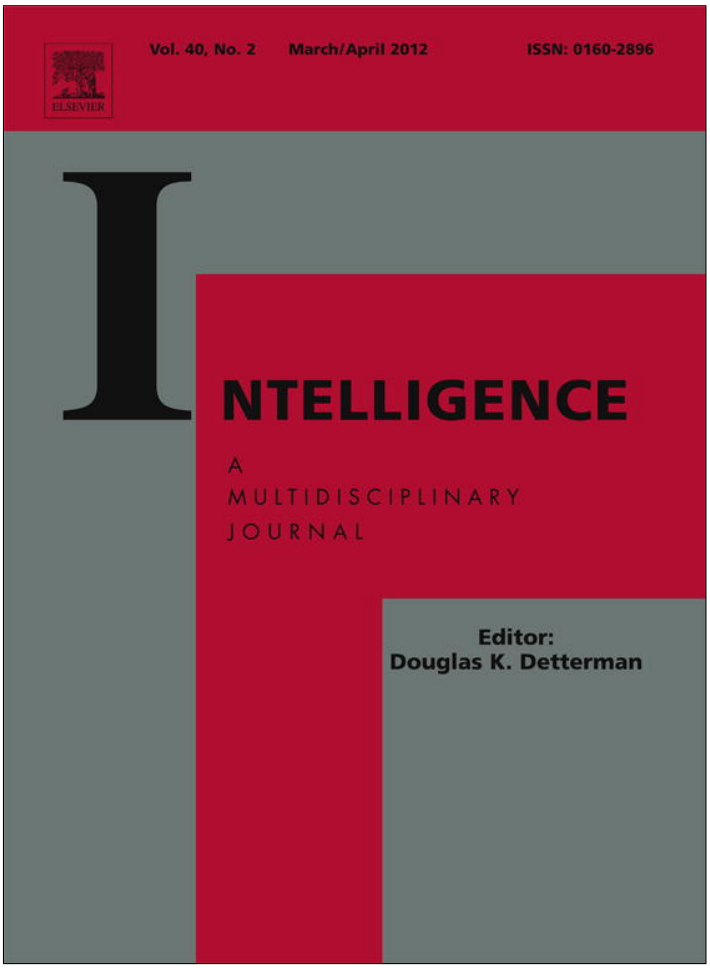
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Intelligence 40 (2012) 123–138

Contents lists available at SciVerse ScienceDirect

Intelligence

Are cognitive *g* and academic achievement *g* one and the same *g*? An exploration on the Woodcock–Johnson and Kaufman tests

Scott Barry Kaufman [a,*], Matthew R. Reynolds [b], Xin Liu [c], Alan S. Kaufman [d], Kevin S. McGrew [e,f]

[a] *Psychology Department, New York University, USA*
[b] *The University of Kansas, USA*
[c] *NCS Pearson, Clinical Assessment, Bloomington, MN, USA*
[d] *Yale Child Study Center, Yale University School of Medicine, New Haven, CT, USA*
[e] *Woodcock-Muñoz Foundation, USA*
[f] *University of Minnesota, Minneapolis, MN, USA*

ARTICLE INFO

Article history:
Received 13 January 2011
Received in revised form 14 January 2012
Accepted 14 January 2012
Available online 7 February 2012

Keywords:
Intelligence
General cognitive ability
Academic achievement
Cattell–Horn–Carroll (CHC) theory
Kaufman Assessment Battery for Children — Second Edition (KABC-II)
Kaufman Test of Educational Achievement — Second Edition (KTEA-II)
Second Edition Comprehensive Form
Woodcock-Johnson
Third Edition
Woodcock-Johnson III
General Intelligence (g)

ABSTRACT

We examined the degree to which the conventional notion of *g* associated with IQ tests and general cognitive ability tests (*COG-g*) relate to the general ability that underlies tests of reading, math, and writing achievement (*ACH-g*). Two large, nationally representative data sets and two independent individually-administered set of test batteries were analyzed using confirmatory factor analysis procedures: (a) the Kaufman-II sample ($N = 2520$), organized into six age groups between 4–5 and 16–19 years, tested on both the Kaufman Assessment Battery for Children-2nd ed. (KABC-II) and the Kaufman Test of Educational Achievement-2nd ed. (KTEA-II) Comprehensive Form; and (b) the WJ III sample ($N = 4969$), organized into four age groups between 5–6 and 14–19 years, tested on both the Cognitive and Achievement batteries of the Woodcock–Johnson-3rd ed. (WJ III). Second-order latent factor models were used to model the test scores. Multi-group confirmatory factor analysis was used to investigate factor loading invariance across the age groups. In general, invariance was tenable, which allowed for valid comparisons of second-order *COG-g* and *ACH-g* factor variance/covariances and correlations across age. Although *COG-g* and *ACH-g* were not isomorphic, they correlated substantially, with an overall mean correlation coefficient of .83, and with the correlations generally increasing with age (ranging from .77 to .94). The nature of the relation between *COG-g* and *ACH-g* was explored and the best measures of *COG-g* were examined.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

One of the central purposes of intelligence testing, dating back to Alfred Binet, is to predict educational achievement (Binet & Simon, 1916). Research has shown a moderate to strong relation between general cognitive ability (*g*) and school grades, ranging from 0.40 to 0.70 (Mackintosh, 1998). Jensen (1998) noted that the median validity coefficient of IQ for educational variables is about .50, with the spread of the validity coefficients varying considerably depending on the variability of the group (the coefficient being higher for those nearer to the lower end of the educational ladder).

Even though the IQ-achievement correlations tend to be moderate to high, typically there is about 50 to 75% of the variance in academic achievement unaccounted for by measures of cognitive ability alone (Rohde & Thompson, 2007). While some of the unaccounted for variance is measurement error, there are certainly many factors besides *g* that systematically play a role in determining school grades, including domain-specific aptitudes (e.g., Gustaffson & Balke, 1993), other student characteristics (e.g., social-emotional functioning, behavior, motivation, grit, affect, metacognition, specific

* Corresponding author at: Department of Psychology, New York University, 6 Washington Place, room 158, New York, NY, 10003, USA.
E-mail address: scott.barry.kaufman@nyu.edu (S.B. Kaufman).

0160-2896/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2012.01.009

Author's personal copy

psychomotor skills), classroom practices, design and delivery of curriculum and instruction, school demographics, climate, politics and practices, home and community environments, and, indirectly, state and school district organization and governance (Wang, Haertel, & Walberg, 1993). Indeed, in his seminal review of the link between cognitive ability and academic achievement, Ceci (1991) showed that the relation between IQ and academic achievement is substantially influenced by the context of the learning environment. School learning is clearly the result of the interaction of a complex set of proximal and distal student and environmental characteristics.

The correlation between IQ and academic achievement is typically higher when looking at standardized tests of achievement rather than grades in school (Mackintosh, 1998), because school performance is more strongly tied to the curriculum, student effort, teacher competency, and other "irrelevant" variables. Research has shown that IQ and achievement test scores have yielded correlation coefficients that usually range from the mid-.60s to mid-.70s (Naglieri & Bornstein, 2003) and sometimes reach the mid-.80s (The Psychological Corporation, 2003, Table 5.15), based on a variety of individually-administered IQ and achievement tests.

Although such standardized achievement tests certainly do not guarantee that all students will be on equal footing in terms of their learning context, the tests do minimize potentially confounding variables such as idiosyncratic teacher grading styles and teacher perceptions. Another benefit of using standardized achievement tests in assessing the relation between intelligence and academic achievement is that factor analysis can be applied to multiple tests, allowing for an assessment of common variance across the tests and minimizing error variance,[1] which can contribute to a less-than-accurate correlation with *g* (e.g., Watkins, Lei, & Canivez, 2007). Lastly, *individually* administered tests eliminate confounds related to group administered tests in that a skilled examiner may minimize examinee related variance related to motivation, comprehension of instructions, effort, attention and so forth, especially for younger school-aged children who are not used to standardized test situations (Kaufman, 1979).

In a recent study, multiple measures of *g* were used to predict group administered standardized national public examination results across 25 academic subjects (Deary, Strand, Smith, & Fernandes, 2007). The correlation between a latent *g* factor measured at age 11 and a latent general educational achievement factor measured at age 16 was 0.81. A longitudinal coefficient as substantial as .81 is remarkably high and suggests that the latent *g* and academic achievement constructs might approach identity when assessed concurrently. Other studies have found that the average IQ of a nation is highly correlated with the academic achievement of that nation (Lynn & Meisenberg, 2010; Rindermann, 2007). Looking at differences in IQ across 86 countries, Lynn and Meisenberg (2010) found a correlation of .92 between a nation's measured IQ and the educational attainment of school students in math, science and reading comprehension. Correcting for attenuation, they found a correlation of 1.0. All of these results are consistent with Spearman (1904), who suggested over 100 years ago that the general factor extracted from a performance battery based on school grades would be almost perfectly correlated with general intelligence.

A related issue is the extent to which the relation between cognitive-*g* (*COG-g*) and academic achievement-*g* (*ACH-g*) varies as a function of age. Studies that have assessed the relation at various ages have reported a declining association with age, which has been attributed to dropout effects, increasing restriction of range, variability in educational complexity, and the increasing role of dispositional factors (e.g., motivation and persistence) (Jensen, 1998). None of these studies, however: (a) estimated the correlations between a latent *g* factor from individually administered tests of cognitive ability and a latent general factor extracted from standardized achievement tests, and (b) tested for factorial invariance across the different age groups. Gustaffson and Balke (1993) investigated the relationship between a latent cognitive ability factor and a latent school achievement factor comprising course grades in 17 different subjects. They found that *COG-g* explained a substantial amount (40%) of the variance in *ACH-g*. Similarly, among a sample of German students, Rindermann and Neubauer (2004) found a correlation of .63 between *COG-g* and an *ACH-g* consisting of school grades. In both of these studies, standardized tests of *ACH-g* were not administered.

