
Cognitive processing speed	.57	.04*	4	.58	.00*	0	.45	.01*	1
	Oral expression			Listening comprehension			Broad achievement		
	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.37*	-	37	.44*	-	44	.66*	-	66
CHC cluster scores (df = 7)	.53	.16*	16	.54	.10*	10	.70	.04*	4
Comprehension-knowledge	.47	.09*	9	.48	.04*	4	.66	.01*	1
Fluid reasoning	.39	.02*	2	.44	.00	0	.66	.01*	1
Auditory processing	.44	.07*	7	.45	.01*	1	.66	.00*	0
Visual-spatial thinking	.38	.01*	1	.47	.03*	3	.66	.00	0
Long-term retrieval	.37	.00	0	.46	.03*	3	.66	.00*	0
Short-term working memory	.37	.00*	0	.44	.00	0	.66	.00*	0
Cognitive processing speed	.38	.01*	1	.44	.00	0	.67	.01*	1

Note. GIA = General Intellectual Ability composite score. CHC = Cattell–Horn–Carroll scores. All coefficients rounded to nearest hundredth, may not equate due to rounding.
[a]Represents proportion of variance accounted for by variables at their entry point into regression equation. *R2/ΔR2* values multiplied by 100.
* $p < .05$.

the case in the present study, other studies have reported such outcomes (e.g., Breit et al., 2024; Nelson & Canivez, 2012) which complicates blanket application of Cohen's (1988) rules of thumb to automatically dismiss incremental validity outcomes that improve from zero. Nevertheless, the successive-level approach to test interpretation rests on the assumption that most, if not all, interpretive weight should be focused on levels beyond the first step in terms of profile analytic applications and nomothetic and/or idiographic strengths and weakness inferences. As such, practitioners should consider whether sufficient variance is accounted by any individual WJ IV score at that stage of interpretation in concert with available psychometric evidence (Kranzler & Floyd, 2020).

Table 4. Incremental Contribution of Woodcock–Johnson IV Tests of Cognitive Abilities CHC Scores in Predicting Woodcock–Johnson IV Tests of Achievement Scores Beyond the GIA for Normative Participants Ages 14-19 (N = 1,685)

	Basic reading skills			Reading comprehension			Reading fluency		
Predictor	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.49*	-	49	.53*	-	53	.44*	-	44
CHC cluster scores (*df* = 7)	.53	.04*	4	.58	.05*	5	.52	.09*	9
Comprehension-knowledge	.50	.01*	1	.54	.01*	1	.47	.03*	3
Fluid reasoning	.49	.00	0	.56	.02*	2	.44	.00	0
Auditory processing	.50	.01*	1	.54	.00*	0	.44	.00	0
Visual-spatial thinking	.49	.00	0	.53	.00	0	.44	.00*	0
Long-term retrieval	.50	.01*	1	.53	.00	0	.44	.00	0
Short-term working memory	.49	.00	0	.54	.01*	1	.44	.00	0
Cognitive processing speed	.50	.01*	1	.55	.02*	2	.47	.03*	3
	Math calculation skills			Math problem solving			Written expression		
	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.56*	-	56	.64*	-	64	.49*	-	49
CHC cluster scores (*df* = 7)	.62	.06*	6	.70	.06*	6	.51	.02*	2
Comprehension-knowledge	.56	.00	0	.65	.01*	1	.50	.01*	1
Fluid reasoning	.57	.01*	1	.68	.04*	4	.49	.00	0
Auditory processing	.59	.03*	3	.64	.00*	0	.49	.00	0
Visual-spatial thinking	.56	.00*	0	.64	.00	0	.49	.00	0
Long-term retrieval	.57	.01*	1	.64	.00	0	.49	.00	0
Short-term working memory	.57	.01*	1	.64	.00*	0	.49	.00*	0
Cognitive processing speed	.58	.02*	2	.65	.01*	1	.49	.00*	0
	Oral expression			Listening comprehension			Broad achievement		
	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]	R^2	ΔR^2	Increment (%)[a]
GIA	.44*	-	44	.50*	-	50	71*	-	71
CHC cluster scores (*df* = 7)	.57	.13*	13	.58	.07*	7	.74	.04*	4
Comprehension-knowledge	.49	.05*	5	.52	.01*	1	.72	.01*	1
Fluid reasoning	.46	.03*	3	.51	.00	0	.71	.01*	1
Auditory processing	.50	.06*	6	.52	.02*	2	.71	.01*	0
Visual-spatial thinking	.44	.01*	1	.53	.02*	2	.71	.00	0
Long-term retrieval	.44	.00	0	.52	.02*	2	.71	.00*	0
Short-term working memory	.44	.00*	0	.51	.00*	0	.71	.00*	0
Cognitive processing speed	.44	.01*	1	.52	.02*	2	.71	.01*	1

Note. GIA = General Intellectual Ability composite score. CHC = Cattell–Horn–Carroll scores. All coefficients rounded to nearest hundredth, may not equate due to rounding.
[a]Represents proportion of variance accounted for by variables at their entry point into regression equation. *R2/ΔR2* values multiplied by 100.
*p < .05.

Further, Tucker-Drob (2009) has provided convincing evidence that there are age effects in terms of the growth and trajectory of broad cognitive abilities as purportedly measured by the WJ IV. Descriptive evidence of differential prediction can be ascertained across the age range in the present study though not designed to assess for those effects specifically. For example, the variance explained by the GIA seems to increase across the age range for Oral Expression and yet the broad abilities for that variable as well as Math Problem Solving seem to vacillate countering the overarching hypothesis proffered by Tucker-Drob (2009) suggesting lower-order abilities grow monotonically in importance across the lifespan. Further, one also must also consider an additional aspect of differentiation and that is one of ability differentiation also known as Spearman's law of dimensioning returns (Briet et al., 2019). That is, as ability level increases, the role of specific cognitive abilities increases and vice versa (Breit et al., 2022). Consequently, it cannot be assumed that the present results will generalize to focal clinical assessment situations where clients may perform at the extreme tails of the normative distribution (e.g. gifted assessment). Future research focused on these aspects of prediction is welcomed but cannot be ascertained from the present design. Particularly, given the conspicuous decline in unique prediction accounted for by CHC abilities across the age range which would seem to contradict intuitive understanding on these matters.

Conclusion

To conclude, adjudicating whether psychometric *g* represents a viable psychological dimension or is merely a statistical artifact is beyond the scope of the present investigation though hotly contested within the literature (e.g., McGrew et al., 2023). Thus, it may be tempting for clinicians to believe that bypassing the initial steps in successive-level systems, which all encourage primary interpretation of the global composite at Step 1, will serve as a sufficient safeguard against interpretive redundancy. While a common practice (e.g., Kranzler et al., 2020), this approach conflates construct-irrelevant variance at lower-levels of measurement, making it difficult to disentangle at the individual level when ascribing performance to any individual subscale attribute (Canivez, 2025). That is, that clinicians should do well to remember that all cognitive measures at lower levels of the successive hierarchy of interpretation are not "pure" measures of their respective constructs. They contain degrees of error variance and common variance that are attributable to elements which the title they are assigned may not be aligned; including a vast majority or variance in some circumstances that is already accounted for in Step 1 (Carroll, 1995). That is not to say that levels beyond Step 1 aren't important. The present results provide several examples in which meaningful variance may be accounted for in Step 2. The important takeaway is that is not always the case.

Regardless of theoretical considerations, the psychometric implications of this and prior incremental validity research suggests that users will likely have to consider the implications of the potential redundancies between general and specific abilities if clinical training and the accompanying training literature continues to encourage the simultaneous interpretation of composite and index-level scores (e.g., clusters) derived from the same indicators as a general matter of course.