Among recent studies that have included standardized measures of academic achievement (Deary et al., 2007; Spinks et al., 2007), the measures of academic achievement were group administered, and the correlations between a latent general cognitive ability factor and latent educational achievement scores were assessed within longitudinal designs, not via cross-sectional methodology. Although these studies provided important insights into the possible causal relations linking *COG-g* to *ACH-g*, they did not directly measure the degree to which the two types of *g* are the same or different, which is best examined when cognitive and achievement tests are administered concurrently. Also, they did not assess whether the *COG-g* and *ACH-g* correlation differs as a function of chronological age.[2]

To cast light on these issues, the current study aimed: (a) to assess the relation between a latent *g*-factor extracted from a battery of individually administered cognitive ability tests (reflecting *COG-g*) and a latent academic achievement factor extracted from a battery of academic achievement tests (reflecting *ACH-g*) using large, nationally representative samples; (b) to test the equivalence of second-order *COG-g* and *ACH-g* latent factor correlations across a wide range of age groups; (c) to cross-validate these findings with a second large battery of cognitive and achievement tests, normed on an independent, nationally representative sample; and (d) to interpret all results, from the perspective of Cattell–Horn–Carroll (CHC) theory (see Carroll, 1993; Horn & Noll, 1997; McGrew, 2005, 2009).

The CHC model represents a merger of the Horn–Cattell Gf–Gc theory (Horn & Cattell, 1966; Horn & Noll, 1997) and Carroll's (1993) three-tiered hierarchical organization of

[1] If latent variable structural equation model-based factor methods are used, the relations between cognitive and achievement latent variables can be estimated that are purified or purged of measurement error.

[2] It is possible to investigate developmental age effects if more complex longitudinal test–retest designs are used that include developmental time-lag components (see McArdle & Woodcock, 1997, for an example).

Author's personal copy

human abilities. Historically, both theories placed a key emphasis on an array of 8–10 cognitive Broad Abilities. The merged, contemporary CHC theory identifies 10 such Broad Abilities, for example, Crystallized Knowledge (Gc), Fluid Reasoning (Gf), Short-term Retrieval (Gsm) and Processing Speed (Gs). CHC theory is particularly pertinent as a theoretical foundation for the present study because the 10 Broad Abilities include eight that are readily identifiable as cognitive and two that fit naturally into the academic achievement domain: Grw (Reading & Writing) and Gq (Quantitative Knowledge). Furthermore, CHC theory is the main theoretical basis for nearly all current individually-administered tests of cognitive ability (Kaufman, DeYoung Gray, Brown & Mackintosh, 2009), including both sets of instruments used to address the *COG-g* and *ACH-g* relation in the present study.

2. Method

2.1. Participants

Kaufman-II. The "Kaufman sample" included the conorming population of the *Kaufman Test of Educational Achievement-Second Edition, Comprehensive Form* (KTEA-II;Kaufman & Kaufman, 2004b) and the *Kaufman Assessment Battery for Children-Second Edition* (KABC-II; Kaufman & Kaufman, 2004a). This sample comprised a total of 2520 students included in the KABC-II norm sample and also in either the age-norm or the grade-norm sample of the KTEA-II. About half the sample was tested on KTEA-II Form A ($n=1227$) and the other half on KTEA-II Form B ($n=1293$). Both KTEA-II and KABC-II norm samples were stratified to be close to population percentages for gender, ethnicity, parental education, and geographic region, according to US population data from the Current Population Survey, March 2001. Analyses were conducted on six age groups: 4–5 ($n=295$), 6 ($n=198$), 7–9 ($n=565$), 10–12 ($n=577$), 13–15 ($n=511$), and 16–19 ($n=374$). Each age sample matched Census figures with reasonable accuracy. The total sample comprised 1257 (49.9%) females and 1263 (50.1%) males; 1569 (62.2%) Caucasians, 375 (14.9%) African Americans, 445 (17.7%) Hispanics, and 131 (5.2%) "Others" (e.g., American Indians, Alaska Natives, Asian Americans, and Pacific Islanders); 362 (14.4%) had parents who completed less than 12 years of formal schooling, 818 (32.5%) had parents who graduated high school, 759 (30.1%) had parents who completed 1–3 years of college, and 581 (23.0%) had parents who graduated college; and 350 (13.9%) lived in the Northeast, 662 (26.3%) lived in the North Central region, 875 (34.7%) lived in the South, and 633 (25.1%) lived in the West.

WJ III. The "WJ III sample" comprised $N=4969$ individuals tested on the *Woodcock–Johnson III* (WJ III; Woodcock, 2001). The sample was drawn from the nationally representative WJ III standardization sample (see McGrew & Woodcock, 2001). It was constructed using a three-stage stratified sampling plan that controlled for 10 individual (e.g., race, gender, educational level, occupational status) and community (e.g., community size, community SES) variables as per the United States Census projection for the year 2000. Analyses were conducted on four age groups: 5 to 6 ($n=639$), 7 to 8 ($n=720$), 9 to13 ($n=1995$), 14 to 19 ($n=1615$). Within each age group, two randomly divided subsamples were used so that the analysis would consist of models calibrated in one sample and then cross-validated in another sample at each respective age grouping.

2.2. Measures

Kaufman-II. The cognitive and achievement tests used in the Kaufman sample were from the KABC-II and KTEA-II test batteries. Descriptions of KTEA-II and KABC-II subtests are presented in the test manuals (Kaufman & Kaufman, 2004a, Table 1.2, Kaufman & Kaufman, 2004b, Table 1.1) and are available in a number of assessment texts (Kaufman, Lichtenberger, Fletcher-Janzen, & Kaufman, 2005; Lichtenberger & Breaux, 2010). Estimates of reliability and evidence of validity for all KTEA-II and KABC-II scores are reported by Kaufman and Kaufman (2004a, 2004b), Kaufman et al. (2005), and Lichtenberger and Breaux (2010); the pros and cons of the instruments, as evaluated by independent reviewers, are also summarized by Kaufman et al. (2005), Lichtenberger and Breaux (2010), and Reynolds, Keith, Fine, Fisher, and Low (2007).

The KTEA-II is an individually administered measure of academic achievement for individuals ages 4.5 through 25 years. It includes 14 subtests, nine of which measure (a) mathematics (Math Computation, Math Concepts & Applications), (b) reading (Letter & Word Recognition, Reading Comprehension, Nonsense Word Decoding), (c) reading fluency (Word Recognition Fluency, Decoding Fluency), and (d) written language (Written Expression, Spelling). In terms of the CHC taxonomy (McGrew, 2005, 2009), the reading, writing, and spelling tests are associated with Grw, and the math tests with the Gq factor. The other five KTEA-II subtests did not fit into reading-writing (Grw) or math (Gq) domains and were best classified as measuring Gc (Listening Comprehension, Oral Expression; Associational Fluency) and Glr (Naming Facility/RAN; Flanagan, Ortiz, & Alfonso, 2007; Flanagan, Ortiz, & Alfonso, in press; Kaufman et al., 2005). These KTEA-II subtests were, therefore, included in this study as measures of *COG-g*, although not all of these subtests were administered to all age groups.

The KABC-II is a popular individually administered measure of intelligence. The scoring structure of the KABC-II includes five CHC broad ability composites: Gc, Glr, Gf, Gsm, and Visual Processing (Gv). A total of 18 KABC-II subtests were included in this study, although not all of the tests were available for each age group. The KABC-II and KTEA-II subtests were organized into CHC broad ability factors, which, in turn were regressed onto cognitive and achievement second-order *g* factors (see Table 1).

WJ III. The cognitive and achievement measures used in the WJ III sample were from the WJ III Cognitive, Achievement, and Diagnostic Supplement test batteries test batteries (Woodcock, McGrew, Mather, & Schrank, 2003). A total of 40 tests were used. The development, standardization, and psychometric properties of the WJ-III battery have generally been evaluated favorably by independent reviewers (Bradley-Johnson, Morgan, & Nutkins, 2004; Cizek, 2003; Sandoval, 2003; Sares, 2005; Thompson, 2005). CHC theory (McGrew, 2005, 2009) was used to organize the tests by CHC broad ability factors. These classifications are shown in Table 2. The variables used were the same across all age groups except in

Author's personal copy

Table 1
Kaufman organization of subtests into broad abilities.

COG-g					*ACH-g*	
Gc	Gv	Gf	Glr	Gsm	Grw	Gq
Verbal knowledge	Triangles	Story completion	Atlantis	Word order	Letter–word recognition	Math concepts and applications
Expressive vocabulary	Block counting	*Pattern reasoning*	Atlantis-delayed	Number recall	*Written expression*	Math calculation
Riddles	Rover	*Hand movements*	Rebus	*Hand movements*	Nonsense word decoding	
Oral expression	*Pattern reasoning*	*Conceptual thinking*	Rebus-delayed		Spelling	
Listening comprehension	*Gestalt closure*		Naming facility		*Reading comprehension*	
Written expression	Face recognition				Decoding fluency	
Reading comprehension	*Conceptual thinking*				Nonsense word decoding fluency	
Gestalt closure						
Associational fluency						

Notes. Italics indicated subtest was cross-loaded.

Table 2
WJ-III organization of tests into broad abilities.

			COG-g				*ACH-g*	
Gc	Gv	Gf	Glr	Gsm	Ga	Gs	Grw	Gq
Verbal comp	Spatial relations	Analysis-synthesis	Memory for names	Memory for words	Auditory attention	Decision speed	Spelling	Math applied problems
General info	Block rotation	Concept formation	Memory for names delayed	*Memory for sentences*	Sound patterns	Visual-matching	Writing sample	Math calculation
Oral comp	Visual closure	*Numerical reasoning*	Picture recognition	Auditory working memory	Incomplete words	*Cross-out*	Letter–word identify	*Math fluency*
Story recall	*Cross-out*		Visual–auditory memory	Numbers reversed	Sound blending	*Rapid naming*	Passage comp	*Numerical reasoning*
Memory for sentences	Planning		Visual–auditory memory delayed			*Writing fluency*	Word attack	
Rapid naming						*Math fluency*	*Math fluency*	
Retrieval fluency							Reading fluency	
							Editing	
							writing fluency	

Notes. Italics indicated test was cross-loaded.

Author's personal copy

the youngest sample (ages 5–6) where the Writing and Reading fluency tests and Editing were not applicable.

2.3. Analytic strategy

Preliminary models. Second-order CFA models were developed for each age group and each test battery. Initial models were based on information from the test manuals, CHC theory, and prior research. The goal was to develop models that were acceptable from theoretical and statistical standpoints. Once acceptable models were estimated in each age group, multi-group confirmatory factor models were estimated. General specifications for the Kaufman-II and WJ III models are presented next with more detailed information presented in the Results section.

Kaufman-II. To estimate the correlation between KABC-II *COG-g* and KTEA-II *ACH-g* across ages, a model with correlated second-order *COG-g* and *ACH-g* factors was proposed in each age group; this model was based on theoretical, empirical, and clinical considerations (Kaufman et al., 2005; Lichtenberger & Breaux, 2010; Reynolds et al., 2007). Because not all subtests were administered at all ages, the number of factor indicators varied by age. Factor to indicators specifications are shown in Table 1.

The same 18 cognitive and 10 achievement subtests were administered to participants across the 10–12, 13–15, and 16–19 age groups. The initial model for these age groups included seven first-order CHC common factors (Gf, Gc, Glr, Gsm, Gv, Grw, Gq), with these first-order factors regressed on two correlated second-order latent factors. Grw and Gq factors were regressed on a second-order *ACH-g* factor and Gf, Gc, Glr, Gsm, and Gv factors were regressed on a second-order *COG-g* factor.

Participants in the 7–9 age group were administered the same 18 cognitive subtests as the older age groups; however, they were given two fewer KTEA-II achievement subtests (See Fig. 1). These fewer subtests, however, did not influence the number of factors (See Figs. 1 and 2).

Although there were few differences in factor indicators in the 7–9, 10–12, 13–15, 16–19 age groups, there were several differences in the youngest two age groups. In the 6 year-old age group, Word Recognition Fluency, Decoding Fluency, Spelling, Nonsense Word Decoding, and Reading Comprehension—all indicators of the Grw factor—are not age appropriate and were not administered. Grw was thus indicated by two subtests at age 6 (viz., Letter & Word Recognition and Written Expression). In addition to these KTEA-II subtests, Atlantis Delayed and Rebus Delayed were not administered. Two subtests not administered to the older age groups, Conceptual Thinking and Gestalt Closure, were administered to children at age 6 (and also ages 4–5). Conceptual Thinking was specified to load on both Gf and Gv factors for 6-year-olds and Gestalt Closure was specified to load on both Gv and Gc.

Last, the age 4–5 model included several departures from the models in the other age groups. First, there was not a Gf factor. Gv and Gf were not differentiated for this age group in the KABC-II (see Kaufman & Kaufman, 2004a). Gv was indicated by Gestalt Closure (cross-loaded on Gc), Conceptual Thinking, Face Recognition, and Triangles. Three KTEA-II achievement subtests were administered at these ages: Math Concepts & Applications, Written Expression, and Letter & Word Recognition. Gq was indicated by the single subtest, Math Concepts & Applications. To identify this factor the residual variance was fixed to 12.97 (Keith, 2006), which was calculated by subtracting the internal reliability estimate (.93) from 1.0, and then multiplying this value (.07) by the variance (185.29).

There were some additional modifications to the specifications outlined above. In addition to second-order factor correlations, some first-order factor unique variances between the cognitive and achievement factors were correlated. These correlations were freed because previous research has supported the influence of broad ability CHC factors on achievement above and beyond the influence of *g* (e.g., Benson, 2008; Keith, 1999). In addition, some KABC-II subtests were found to cross-load on two factors in previous research (Reynolds et al., 2007). We allowed these subtests (Hand Movements on Gf and Gsm; Pattern Reasoning on Gf and Gv; Gestalt Closure, when administered, on Gc and Gv) to cross-load on two cognitive factors across the age groups. Both Written Expression and Reading Comprehension cross-loaded on Gc. Delayed subtest residuals correlated freely with their respective tests from the initial measurement (e.g., Atlantis with Atlantis Delayed). Lastly, three additional measurement residual correlations were estimated freely because they represented common content or method: Word Recognition Fluency and Decoding Fluency; Decoding Fluency and Nonsense Word Decoding; and Associational Fluency and Naming Facility. Additional age-group specific adjustments are discussed in the Results section (model calibration sample, See Fig. 2).

WJ III. Initial models were developed using Subsample 1 (model calibration sample). The same model specifications were utilized when using Subsample 2 (model cross-validation sample) data for analysis. The general *COG-g* model was specified to include seven first-order CHC latent factors (Gf, Gc, Glr, Ga, Gv, Gsm, Gs) which were, in turn, regressed on a second-order *COG-g* factor. The general *ACH-g* model used across ages specified two broad first-order latent achievement factors (Gq, Grw), with those factors regressed on a second-order *ACH-g* factor. The *COG-g* and *ACH-g* second-order factors were correlated. Some correlated residuals and cross-loadings were also included across all age groups. Delayed recall test residuals were correlated with each other and with their corresponding residual from the initial measurement. In addition, the three achievement fluency test residuals were correlated when they were administered. Five cognitive tests had cross-loadings, which were allowed across ages. Retrieval Fluency loaded on Gc and Gs, Numerical Reasoning on Gf and Gq, Memory for Sentences on Gc and Gsm, Rapid Naming on Gc and Gs, and Cross Out on Gv and Gs. In addition, Writing (Grw) and Math (Gq) Fluency tests were cross-loaded on the Gs factor. The organization of tests by broad abilities is shown in Table 2.

There were some age-group specific correlations between the cognitive and achievement first-order unique variances. Moreover, there were some age-group specific measurement residual correlations, as explained in the Results section.

Multi-group models. The primary purpose of this research was to test the equivalence of *g* factor covariances (correlations) and variances across age levels. Valid comparisons of

Author's personal copy

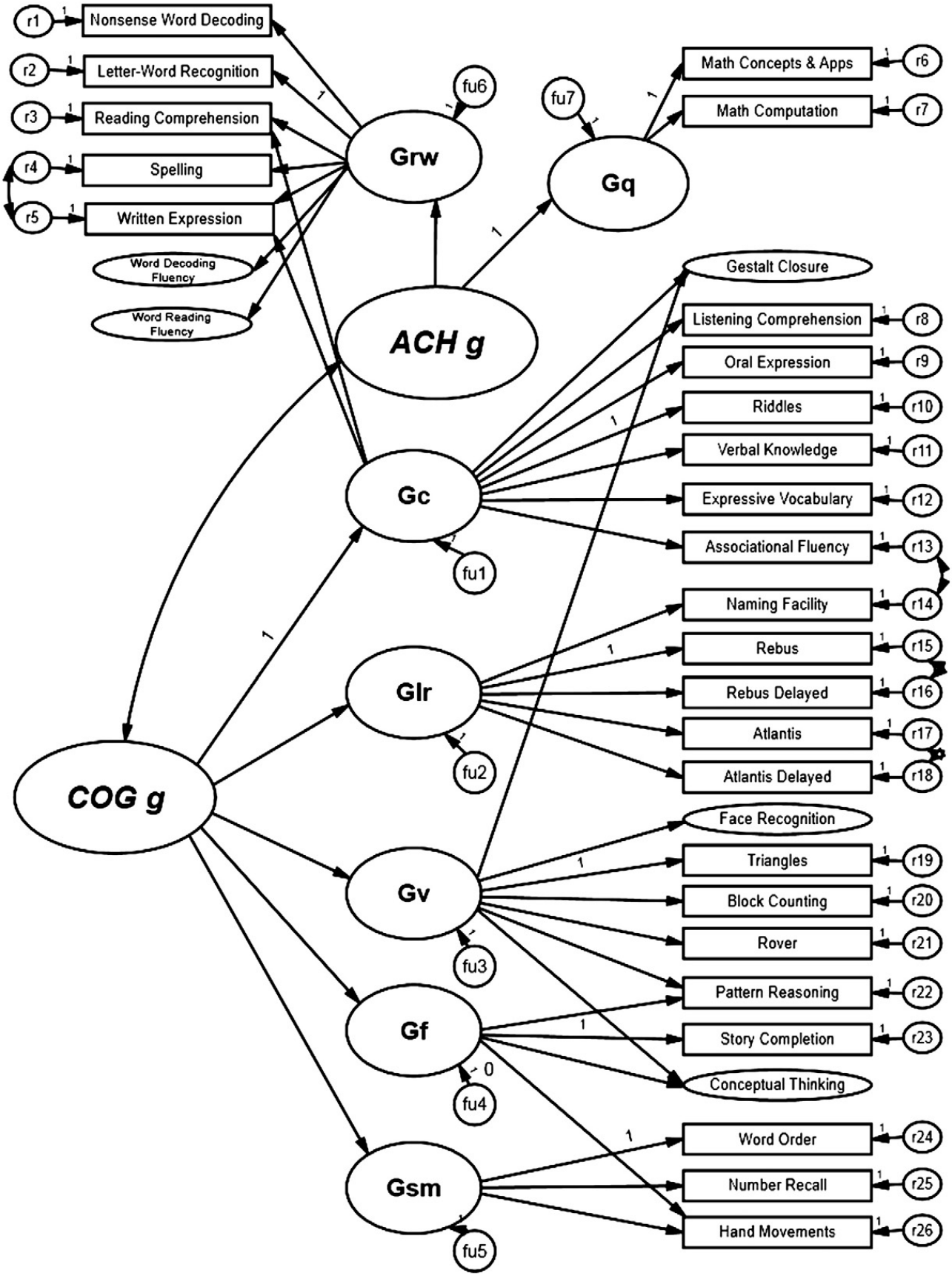


Fig. 1. Kaufman *COG-g/ACH-g* second-order model for children aged 7–9. *Note.* Second-order factor model for the Kaufman-II data, with two correlated second-order *g* factors. The correlation between the Glr–Grw uniqueness was not included in the figure.