ORCID iDs

Ryan J. McGill https://orcid.org/0000-0002-5138-0694
Gary L. Canivez https://orcid.org/0000-0002-5347-6534
Stefan C. Dombrowski https://orcid.org/0000-0002-8057-3751

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. The term "successive-level" (a.k.a., "levels of analysis") can be contemporaneously sourced to Sattler (2024) but the approach is synonymous with the "intelligent testing" method previously espoused by Kaufman (1979) which is virtually identical in its original form to the Rappaport et al. system.
2. Prominently cited cognitive-achievement relations studies (see McGrew & Wendling, 2010) often omit exploring the influence of *g* or attempt to underexplain its importance when that variable is included which is itself a form of specification bias.
3. Please see Sattler (2024) for the stepwise interpretive procedures recommended for the measurement instrument targeted in the current study that align with conventional guidance on the matter.
4. The WJ IV GIA score is composed of a differentially weighted combination of only the first seven subtests in the standard battery as opposed to the 14 subtest combination that comprised the GIA-Extended in the previous version featuring all tests that combine to form the primary seven CHC clusters. This would *seemingly* allow for greater degrees of freedom and differentiation between the omnibus full scale score and the clusters given the indicators are already imperfectly correlated, allaying prior fears of collinearity (i.e., redundancy in tests) previously raised by Hale et al. (2007) and summarily dismissed in a special issue dedicated to the topic in *Applied Neuropsychology*.
5. It is not clear how it is possible to extract a surfeit of *k* latent factors that are equivalent to the number of indicators given they are mathematically under-identified (Dombrowski et al., 2019).
6. Technically, these studies are not "predictive" because measurement occurs concurrently between the DVs and IVs. Nevertheless, the term is used commonly in the cognitive assessment literature.
7. Results from incremental validity studies dealing with observed scores are not directly exchangeable with studies such as this that rely on latent scores (Canivez, 2025).

References

American Educational Research Association, American Psychological Association, & National Council on Measurement on Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Benson, N. F., Floyd, R. G., Kranzler, J. H., Eckert, T. L., Fefer, S. A., & Morgan, G. B. (2019). Test use and assessment practices of school psychologists in the United States: Findings from the 2017 National Survey. *Journal of School Psychology*, *72*, 29–48. https://doi.org/10.1016/j.jsp.2018.12.004

Breit, M., Brunner, M., Molenaar, D., & Preckel, F. (2022). Differentiation hypotheses of intelligence: A systematic review of the empirical evidence and an agenda for future research. *Psychological Bulletin*, *148*(7-8), 518–554. https://doi.org/10.1037/bul0000379

Breit, M., Brunner, M., & Preckel, F. (2019). Spearman's law of diminishing returns and its implications for theories of general intelligence and intelligence testing. In D. J McFarland (Ed.), *General and specific mental abilities* (pp. 340–358). Cambridge Scholars Publishing.

Breit, M., Scherrer, V., & Preckel, F. (2024). How useful are specific cognitive ability scores? An investigation of their stability and incremental validity beyond general intelligence. *Intelligence*, *103*, 101816. https://doi.org/10.1016/j.intell.2024.101816

Caemmerer, J. M., Maddocks, D. L. S., Keith, T. Z., & Reynolds, M. R. (2018). Effects of cognitive abilities on child and youth academic achievement: Evidence from the WISC-V and WIAT-III. *Intelligence*, *68*, 6–20. https://doi.org/10.1016/j.intell.2018.02.005

Canivez, G. L. (2013). Incremental criterion validity of WAIS-V factor index scores: Relationships with WIAT-II and WIAT-III subtest and composite scores. *Psychological Assessment*, *25*(2), 484–495. https://doi.org/10.1037/a0032092

Canivez, G. L. (2017). Review of the woodcock–Johnson IV. In J. F. Carlson, K. F. Geisinger, & J. L. Jonson (Eds.), *The twentieth mental measurements yearbook* (pp. 875–882). Buros center for testing.