Author's personal copy

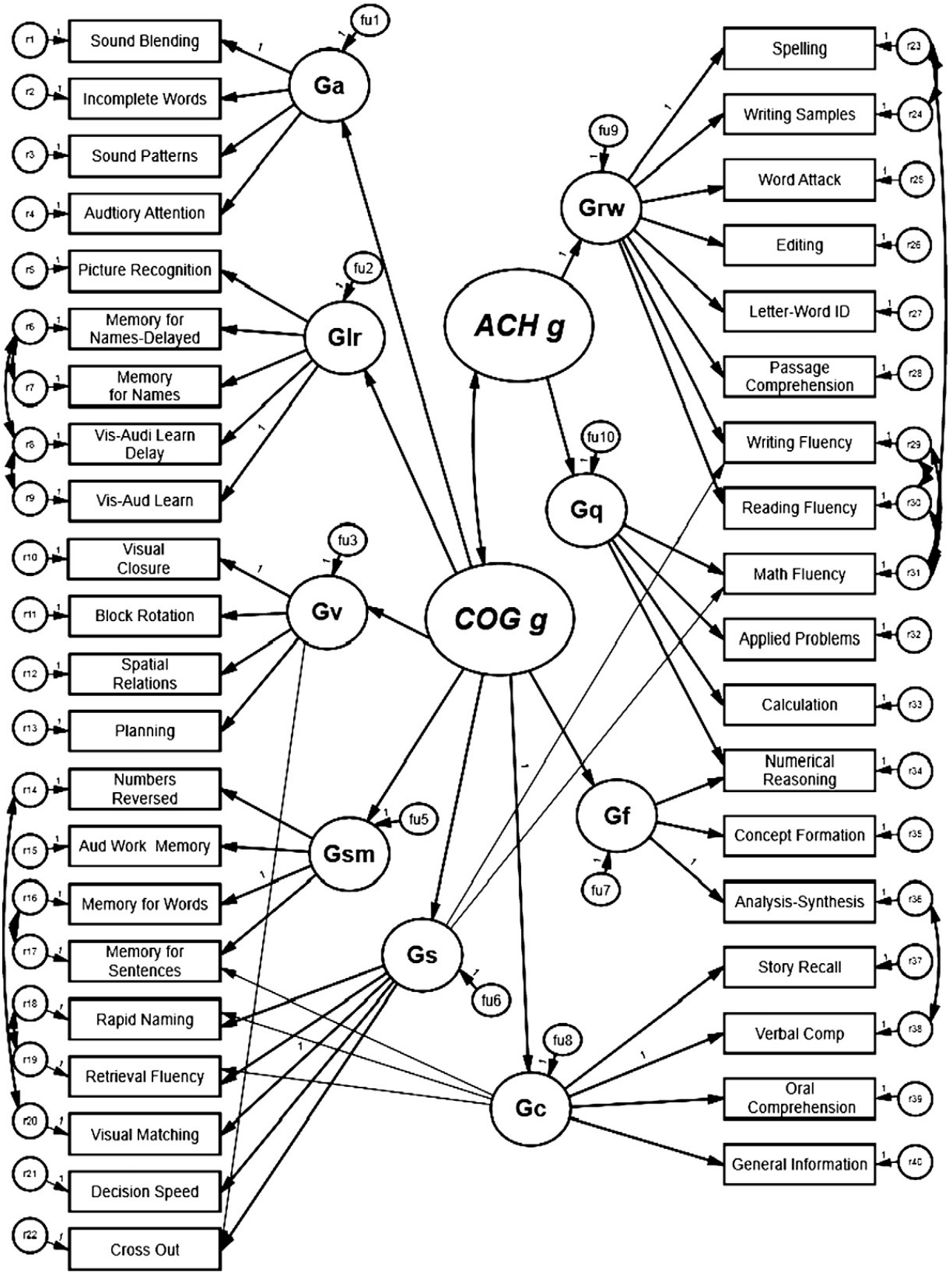


Fig. 2. Woodcock Johnson-III *COG-g/ACH-g* second-order model for children aged 9–13. *Note.* Second-order factor model for the Woodcock Johnson-III data, with two correlated second-order *g* factors. Correlations between the Grw–Gc, Gf–Gq, Gs–Gq, and Gs–Grw uniqueness were not included in the figure.

Author's personal copy

factor variances and covariances required that factors had the same meaning across age. This assumption was tested empirically via tests of factorial invariance. Higher-order models were used in this research; thus, age-invariance of both first- and second-order factor loadings was evaluated in a set of nested, multi-group models across age groups. Invariant first- and second-order loadings allowed for valid quantitative comparisons of factor variances and covariances across the age groups (Gregorich, 2006).

Substantive questions were addressed using multi-group analysis in which like factor variances and covariances (correlations), in addition to invariant first- and second-order factor loadings, were constrained across age groups. Before factor covariances were compared, the factor variances were tested for equivalence across age. If equality of factor variances was tenable, these equality constraints were maintained and the equivalence of *COG-g/ACH-g* factor covariances (correlations if factor variances were equal) across age was tested.

One issue that arose while testing for factorial invariance was that some subtests were not administered at all ages. Consequently some factor indicators varied across age groups. If a subtest indicator of a common factor was missing in one age group, it was specified as a latent variable with zero variance (i.e., missing data). This specification, along with common indicators within each factor available across ages, allowed for us to proceed with invariance tests in the multi-group models despite some differences in factor indicators across age (see Keith, Low, Reynolds, Patel, & Ridley, 2010; Keith & Reynolds, 2012; McArdle, 1994; Reynolds et al., 2007).

2.4. Model fit

Values for model chi squared (χ^2), root-mean square error of approximation (RMSEA; Steiger & Lind, 1980), and comparative fit index (CFI; Bentler, 1990) were reported and used for evaluation of individual models. The standardized root mean square residual (SRMR; Hu & Bentler, 1999) was reported when available.

There is no definitive method for evaluating the fit for tests of factorial invariance. To compare the fit of hypothesized models while testing for invariance, we used the likelihood ratio test (Bentler & Bonett, 1980) and ΔCFI. In a simulation study, Cheung and Rensvold (2002) found that ΔCFI > .01 was considered meaningful change, and that the index was not overly sensitive to small errors of approximation when applied to first-order factor models. Some may consider ΔCFI as a more liberal criterion. Alternatively, the likelihood ratio test is often considered to be overpowered at detecting unimportant and tiny differences in fit when there are a large number of constraints and large sample size. The more liberal criterion was given more weight for measurement invariance models due to the complexity of the model, number of constraints, and large sample size. The likelihood ratio test was used when a test of one or very few specific parameters related to structural level substantive hypotheses was required. In addition, RMSEA values and SRMR (when available) were reported for all models. Steiger's (1998) multi-group correction (RMSEA × √# of groups) was applied to RMSEA values. Some general guidelines for changes in these index values for demonstrating factor loading invariance have been provided in previous research: ΔCFI < .01; ΔRMSEA < .015; ΔSRMR < .03 (Chen, 2007).

3. Results

3.1. Descriptive statistics

Age-standardized scores were used for the analysis of Kaufman-II data. Because the scores were from the norming sample, the means and standard deviations for all of the age-groups are close to the designated sample mean and standard deviation of 10 and 3, respectively for the KABC-II subtests, and 100 and 15 for KTEA-II subtests. There were very small amounts of incomplete data (5% or less) for a very few number of variables within each age group. Little's MCAR tests were performed in each age, none of which was statistically significant at the $p < .01$ level. Regardless, all of the available data were analyzed using maximum likelihood estimation procedures, which required the less strict assumption of missing at random.

In the WJ III sample data, we used covariance matrices that were developed in prior research. The covariance matrices were split into calibration and validation samples (here referred to as Subsamples 1 and 2) in prior research. Descriptive statistics for tests for all ages in the standardization sample are presented in McGrew and Woodcock (2001). Again, because standardization sample data were used, all test mean and standard deviations were close to the designated mean of 100 and standard deviation of 15. Detailed descriptive statistics for all tests in the same age-differentiated calibration and cross-validation samples have been described previously (Floyd, Keith, Taub, & McGrew, 2007; Taub, Floyd, Keith, & McGrew, 2008).

3.2. Kaufman-II analysis

KTEA-II data included scores from parallel Forms A and B. An analysis was conducted to determine whether it was feasible to collapse the Forms into a single data set. A four-factor Grw, Gq, Oral Language, and Oral Fluency model was specified for each Form. A test of strict factorial invariance was performed. The Configural Invariance model fit was acceptable, χ^2 (110) = 1084.77, CFI = .943. Next, a strict factorial invariance model was imposed, χ^2 (145) = 1213.65, CFI = .938, with $\Delta\chi^2$ (35) = 128.88, $p < .001$. The ΔCFI (.005) was negligible. Given the sensitivity of the likelihood ratio test to sample size, it was deemed appropriate to collapse KTEA-II Forms A and B for the purpose of this research.

Once Form A and Form B KTEA-II data were merged, acceptable second-order models were developed in each age group. In the *COG-g* model, the Gf factor unique variance was not statistically significantly different from zero in any age group; this variance was constrained to zero in each age group. This finding of a perfect correlation between Gf and *g* is not uncommon (Gustaffson, 1984; Reynolds & Keith, 2007).

The fit of the models for ages 10–12, 13–15, and 16–19 were acceptable (Table 3). Some first-order cognitive and achievement factor unique variances were correlated freely in each of these three age groups: Gc with Gq ($r = .50$, .48, .42, respectively) and Gv with Gq ($r = .64$, .77, .41,

Author's personal copy

Table 3
Age group specific fit of the Kaufman-II models.

Age-group model	$\chi^2(df)$	CFI	RMSEA
4–5 year-olds	201.78(127)	.963	.045
6 year-olds	344.94(216)	.945	.055
7–9 year-olds	560.32(283)	.966	.042
10–12 year-olds	794.37(329)	.955	.050
13–15 year-olds	807.55(329)	.953	.053
16–19 year-olds	679.80(329)	.947	.053

respectively). In addition, Glr was correlated with Grw in the 10–12 ($r=.21$) and 13–15 ($r=.25$) age groups, and Gc with Grw in the 13–15 ($r=.50$) and 16–19 ($r=.48$) age groups. Lastly, Gsm correlated with Gq in the 10–12 ($r=.22$) age group, and with Grw in the 16–19 ($r=.25$) age group. These correlations represent the relations between specific cognitive and achievement areas.

A model was also estimated in each of the three remaining age groups, and each had acceptable fit (Table 3). Glr and Grw unique variances were correlated freely in the models for the 7–9 ($r=.45$) and 6 ($r=.66$) age groups.

3.3. Multi-group COG-g/ACH-g models

Loading invariance. The Configural model fit well (Table 4). According to the likelihood ratio test, model fit degraded when the first-order loadings were constrained equal. ΔCFI, however, was negligible[3] and the RMSEA did not change. Model fit did not degrade when constraints were added to corresponding second-order factor loadings (Table 4). The constraints added in this Second-Order Factor loading model were maintained in subsequent models.

Factor variance/covariance equivalence. The next set of tests focused on substantive questions. Model fit comparisons are shown in Table 4. The dispersion of corresponding *g* factors was the same across age groups as indicated by the lack of degradation in model fit when they were fixed to be equal. These constraints were maintained and we tested for factor correlation equivalence across age.

The application of equality constraints on the *COG-g/ACH-g* correlations resulted in degradation in model fit. The primary source of misfit was due to the correlation between the *COG-g* and *ACH-g* factors in the 9–10 age group. When this constraint was released, there was a statistically significant improvement in model fit, $\Delta\chi^2(1)=19.26$, $p<.001$. The factors were less differentiated in this age group ($r=.94$). This model, however, still demonstrated a statistically significant degradation compared to the *ACH-g* variance model (Table 4). We released the equality constraint from the correlation between *COG-g* and *ACH-g* factors for the age 4–5 group, which resulted in a statistically significant improvement in model fit, $\Delta\chi^2(1)=7.41$, $p=.007$. The factors were more differentiated in this group ($r=.77$). Moreover, the model fit did not differ statistically significantly from the fit of the *ACH-g* model (Table 4). The correlation between factors, constrained equal across the remaining four age groups, was .86.

3.4. WJ III analysis

Although the general models were described previously, there were some slight sample specific correlated measurement residuals across the age groups. Because the investigation of all potential specific factors that may arise in each sample was not a purpose of this research, and because we wanted to maintain *g* factors of the same order, a few measurement residual correlations were allowed in each age group. These correlations were generally trivial, and they did not interfere with an assumption of configural invariance for the multi-group models so to conserve space all of these are not reported.

First, an acceptable model was specified for the 14–19 age group. Model fit was reasonably acceptable, though the CFI was less than desired (See Table 5). Correlated factor unique variances included Gf with Gq ($r=.60$) and Gc with Gq ($r=.29$). The 9–13 age group model fit well. Four first-order unique variances were correlated: Gc with Grw ($r=.62$), Gf with Gq ($r=.35$), and Gs with Gq ($r=.34$) and Grw ($r=.38$). Model fit for the 7–8 year-old age group model was acceptable. Three first-order factor unique variances were correlated, including Gs with Gq ($r=.43$) and Grw ($r=.47$), and Gc with Grw ($r=.34$). Lastly, an acceptable model for the 5–6 age group was developed. The first-order Gf unique variance correlated with the first-order Gq unique variance ($r=.72$).

3.5. Multi-group COG-g/ACH-g models

Loading invariance. In Subsample 1 the models with first- and second-order factor loadings fixed to equality demonstrated degradation in fit using the likelihood ratio test, but the ΔCFI, ΔRMSEA, and ΔSRMR were not considered substantial or important (Table 6). Using the models specified in Subsample 1, the same procedure was applied to Subsample 2. Although the model fit indexes suggested slightly worse fit in Subsample 2, the fit was adequate (Table 6).

3.6. Factor variance/covariance equivalence

Subsample 1. No statistically significant degradation in fit was observed when the *COG-g* variances were fixed to be equal across age (Table 6). There was a statistically significant degradation, however, in fit for the *ACH-g* model. The variance for the 7–8 group was statistically significantly different, with *ACH-g* being more heterogeneous in this age group. The three remaining *ACH-g* variances did not differ significantly from each other.

Lastly, factor correlation equivalence was tested across age groups. The variance of the *ACH-g* factor for the 7–8 age group remained unconstrained across the ages, while the remaining corresponding *COG-g* and *ACH-g* variances were constrained equal. The correlations between all the groups, excluding the 7–8 year-old group, were fixed to be equal. There was a statistically significant degradation in model fit. The degradation was due to the correlation in the 14–19 age group. Model fit improved when this equality constraint

[3] We followed up and tried to locate local sources of non-invariance. Freeing these constraints or even deleting the subtests (e.g., Reading comprehension) from the model did not affect the factor correlations or substantive conclusions.

Author's personal copy

Table 4
Tests of invariance for the Kaufman-II models.

Model	$\chi^2(df)$	$\Delta\chi^2(\Delta df)$	p	CFI	RMSEA[a] [a]
1. Configural	3388.90(1613)			.955	.051
2. First-order loadings	3694.09(1715)	305.20(102)	<.001	.950	.051
3. Second-order loadings	3739.86(1739)	45.77(24)	= .005	.949	.051
4. *COG-g* factor variance	3743.53(1744)	3.67(5)	= .598	.949	.051
5. *ACH-g* factor variance	3748.03(1749)	4.50(5)	= .480	.950	.051
6. *COG-g*/*ACH-g* correlation	3779.57(1754)	31.54(5)	<.001	.949	.051
7. *COG-g*/*ACH-g* correlation	3760.31(1753)	12.28(4)	= .015	.949	.051
8. *COG-g*/*ACH-g* correlation	3752.91(1752)	4.87(3)	= .181	.949	.051

Note. [a]RMSEA corrected for multi-groups (Steiger, 1998). Compare models 1–6 with the previous model. Compare models 7 and 8 with model 5.

Table 5
Age group specific models for the WJ III.

Age-group model	$\chi^2(df)$	CFI	RMSEA	SRMR
5–6 year-olds	763.05(602)	.959	.028	.046
7–8 year-olds	1050.51(706)	.945	.035	.043
9–13 year-olds	1251.34(706)	.954	.031	.037
14–19 year-olds	1318.77(707)	.940	.037	.043

was removed, $\Delta\chi^2$ (1) = 10.8, p = .001. The correlations between *COG-g*/*ACH-g* were the same in the 5–6 and 9–13 age groups (r = .80), and these factors were more differentiated than were the *COG-g*/*ACH-g* factor correlation in the oldest age group (r = .87). The correlation between these two factors in the 7–8 age group was .78.

Subsample 2. *COG-g* and *ACH-g* variances were not statistically significantly different from each other. The *COG-g* and *ACH-g* factor correlations were constrained equal across age groups, but there was statistically significant degradation in fit. The correlation between these two factors in the 14–19 age group was larger than the correlation in the other age groups, $\Delta\chi^2$ (1) = 20.2, p < .001. Once again, *COG-g* and *ACH-g* correlations were more differentiated in the younger age groups (r = .77) than they were in the 14–19 age group (r = .88). In both samples the oldest age group contained the least differentiated *COG-g* and *ACH-g* factors and the magnitude of the correlations were similar across samples.

3.7. Overview of COG-g/ACH-g correlations

Table 7 summarizes the coefficients of correlation between *COG-g* and *ACH-g*, by age group, for the Kaufman-II and the two WJ III subsamples. The values range from .77 to .94 across ages; mean coefficient is .86 for the Kaufman-II and .80 for the WJ III, with an overall mean correlation coefficient of .83 and median of .80. Based on the median age for the various age groups on the two sets of test batteries, 10 coefficients are shown for ages 4½–11 in Table 7 and four coefficients are shown for ages 14–17½. Notably, seven of the 10 values for ages 4½–11 are .80 or lower whereas all values for the adolescent samples were greater than .85. Thus, there is a small trend for the *COG-g*/*ACH-g* coefficients to be larger for high school students than for preschool and elementary school children.

Lastly, a few additional tests were performed to determine whether the correlations between *COG-g* and *ACH-g* were statistically significantly different from one. The second-order

Table 6
Tests of invariance for the WJ III models in subsamples 1 and 2.

Model	$\chi^2(df)$	$\Delta\chi^2(\Delta df)$	p	CFI	RMSEA[a]	SRMR
Subsample 1						
1. Configural	4384.1(2721)			.948	.034	.046
2. First-order loadings	4801.0(2828)	416.9(107)	<.001	.939	.036	.069
3. Second-order loadings	4873.1(2849)	72.7(21)	<.001	.937	.036	.073
4. *COG-g* factor variance	4879.1(2852)	6.0(3)	= .112	.937	.036	.075
5. *ACH-g* factor variance	4893.7(2855)	20.6(3)	<.001	.937	.036	.076
6. *ACH-g* factor variance, without 7–8 groups[b]	4879.5(2854)	0.6(2)	= .74	.937	.036	.075
7. *COG-g*/*ACH-g* correlation, with 14–19 free[c]	4880.1(2855)	0.6(1)	= .439	.937	.036	.075
8. *COG-g*/*ACH-g* correlation[c]	4890.9(2856)	11.4(2)	= .003	.937	.036	.076
Subsample 2						
1. Configural model	4695.2(2721)			.936	.036	.046
2. First-order loadings	5061.2(2828)	366.0(107)	<.001	.928	.038	.066
3. Second-order loadings	5124.2(2849)	63.0(21)	<.001	.926	.038	.071
4. *COG-g* factor variance	5128.5(2852)	4.3(3)	= .233	.926	.038	.071
5. *ACH-g* factor variance	5133.8(2855)	5.3(3)	= .150	.926	.038	.072
6. *COG-g*/*ACH-g* correlation	5155.2(2858)	21.3(3)	<.001	.926	.038	.072
7. *COG-g*/*ACH-g* correlation, 14–19 free[d]	5135.0(2857)	1.2(2)	= .552	.926	.038	.072

[a] RMSEA corrected for multi-groups (Steiger, 1998). Compare models 1–5 with the previous model.
[b] Compare with Model 4.
[c] Compare with model 6 in Subsample 1.
[d] Compare with model 5 in Subsample 2.

Author's personal copy

Table 7
Correlations between *COG-g* and ACH-g on the Kaufman-II and WJ III, By Age.

Age group			
(Median age)	Kaufman-II	WJ III # 1	WJ III # 2
4½	.77	–	–
5½–6	.86	.80	.77
7½–8	.94	.78	.77
11	.86	.80	.77
14	.86	–	–
16½	–	.87	.88
17½	.86	–	–

Note: Age ranges for the Kaufman-II are 4–5, 6, 7–9, 10–12, 13–15, and 16–19; age ranges for the two WJ III subsamples are 5–6, 7–8, 9–13, and 14–19. The coefficient of .78 for WJ III Subsample 1 was not compared across samples because of the variance difference.

factor variances were fixed to one, and then the covariance was fixed to one. All of the tests indicated that indeed there was *not* a perfect correlation between these two second-order factors in any of the age groups. They were strongly correlated, yet distinct factors.

3.8. Summary of g loadings for broad abilities and subtests

In Table 8, loadings on *COG-g* are presented, by age, for the Kaufman-II and WJ III Broad Abilities, and in Table 9 comparable data are presented for individual subtests (for the WJ III, *g* loadings were averaged for Subsamples 1 and 2).