Canivez, G. L. (2025). Ethics, test standards, and test interpretation: Measurement matters. In G. L. Canivez (Ed.), *Assessing psychometric fitness of intelligence tests: Toward evidence-based interpretation practices* (pp. 1–84). Rowman and Littlefield.

Carroll, J. B. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research, 30*(3), 429–452. https://doi.org/10.1207/s15327906mbr3003_6

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence 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*(8), 787–803. https://doi.org/10.1002/pits.21945

Cormier, D. C., Bulut, O., McGrew, K. S., & Singh, D. (2017). Exploring the relations between cattell-horn-carroll (CHC) cognitive abilities and mathematics achievement. *Applied Cognitive Psychology*, *31*(5), 530–538. https://doi.org/10.1002/acp.3350

Dombrowski, S. C., Beaujean, A. A., McGill, R. J., & Benson, N. F. (2019). The Woodcock-Johnson IV tests of achievement provides too many scores for clinical interpretation. *Journal of Psychoeducational Assessment*, *37*(7), 819–836. https://doi.org/10.1177/0734282918800745

Dombrowski, S. C., & McGill, R. J. (2024). Clinical assessment in school psychology: Impervious to scientific reform? *Canadian Journal of School Psychology*, *39*(4), 297–306. https://doi.org/10.1177/08295735231224052

Farmer, R. L., McGill, R. J., Dombrowski, S. C., & Canivez, G. L. (2021). Why questionable assessment practices remain popular in school psychology: Instructional materials as pedagogic vehicles. *Canadian Journal of School Psychology*, *36*(2), 98–114. https://doi.org/10.1177/0829573520978111

Flanagan, D. P., & Alfonso, V. C. (2017). *Essentials of WISC-V assessment*. Wiley.

Garb, H. N., Lilienfeld, S. O., & Fowler, K. A. (2008). Psychological assessment and clinical judgment. In J. E. Maddux & B. A. Winstead (Eds.), *Psychopathology: Foundations for a contemporary understanding* (2nd ed., pp. 103–124). Routledge.

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. *The Journal of Special Education*, *40*(2), 103–114. https://doi.org/10.1177/00224669060400020101

Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook*. Guilford Press.

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*(1), 37–51. https://doi.org/10.1080/09084280701280445

Haynes, S. N., Smith, G. T., & Hunsley, J. D. (2019). *Scientific foundations of clinical assessment* (2nd ed.). Routledge.

Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. Wiley.

Kranzler, J. H., & Floyd, R. G. (2020). *Assessing intelligence in children and adolescents: A practical guide for evidence-based assessment* (2nd ed.). Rowman and Littlefield.

Kranzler, J. H., Maki, K. E., Benson, N. F., Eckert, T. L., Floyd, R. G., & Fefer, S. A. (2020). How do school psychologists interpret intelligence tests for the identification of specific learning disabilities? *Contemporary School Psychology*, *24*(4), 445–456. https://doi.org/10.1007/s40688-020-00274-0

Lockwood, A. B., Benson, N., Farmer, R. L., & Klatka, K. (2022). Test use and assessment practices of school psychology training programs: Findings from a 2020 survey of US faculty. *Psychology in the Schools*, *59*(4), 698–725. https://doi.org/10.1002/pits.22639

McGill, R. J. (2015). Interpretation of KABC-II scores: An evaluation of the incremental validity of Cattell-Horn-Carroll (CHC) factor scores in predicting achievement. *Psychological Assessment*, *27*(4), 1417–1426. https://doi.org/10.1037/pas0000127

McGill, R. J., & Busse, R. T. (2015). Incremental validity of the WJ III COG: Limited predictive effects beyond the GIA-E. *School Psychology Quarterly: The Official Journal of the Division of School Psychology, American Psychological Association*, *30*(3), 353–365. https://doi.org/10.1037/spq0000094

McGrew, K. S., LaForte, E. M., & Shrank, F. A. (2014). *Technical manual: Woodcock-johnson IV.* Riverside.