3.9. Broad abilities

Gf demonstrated the strongest relation to *COG-g*, averaging 1.0 for the Kaufman-II (a direct function of the Gf error variance being set to zero in the models) and .95 for the WJ III. Average *g* loadings were in the .80s for Gc, Gv, and Glr for both test batteries, suggesting good consistency across data sets (Table 8). The one area of disparity concerns Gsm, which had by far the weakest average *g* loading (.71) on the Kaufman-II, but ranked second to Gf with a mean of .92 on the WJ III. Auditory Processing (Ga; mean *g* loading = .82) and Gs (.62) round out the WJ III picture (Gs is not measured by the Kaufman-II and Ga—only measured at age 6 on the KTEA-II—was excluded from the Kaufman-II analyses.). Thus, Gf had the strongest relation with *COG-g* for both data sets; Gsm had the weakest relation for the Kaufman-II and Gs had the weakest relation for the WJ III. These results are consistent with current confirmatory factor analysis research with the WAIS-IV (Lichtenberger & Kaufman, 2009), WISC-IV (Keith, Fine, Taub, Reynolds, & Kranzler, 2006), KABC-II (Reynolds et al., 2007), DAS-II (Keith et al., 2010), WJ-R (Bickley, Keith, & Wolfle, 1995; McGrew, Werder, & Woodcock, 1991), and WJ-III (Woodcock, 2001). All have shown that *Gf* has the highest loading on *COG-g*. In addition, *Gf*, *Gc*, and *Glr* almost always have higher loadings on *g* than either *Gsm* or *Gs*.

Table 9
Subtest analysis: the six best measures of *COG-g* on the Kaufman-II and WJ III.

Subtest	Broad Ability	Mean COG-g Loading (and range)
Kaufman-II		
Riddles	Gc	.72 (range = .69–.77)
Verbal knowledge	Gc	.71 (range = .67–.75)
Pattern reasoning	Gf	.70 (range = .67–.75)
Expressive vocabulary	Gc	.69 (range = .64–.73)
Story completion	Gf	.63 (range = .60–.72)
Rebus	Glr	.61 (range = .51–.72)
WJ III		
General information	Gc	.67 (range = .64–.70)
Concept formation	Gf	.66 (range = .64–.69)
Verbal comprehension	Gc	.66 (range = .64–.68)
Visual–auditory learning	Glr	.61 (range = .59–.62)
Oral comprehension	Gc	.58 (range = .57–.61)
Numbers reversed	Gsm	.58 (range = .56–.60)

Note: Mean loadings are based on six age groups between 4–5 and 16–19 years on the Kaufman-II and for four age groups between ages 5–6 through 14–19 years on the WJ III. Values shown for WJ III represent the mean value for Subsample 1 and Subsample 2. For the Kaufman-II, Word Order (Gsm) and Conceptual Thinking (Gv/Gf) were among the highest-loading subtests for ages 4–5 years.

3.10. Subtests

Although Gf was the strongest measure of *g* for both test batteries, the Gc variables emerged as the best measures of *g* among the subtests. Table 9 presents average *g* loadings, across the age range, for the six best measures of *g* on both the Kaufman-II and WJ III. Three of the six subtests for each test battery primarily measure Gc, including the best

Table 8
COG-g loadings of broad abilities on the Kaufman-II and WJ III, by age.

Age group (Median age)	Gf		Gc		Gv		Glr		Gsm		Ga	Gs
	K	WJ	K	WJ	K	WJ	K	WJ	K	WJ	WJ	WJ
4½	–	–	.88	–	.99	–	.97	–	.84	–	–	–
5½–6	1.0	.97	.83	.85	.88	.75	.82	.85	.79	1.0	.84	.75
7½–8	1.0	.97	.88	.85	.89	.93	.84	.84	.65	.90	.81	.56
11	1.0	.88	.88	.86	.82	.77	.78	.83	.64	.93	.82	.55
14	1.0	–	.86	–	.88	–	.85	–	.66	–	–	–
16½	–	.96	–	.90	–	.74	–	.78	–	.84	.79	.60
17½	1.0	–	.83	–	.82	–	.85	–	.68	–	–	–
Mean	1.0	.95	.86	.86	.88	.80	.85	.83	.71	.92	.82	.62

Note: K = Kaufman; WJ = WJ III. Age ranges for the Kaufman-II are 4–5, 6, 7–9, 10–12, 13–15, and 16–19; age ranges for the two WJ III subsamples are 5–6, 7–8, 9–13, and 14–19. Values shown for WJ III represent the mean value for Subsample 1 and Subsample 2, by age group.

Author's personal copy

measures of *g* on the Kaufman-II (Riddles, Verbal Knowledge) and WJ III (General Information). Gf is represented among the top six by two Kaufman-II subtests and one WJ III subtest, and Glr is represented for each test battery. The six best measures of *g* on the WJ III were remarkably consistent from age group to age group. That age-to-age consistency was demonstrated by most Kaufman-II subtests, with one notable exception: Word Order (Gsm) was among the best measures of *g* for ages 4–5 (.72) and 6 (.72), but not for the older four age groups (.49–.54). Like the results for the Broad Abilities, these subtest results—namely, the highest *g* loadings by Gc subtests—were entirely consistent with the Wechsler literature and with empirical analyses of the KABC-II and WJ III.

4. Discussion

The central purpose of this study was to assess the relation between a second-order latent *g*-factor from a battery of individually administered cognitive ability tests (*COG-g*) and a second-order latent *g*-factor from a battery of individually administered academic achievement tests (*ACH-g*). Multi-group confirmatory factor analysis was used to demonstrate invariance of the first- and second-order factor loadings across age groups, allowing for valid comparison of factor variance/covariances and correlations across age. Remarkable consistency was found across two large, nationally representative samples and two independent sets of batteries. Like Deary et al.'s (2007) value of .81, the overall mean correlation coefficient between *COG-g* and *ACH-g* was .83 ($Mdn = .80$), generally ranging from .77 to .88 (depending on age and sample, with a .94 correlation as an outlier), spanning the entire school age range (4–19).

Therefore, Binet and Simon (1916) succeeded in their goal of predicting broad academic achievement from tests of general cognitive ability. Even present day tests of general cognitive ability show extremely high correlations with general academic achievement. In the title of this paper, however, we posed the question: Is *COG-g* and *ACH-g* one and the same *g*? The answer to this question is no. They are highly related, yet distinct constructs. And although first-order unique correlations were not discussed in detail, their presence also suggests that beyond the general factor, more specific cognitive factors are important for explaining specific aspects of achievement.

In terms of developmental trends, for both WJ III subsamples, *g*-factors in the oldest age ranges were less differentiated than the youngest age groups. In the Kaufman-II samples the strongest correlation was .94 for 7–9 year olds and the weakest was .77 for ages 4–5, but the correlations were equivalent ($r = .87$) for age 6 and the age groups between 10 and 19 years. As mentioned previously, Table 7 shows a very small, and certainly not conclusive, trend for the coefficients to increase with age, with the values for adolescents tending to be higher than the values for children.

These results counter the findings of prior studies, which have demonstrated *declining* correlations between intelligence and academic achievement with school age (Jensen, 1998). One explanation for this decline has been that restriction of range occurs because students drop out of school as education becomes more advanced (Pind, Gunnarsdottir, & Johannesson, 2003). Indeed, after controlling for restriction of range, Laidra, Pullmann, and Allik (2007) found that academic achievement (as assessed by GPA) relies on general cognitive ability (as assessed by the Raven's Standard Progressive Matrices) through the school years. Before controlling for restriction of range, however, they found a decreasing relation between GPA and IQ. In the current study, since age invariance was demonstrated, we were able to make valid comparisons factor variance/covariances across the different age groups. Therefore, our finding of an increase or no change in the *COG-g* and *ACH-g* correlation with age fully takes into account potential changes in the range of the constructs.

Why are correlations relatively lower for preschool children and in the early school years? Messick (1984) noted that cognitive ability interacts with knowledge structures during the learning process. According to Messick, at initial stages of learning in a specific area, it is the cognitive abilities that drive or aid in the acquisition of new knowledge and skills. As learning proceeds, "these cognitive thinking skills are applied to address problems of a higher order of complexity than at less advanced levels (p. 223)." Thus if *COG-g* represents or underlies some type of general learning mechanism that is a primary driver of the acquisition and structuring of new knowledge within and between academic specific domains, then individual differences in general academic achievement might mostly be explained by *COG-g*. Moreover, over time and with more exposure to school curriculum, individual differences in the accumulation of skills across academic domains may more likely be a reflection of individual differences in *COG-g*. Perhaps this also explains the higher correlations in the oldest age groups in the WJ III sample.

4.0.1. How are COG-g and ACH-g related?

Regardless of the developmental trend, an important but unresolved issue is how *COG-g* and *ACH-g* are related. The results of the current study provide strong evidence that second-order *COG-g* and *ACH-G* are highly related constructs from K-12. There have been many different hypotheses and theories put forth about the relation between cognitive and achievement constructs (and measures), not all of which are mutually exclusive. Rather than select a "pet" theory we will provide a number of different possibilities.

First, some researchers have argued that *COG-g* and *ACH-g* are perfectly correlated or even identical constructs (Spearman, 1904). Such arguments seem reasonable because it has been shown that measures of general academic achievement mostly measure *COG-g* (Frey & Detterman, 2004; Koening, Frey, & Detterman, 2008). Moreover, *COG-g* factors extracted from different intelligence test batteries have been found to correlate perfectly with each other (Johnson et al., 2004; Keith, Kranzler, & Flanagan, 2001), and our study shares a similar methodology in that scores from individually administered test batteries were analyzed. Nevertheless, the findings from the current study, as well as those from others (e.g., Deary et al., 2007), indicate that although latent *COG-g* and *ACH-g* factors are highly correlated, they are not perfectly correlated.

One reason for the lack of perfect correlation may be that non-cognitive variables might be captured in performance across academic achievement measures to a lesser extent than cognitive measures. Individual influences such as Conscientiousness, motivation, and self-belief influence the application

Author's personal copy

of cognitive ability to academic learning (Furnham, Chamorro-Premuzic, & McDougall, 2002; Sternberg, 1999). Models of school learning indicate that although *COG-g* has an important influence on school learning, so does academic motivation and other variables such as time spent learning and the quality and amount of instruction (e.g., Carroll, 1963; Haertel, Walberg, & Weinstein, 1983). [4]

Note that most of these theories include cognitive ability (or aptitude) as an explanatory variable. Temporal precedence has been supported in some research. For example, Watkins et al. (2007) using data from a sample of special education students, showed that psychometric intelligence was predictive of future achievement whereas achievement was not predictive of future psychometric intelligence. The findings supported the idea of temporal precedence as well as evidence that they are distinct constructs. Thus, although individual differences in general academic achievement may to a large extent reflect individual differences in *COG-g*, they are not identical constructs as other important influences are involved in general achievement.[5]

Other, and not necessarily mutually exclusive theories, from cognitive and developmental psychology, may also describe why there is a strong relation between *COG-g* and *ACH-g* factors. For example, according to Cattell's (1987) investment hypothesis, a child's "rate of learning in fields demanding insights into complex relations—and these fields include especially the problems of reading, arithmetic, and abstract reasoning with which he struggles at school—will depend appreciably on his level of fluid intelligence (though motivation, goodness of teaching, etc., will still play a part, as with the acquisitions of low relational complexity)" (p. 139). Ferrer and McArdle (2004) applied bivariate dual change score models to longitudinal data on mainstream education students and found that fluid intelligence (*Gf*) was indeed a leading indicator of changes in academic achievement, while achievement was a lagging indicator of Gf. Their study supported one aspect of the investment hypothesis in that Gf is a primary driver of academic achievement. Again, the findings also supported the notion that intelligence and achievement are interrelated, yet separable constructs. [6]

What are some of the most important cognitive processes that may drive academic achievement? Various cognitive mechanisms underlie *COG-g*, such as working memory, explicit associative learning, and processing speed (Kaufman, DeYoung, Gray, Brown, & Mackintosh, 2009).[7] These cognitive mechanisms are also associated with academic achievement. In a recent study, Vock, Preckel, and Holling (2011) found that the "basic cognitive abilities" mental speed and short-term memory exerted their effects on achievement indirectly by affecting the more "complex cognitive abilities" reasoning and divergent thinking. Similar results have been found by Rindermann and Neubauer (2004). Luo, Thomson, and Detterman (2003) also found that mental speed and more basic elementary processes played an important role in the relation between *COG-g* and academic achievement, although they described the role as of these processes either as a mediator between *COG-g* and achievement, or as a potential common cause.

Other research and theories indicate that there is a reciprocal relation between cognitive abilities and achievement or knowledge acquisition. In a recent study, Rindermann, Flores-Medoza, and Mansur-Alves (2010) found in Brazilian and German samples similar effects of fluid intelligence on crystallized intelligence and vice versa. Their model supports a mutually beneficial relationship between intelligence and knowledge. The mutually reinforcing nature of *COG-g* and *ACH-g* is also explained by van der Maas et al. (2006), who argue that the positive manifold arises due to positive beneficial interactions between cognitive processes during development. For example, school learning may boost various cognitive processes, which in turn impact on cognitive development. Or as already noted, Messick (1984) argued that cognitive ability constantly interacts with knowledge structures during the learning process. Indeed, Ceci (1991) reviewed evidence that learning context does have an important impact on cognitive development, suggesting that future research should investigate the mutually reinforcing nature of abilities and expertise during the development of academic knowledge (also see Sternberg, 1999).

Another plausible explanation for the correlation between these two second-order latent *COG-g* factors is that there may be a common cause that produced this correlation. Sternberg (1999) proposes that a collection of skills, including metacognitive skills, learning skills, and thinking skills, underlie performance on both tests of *COG-g* and tests of *ACH-g*. Indeed, Sternberg argues that performance on tests of *COG-g* and *ACH-g* both require a common set of developing expertise, which explains their high correlation.

In a different, but not incompatible, line of research, Petrill and Wilkerson (2000) explored various environmental and genetic causes of the correlation between standardized intelligence and academic achievement tests. Reviewing evidence for an environment-only position, a gene + environment hypothesis, and a gene-only hypothesis, they concluded that "although intelligence and achievement themselves are influenced by both genetic and environmental factors, the correlation between intelligence and achievement appears to be driven mostly by shared genetic influences (p. 191)." This conclusion was supported by research such as that by Thompson, Detterman, and Plomin (1991) who found that cognitive and achievement correlations were largely explained by a common set of genes. Common and within-pair environmental differences, however, were primarily responsible for ability-achievement discrepancies. These results are intriguing, and future research should be conducted to understand the genetic and environmental factors that influence *COG-g* and *ACH-g* correlations as well as discrepancies.

[4] See Ceci (1991) for evidence that schooling influences IQ scores. It is important, however, to try to separate changes in the constructs (*g*) from changes in the vehicles (IQ scores) of measurement (Jensen, 1998).

[5] If a developmental trend of higher correlations is clearly established in future research, two possible, among other, reasons may be that that variables related to learning are also captured more in intelligence test scores with increasing age or that the *COG-g* factor tends to "dominate" even more with age.

[6] Ferrer and McArdle (2004) also argued aggregating achievement and cognitive scores obscured important findings, but that discussion is beyond the scope of this study.

[7] In our study some of these more specific cognitive factors (e.g., short-term memory) also had strong correlations with specific achievement factors, which suggests that more narrow cognitive abilities or processes influenced specific achievement areas.

Author's personal copy

Whatever the causes of the higher *COG-g*/*ACH-g* correlation, it is clear that future research is needed to more clearly delineate the causal cognitive mechanisms and direct and indirect pathways that determine the *COG-g*/*ACH-g* relationship across the lifespan. A number of different causes, including the possibilities mentioned here, most likely contribute to the high correlation we found in our study.

4.0.2. Is ACH-g a reflective latent variable?

These issues relate to the proper conceptualization of *ACH-g*. Although we *modeled* both *COG-g* and *ACH-g* as reflective latent variables, *ACH-g* is likely better *conceptualized* as a composite or formative latent variable (Bollen & Bauldry, 2011). If *ACH-g* were a reflective latent variable then an increase in *ACH-g* should improve all achievement areas, including reading, writing, and math, or *ACH-g* is a common cause of correlations between those academic areas. It seems more appropriate that for *ACH-g*, the causal arrows would move out from the subtests and flow into the latent variable, where *ACH-g* is merely a composite, a combination of scores across various academic achievement areas.

Consistent with this conceptualization, Kan, Kievit, Dolan, and van der Maas (2011) recently argued that crystallized intelligence, as hypothesized in Cattell's investment hypothesis, is purely a statistical entity, and is *constructed* from measured variables whereas *COG-g* is more likely a latent variable that exists independently of the scientist. They found that *COG-g* accounted for the correlations between crystallized abilities, and *Gc* (using the CHC model as a framework) was essentially verbal comprehension (also see Johnson & Bouchard, 2005). Therefore, for example, if *ACH-g*, as it sometimes is, were to be equated with "crystallized intelligence", *ACH-g* would be best conceptualized as a statistical factor that *summarizes* the covariances among various crystallized abilities, whereas *COG-g* is best thought of as a substantive underlying variable. Regardless of whether one views crystallized intelligence and academic achievement as identical constructs (and for the most part people do not), *ACH-g* may be better conceptualized as a composite variable.

4.0.3. Study limitations and suggestion for future research

The current investigation is not without limitations that should be addressed in future studies. First, although a stated strength of this study was the cross-sectional nature of the data, our discussion of hypothesized causal relations between *COG-g* and specific CHC broad factors causing or influencing achievement must be tempered by the fact that cross-sectional data were utilized. The current findings, with similar comprehensive conformed measures of *COG-g* and *ACH-g*, in longitudinal or time-lagged research designs, are needed to validate or modify the conclusions tendered here.

Second, while we used individually-administered standardized measures of reading, writing, and math to form *ACH-g*, we acknowledge that this is not the only way to measure academic achievement; other researchers have used different measures to estimate *ACH-g*, such as GPA or group-administered national public examination results across 25 academic subjects (e.g., Deary et al., 2007). While we believe using standardized measures of academic achievement has its methodological advantages, we acknowledge that relations between *COG-g* and *ACH-g* could be higher when standardized measures are used, since other measures, such as school grades and group-administered test scores, may allow for dispositional variables (e.g., conscientiousness and effort) to come more into play (Duckworth, Quinn, & Tsukayama, 2011). Future research should use a variety of different academic achievement outcome measures to assess the relation between *COG-g* and *ACH-g*, and investigate how the relation differs depending on the measures employed.

Third, researchers often use achievement measures as proxy variables for *g*. Students often have achievement, rather than cognitive, test scores in their records from repeated group testing, so, for example, the composite of their reading, writing, and math scores may be used to represent a proxy for *COG-g* in research investigations. These proxy variables are often the best estimate of *COG-g* available and our findings suggest composites scores from standardized achievement tests are adequate proxies. The two constructs are not, however, identical and it would be ideal to use measures of cognitive ability when available.

Alternatively, maintaining separate cognitive and achievement batteries for purposes of individualized psychoeducational assessments is important even though the overall composites scores may have a high correlation. Messick (1984) notes, "Because student characteristics as well as social and educational experiences influence current performance, the interpretation and implications of educational achievement measures must be relative to interpersonal and situational contexts (p. 215)." This idea is also represented in the Intelligent Testing approach (Kaufman, 1979), in which the test administrator interprets the students test profile within the context of his or her pertinent background information, clinical behaviors observed during testing, and the latest theories in cognitive science and learning theory. Therefore, while *COG-g* and *ACH-g* may be strongly related at the group level of analysis, caution but be made on inferring cognitive ability from achievement at the individual level in situations where practical decisions and recommendations must be made about the individual. For some individuals, low academic achievement scores may be more of a reflection of a lack of opportunity for learning, or a specific learning disability, whereas for others, it may be more of a reflection of *COG-g*.

Lastly, this study included two sets of batteries administered to large representative samples, which offers a certain degree of cross-validation of the findings. Nonetheless, studies involving Wechsler's scales and other standardized measures of intelligence and achievement are needed for additional cross-validation. Because archival data are available for a variety of populations, across the age range, who have been tested on measures of both cognitive ability and achievement, we urge researchers who have access to these data to attempt to replicate our findings.

Another caveat pertains to any attempt to explore the relation between *COG-g* and *ACH-g*, namely the distinction between what constitutes intelligence and what constitutes achievement is sometimes fuzzy. We limited achievement to the academic areas of reading, math, and writing. However, the Arithmetic subtest has long been a component of Wechsler's Full Scale IQ; additionally, a variety of measures of listening comprehension, oral expression, and quantitative reasoning have been included on standardized measures of

Author's personal copy

both intelligence and achievement (including the tests in the present study). We used the CHC theoretical model to support our division of subtests into *COG-g* and *ACH-g* (even though that meant including Kaufman and Woodcock "achievement" subtests in the *COG-g* analysis). When other theoretical or clinical or practical models are followed, or when other measures of academic achievement are used such as national public examinations (Deary et al., 2007), some different decisions are likely to be made when labeling subtests COG or ACH and those decisions may affect the relation between the *COG-g* and *ACH-g* constructs across different age groups.