McGrew, K. S., Schneider, W. J., Decker, S. L., & Bulut, O. (2023). A psychometric network analysis of CHC intelligence measures: Implications for research, theory, and interpretation of broad CHC scores beyond *g*. *Journal of Intelligence*, *11*(1), 1–33. https://doi.org/10.3390/jintelligence11010019

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*(7), 651–675. https://doi.org/10.1002/pits.20497

Nelson, J. M., & Canivez, G. L. (2012). Examination of the structural, convergent, and incremental validity of the reynolds intellectual assessment scales (RIAS) with a clinical sample. *Psychological Assessment*, *24*(1), 129–140. https://doi.org/10.1037/a0024878

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Harcourt Brace.

Rapaport, D., Gill, M., & Schafer, R. (1945) *Diagnostic psychological testing* (*1*). Year Book Medical. https://doi.org/10.1037/10834-000

Sattler, J. M. (2024). *Assessment of children: Cognitive foundations and applications* (7th ed.). Sattler Publishing.

Schneider, W. J. (2008). Playing statistical ouija board with commonality analysis: Good questions, wrong assumptions. *Applied Neuropsychology*, *15*(1), 44–53. https://doi.org/10.1080/09084280801917566

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

Schrank, F. A., Mather, N., & McGrew, K. S. (2014a). *Woodcock-johnson IV tests of achievement*. Riverside.

Schrank, F. A., Mather, N., & McGrew, K. S. (2014b). *Woodcock-Johnson IV tests of oral language*. Riverside.

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

Schrank, F. A., McGrew, K. S., & Mather, N. (2014b). *Woodcock-Johnson IV tests of cognitive abilities*. Riverside.

Sechrest, L. (1963). Incremental validity: A recommendation. *Educational and Psychological Measurement*, *23*(1), 153–158. https://doi.org/10.1177/001316446302300113

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Taub, G. E., & McGrew, K. S. (2014). The Woodcock–Johnson tests of cognitive abilities III's cognitive performance model: Empirical support for intermediate factors within CHC theory. *Journal of Psychoeducational Assessment*, *32*(3), 187–201. https://doi.org/10.1177/0734282913504808

Tucker-Drob, E. M. (2009). Differentiation of cognitive abilities across the life span. *Developmental Psychology*, *45*(4), 1097–1118. https://doi.org/10.1037/a0015864

VanDerHeyden, A. M. (2018). Why do school psychologists cling to ineffective practices? let's do what works. *School Psychology Forum: Research in Practice*, *12*(1), 44–52.

Youngstrom, E. A., Kogos, J. L., & Glutting, J. J. (1999). Incremental efficacy of differential ability scales factor scores in predicting individual achievement criteria. *School Psychology Quarterly*, *14*(1), 26–39. https://doi.org/10.1037/h0088996

Zaboski, B. A. I. I., Kranzler, J. H., & Gage, N. A. (2018). Meta-analysis of the relationship between academic achievement and broad abilities of the Cattell-Horn-Carroll theory. *Journal of School Psychology*, *71*, 42–56. https://doi.org/10.1016/j.jsp.2018.10.001

Appendix

A. Sample SPSS Syntax for HMR Analyses

```
*Incremental Validity Using Matrix as Input.
MATRIX DATA Variables = rowtype_ X1 X2 X3 X4 Y1
/format = lower diagonal.
BEGIN DATA.
MEAN 100 100 100 100 100
STDDEV 15 15 15 15 15
N 100 100 100 100 100
CORR 1.00
CORR 0.00 1.00
CORR 0.00 0.00 1.00
CORR 0.00 0.00 0.00 1.00
CORR 0.00 0.00 0.00 0.00 1.00
END DATA.
regression matrix = in (*)
/STATISTICS COEFF OUTS R ANOVA CHANGE COLLIN
/CRITERIA = PIN(.05) POUT (.10)
/DEPENDENT Y1
/METHOD = ENTER X1
/METHOD = ENTER X2 X3 X4.
```

B. Intercorrelation Matrix for Cognitive-Achievement Cluster Scores from the WJ IV Normative Sample Ages 6-8

	GIA	Gc	Gf	Gsm	Gs	Ga	Glr	Gv	OE	LC	BR	RC	RF	MC	MP	WE	BA
GIA	–																
Gc	0.60	–															
Gf	0.80	0.43	–														
Gsm	0.72	0.40	0.54	–													
Gs	0.63	0.29	0.37	0.42	–												
Ga	0.70	0.49	0.54	0.56	0.33	–											
Glr	0.58	0.38	0.47	0.40	0.28	0.48	–										
Gv	0.54	0.36	0.45	0.37	0.28	0.48	0.57	–									
OE	0.52	0.62	0.35	0.40	0.24	0.59	0.39	0.40	–								
LC	0.64	0.56	0.36	0.53	0.26	0.57	0.52	0.54	0.69	–							
BR	0.70	0.46	0.64	0.49	0.32	0.55	0.37	0.42	0.51	0.58	–						
RC	0.73	0.46	0.69	0.46	0.33	0.51	0.39	0.44	0.47	0.55	0.80	–					
RF	0.66	0.47	0.57	0.47	0.52	0.48	0.35	0.37	0.46	0.45	0.76	0.74	–				
MC	0.73	0.40	0.63	0.48	0.58	0.36	0.34	0.34	0.33	0.38	0.60	0.61	0.71	–			
MP	0.74	0.50	0.75	0.53	0.39	0.48	0.44	0.43	0.47	0.53	0.54	0.53	0.50	0.70	–		
WE	0.70	0.30	0.59	0.49	0.42	0.49	0.42	0.40	0.32	0.44	0.76	0.76	0.65	0.65	0.46	–	
BA	0.81	0.50	0.71	0.55	0.55	0.52	0.43	0.45	0.48	0.54	0.86	0.83	0.84	0.86	0.70	0.85	–

Note. GIA = General Intellectual Ability, Gc = Comprehension-Knowledge, Gf = Fluid Reasoning, GSM = Short-Term Working Memory, Gs = Cognitive Processing Speed, Ga = Auditory Processing, Glr = Long-Term Retrieval, Gv = Visual-Spatial Thinking, OE = Oral Expression, LC = Listening Comprehension, BR = Basic Reading Skills, RC = Reading Comprehension, RF = Reading Fluency, MC = Math Calculation Skills, MP = Math Problem Solving, WE = Written Expression, BA = Broad Achievement.

C. Intercorrelation Matrix for Cognitive-Achievement Cluster Scores from the WJ IV Normative Sample Ages 9-13

	GIA	Gc	Gf	Gsm	Gs	Ga	Glr	Gv	OE	LC	BR	RC	RF	MC	MP	WE	BA
GIA	–																
Gc	0.65	–															
Gf	0.79	0.46	–														
Gsm	0.72	0.45	0.51	–													
Gs	0.60	0.28	0.39	0.38	–												
Ga	0.67	0.48	0.50	0.52	0.31	–											
Glr	0.55	0.37	0.47	0.38	0.25	0.40	–										
Gv	0.50	0.35	0.38	0.34	0.30	0.39	0.49	–									
OE	0.61	0.63	0.40	0.47	0.29	0.60	0.35	0.36	–								
LC	0.66	0.58	0.54	0.50	0.28	0.53	0.50	0.49	0.69	–							
BR	0.69	0.51	0.56	0.50	0.35	0.52	0.27	0.32	0.55	0.55	–						
RC	0.70	0.53	0.64	0.44	0.32	0.46	0.36	0.34	0.51	0.54	0.72	–					
RF	0.65	0.56	0.53	0.46	0.53	0.45	0.34	0.29	0.48	0.45	0.70	0.68	–				
MC	0.73	0.47	0.63	0.49	0.59	0.36	0.33	0.33	0.40	0.42	0.59	0.60	0.70	–			
MP	0.76	0.54	0.77	0.52	0.42	0.46	0.42	0.39	0.52	0.52	0.53	0.56	0.55	0.71	–		
WE	0.67	0.34	0.55	0.44	0.46	0.43	0.38	0.33	0.35	0.41	0.69	0.70	0.66	0.65	0.48	–	
BA	0.81	0.59	0.69	0.55	0.57	0.50	0.40	0.39	0.54	0.57	0.82	0.80	0.83	0.88	0.73	0.83	–

Note. GIA = General Intellectual Ability, Gc = Comprehension-Knowledge, Gf = Fluid Reasoning, GSM = Short-Term Working Memory, Gs = Cognitive Processing Speed, Ga = Auditory Processing, Glr = Long-Term Retrieval, Gv = Visual-Spatial Thinking, OE = Oral Expression, LC = Listening Comprehension, BR = Basic Reading Skills, RC = Reading Comprehension, RF = Reading Fluency, MC = Math Calculation Skills, MP = Math Problem Solving, WE = Written Expression, BA = Broad Achievement.

D. Intercorrelation Matrix for Cognitive-Achievement Cluster Scores from the WJ IV Normative Sample Ages 14–19

	GIA	Gc	Gf	Gsm	Gs	Ga	Glr	Gv	OE	LC	BR	RC	RF	MC	MP	WE	BA
GIA	–																
Gc	0.73	–															
Gf	0.81	0.52	–														
Gsm	0.75	0.51	0.53	–													
Gs	0.64	0.38	0.42	0.45	–												
Ga	0.70	0.53	0.52	0.57	0.39	–											
Glr	0.59	0.45	0.51	0.41	0.34	0.43	–										
Gv	0.50	0.35	0.40	0.35	0.31	0.39	0.50	–									
OE	0.66	0.64	0.44	0.52	0.36	0.64	0.41	0.39	–								
LC	0.71	0.60	0.59	0.56	0.36	0.59	0.53	0.49	0.71	–							
BR	0.70	0.58	0.57	0.53	0.38	0.56	0.32	0.34	0.50	0.60	–						
RC	0.73	0.58	0.68	0.49	0.37	0.47	0.44	0.39	0.56	0.59	0.74	–					
RF	0.66	0.60	0.54	0.48	0.56	0.46	0.38	0.29	0.47	0.43	0.55	0.59	–				
MC	0.75	0.54	0.66	0.51	0.58	0.40	0.38	0.34	0.43	0.48	0.57	0.60	0.67	–			
MP	0.80	0.63	0.77	0.58	0.46	0.53	0.47	0.39	0.56	0.58	0.56	0.59	0.57	0.76	–		
WE	0.70	0.44	0.57	0.49	0.49	0.47	0.44	0.37	0.43	0.49	0.69	0.74	0.67	0.63	0.53	–	
BA	0.84	0.68	0.72	0.60	0.60	0.54	0.46	0.41	0.60	0.63	0.81	0.82	0.87	0.87	0.78	0.82	–

Note. GIA = General Intellectual Ability, Gc = Comprehension-Knowledge, Gf = Fluid Reasoning, GSM = Short-Term Working Memory, Gs = Cognitive Processing Speed, Ga = Auditory Processing, Glr = Long-Term Retrieval, Gv = Visual-Spatial Thinking, OE = Oral Expression, LC = Listening Comprehension, BR = Basic Reading Skills, RC = Reading Comprehension, RF = Reading Fluency, MC = Math Calculation Skills, MP = Math Problem Solving, WE = Written Expression, BA = Broad Achievement.