5. Conclusions

Collectively, the findings of this set of studies result in the conclusion that *COG-g* and *ACH-g* are separate but highly related constructs during childhood and adolescence. The finding was consistent across two different test batteries. Considering the prevalence of both standardized tests of cognitive ability and standardized tests of academic achievement in education, we hope researchers will continue to investigate the fascinating linkages among these two constructs, their development, and etiology.

Acknowledgements

We would like to thank Timothy Z. Keith for helpful suggestions on an earlier draft of this article.

References

Benson, N. (2008). Cattell–Horn–Carroll cognitive abilities and reading achievement. *Journal of Psychoeducational Assessment*, *26*, 27–41.
Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.
Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–600.
Bickley, P. G., Keith, T. Z., & Wolfle, L. M. (1995). The three-stratum theory of cognitive abilities: Test of the structure of intelligence across the life span. *Intelligence*, *20*, 309–328.
Binet, A., & Simon, T. (1916). *The development of intelligence in children.* Baltimore, MD: Williams & Wilkens (E. S. Kite, Trans.).
Bollen, K. A., & Bauldry, S. (2011). Three Cs in measurement models: Causal indicators, composite indicators, and covariates. *Psychological Methods*, *16*, 265–284.
Bradley-Johnson, S., Morgan, S. K., & Nutkins, C. (2004). Review of the Woodcock–Johnson III tests of achievement. *Journal of Psychoeducational Assessment*, *22*, 261–274.
Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, *64*, 723–733.
Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.
Cattell, R. B. (1987). *Intelligence: Its structure, growth and action.* North-Holland: New York, NY.
Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, *27*, 703–722.
Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464–504.
Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.
Cizek, G. J. (2003). [Review of the Woodcock–Johnson III.]. In B. S. Plake, & J. C. Impara (Eds.), *The fifteenth mental measurements yearbook* (pp. 1020–1024). Lincoln,NE: Buros Institute of Mental Measurements.
Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13–21.
Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2011). What no child left behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, doi:10.1037/a0026280.
Ferrer, E., & McArdle, J. J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*, *40*, 935–952.
Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2007). *Essentials of cross-battery assessment with CD/ROM* (2nd ed.). Hoboken, N. J.: Wiley.
Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (in press). *Essentials of cross-battery assessment* with CD/ROM (3rd ed.). Hoboken, N. J.: Wiley.
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.
Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the Scholastic Assessment Test and general cognitive ability. *Psychological Science*, *15*, 373–378.
Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2002). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, *14*, 47–64.
Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, *44*(11 Suppl. 3), S78–S94.
Gustaffson, J. -E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*, 179–203.
Gustaffson, J. -E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, *28*, 407–434.
Haertel, G. D., Walberg, H. J., & Weinstein, T. (1983). Psychological models of educational performance: A theoretical synthesis of constructs. *Review of Educational Research*, *53*, 75–91.
Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, *57*, 253–270.
Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf–Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (pp. 53–91). New York: Guilford Press.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.
Jensen, A. R. (1998). *The g factor: The science of mental ability.* Westport, CT: Praeger.
Johnson, W., & Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, *33*, 431–444.
Johnson, W., Bouchard, T. J., Jr., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence*, *32*, 95–107.
Kan, K. -J., Kievit, R. A., Dolan, C., & van der Maas, H. (2011). On the interpretation of the CHC factor Gc. *Intelligence*, *39*, 292–302.
Kaufman, A. S. (1979). *Intelligent testing with the WISC-R.* New York: Wiley.
Kaufman, A. S. (2009). *IQ testing 101.* New York: Springer.
Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children-Second Edition (K-ABC-II).* Circle Pines, MN: American Guidance Service.
Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Test of Educational Achievement-Second Edition (KTEA-II). Comprehensive form.* Circle Pines, MN: American Guidance Service.
Kaufman, A. S., Lichtenberger, E. O., Fletcher-Janzen, E., & Kaufman, N. L. (2005). *Essentials of KABC-II assessment.* New York: Wiley.
Kaufman, S. B., DeYoung, C. G., Gray, J. R., Brown, J., & Mackintosh, N. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*, *37*, 374–382.
Keith, T. Z. (2006). *Multiple regression and beyond.* Boston, MA: Pearson.
Keith, T. Z. (1999). Effects of general and specific abilities on student achievement: Similarities and differences across ethnic group. *School Psychology Quarterly*, *14*, 239–262.
Keith, T. Z., & Reynolds, M. R. (2012). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. F. Flanagan, & P. H. Harrison (Eds.), *Contemporary Intellectual Assessment* (3rd ed.) (pp. 758–799). New York, NY: Guilford.
Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler Intelligence Scale for Children-Fourth Edition: What does it measure? *School Psychology Review*, *35*, 108–127.
Keith, T. Z., Kranzler, J. H., & Flanagan, D. P. (2001). What does the Cognitive Assessment System (CAS) measure? Joint confirmatory factor analysis of the CAS and the Woodcock–Johnson Tests of Cognitive Ability (3rd Edition). *School Psychology Review*, *30*, 89–119.
Keith, T. Z., Low, J. A., Reynolds, M. R., Patel, P. G., & Ridley, K. P. (2010). Higher-order factor structure of the Differential Ability Scales-II: Consistency across ages 4 to 17. *Psychology in the Schools*, *47*, 676–697.

Author's personal copy

Koening, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence*, *36*, 153–160.
Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, *42*, 441–451.
Lichtenberger, E. O., & Breaux, K. C. (2010). *Essentials of WIAT-III and KTEA-II assessment.* New York: Wiley.
Lichtenberger, E. O., & Kaufman, A. S. (2009). *Essentials of WAIS-IV assessment.* New York, NY: Wiley.
Luo, D., Thomson, L. A., & Detterman, D. K. (2003). The causal factor underlying thecorrelation between psychometric g and scholastic performance. *Intelligence*, *31*, 67–83.
Lynn, R., & Meisenberg, G. (2010). National IQs calculated and validated for 108 nations. *Intelligence*, *38*, 353–360.
Mackintosh, N. J. (1998). *IQ and human intelligence.* Oxford, UK: Oxford University Press.
McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research*, *29*, 409–454.
McArdle, J. J., & Woodcock, R. W. (1997). Expanding test-retest designs to include developmental time-lag components. *Psychological Methods*, *2*, 403–435.
McGrew, K. A. (2005). The Cattell–Horn–Carroll theory of cognitive ablities: Past, present, and future. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 136–181). (2nd ed.). New York: Guilford.
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., Werder, J., & Woodcock, R. (1991). *WJ-R technical manual.* Chicago, IL: Riverside Publishing.
McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual. Woodcock–Johnson III.* Itasca, IL: Riverside Publishing.
Messick, S. (1984). The psychology of educational measurement. *Journal of Educational Measurement*, *21*, 215–237.
Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment*, *21*, 244–260.
Petrill, S. A., & Wilkerson, B. (2000). Intelligence and achievement: A behavioral genetic perspective. *Educational Psychology Review*, *12*, 185–199.
Pind, J., Gunnarsdottir, E. K., & Johannesson, H. S. (2003). Raven's standard progressive matrices: New school age norms and a study of the test's validity. *Personality and Individual Differences*, *34*, 375–386.
Psychological Corporation (2003). *WISC-IV technical and interpretive manual.* San Antonio, TX: The Psychological Corporation.
Reynolds, M. R., Keith, T. Z., Fine, J. G., Fisher, M. E., & Low, J. A. (2007). Confirmatory factor structure of the Kaufman Assessment Battery for Children-Second Edition: Consistency with Cattell–Horn–Carroll theory. *School Psychology Quarterly*, *22*, 511–539.
Rindermann, H. (2007). The big G-factor of national cognitive ability. *European Journal of Personality*, *21*, 767–787.
Rindermann, H., & Neubauer, A. C. (2004). Processing speed, intelligence, creativity, and school performance: Testing of causal hypotheses using structural equation models. *Intelligence*, *32*, 573–589.
Rindermann, H., Flores-Medoza, C., & Mansur-Alves, M. (2010). Reciprocal effects between fluid and crystallized intelligence and their dependence on parents' socioeconomic status and education. *Learning and Individual Differences*, *20*, 544–548.
Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, *35*, 83–92.
Sandoval, J. (2003). Review of the Woodcock–Johnson III. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook* (pp. 1024–1028). Lincoln, NE: Buros Institute of Mental Measurements.
Sares, T. (2005). Review of the Woodcock–Johnson III diagnostic supplement to the tests of cognitive abilities. In R. A. Spies, & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook [Electronic version]. Retrieved July 29, 2005 from the Buros Institute's Test Reviews Online website* http://www.unl.edu/buros
Spearman, C. (1904). "General intelligence", objectively determined and measured. *The American Journal of Psychology*, *15*, 201–293.
Spinks, R., Arndt, S., Caspers, K., Yucuis, R., McKirgan, L. W., Pfalzgraf, C., et al. (2007). School achievement strongly predicts midlife IQ. *Intelligence*, *35*, 563–567.
Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. *Structural Equation Modeling*, *5*, 411–419.
Steiger, J. H., & Lind, J. (1980). Statistically based tests for the number of common factors. June. *Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.*
Sternberg, R. J. (1999). Intelligence as developing expertise. *Contemporary Educational Psychology*, *24*, 359–375.
Taub, G., Floyd, R. G., Keith, T. Z., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics. *School Psychology Quarterly*, *23*(2), 187–198.
Thompson, D. T. (2005). Review of the Woodcock–Johnson III diagnostic supplement to the tests of cognitive abilities. In R. A. Spies, & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook [Electronic version]. Retrieved July 29, 2005 from the Buros Institute's Test Reviews Online website* http://www.unl.edu/buros
Thompson, L. A., Detterman, D. K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science*, *2*, 158–165.
van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, *113*, 842–861.
Vock, M., Preckel, F., & Holling, H. (2011). Mental abilities and school achievement: A test of a mediation hypothesis. *Intelligence*, *39*, 357–369.
Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, *63*, 249–294.
Watkins, M. W., Lei, P., & Canivez, G. L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis. *Intelligence*, *35*, 59–68.
Woodcock, R. W. (2001). *Technical manual. Woodcock Johnson III.* Itasca, IL: Riverside Publishing.
Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. A. (2003). *Woodcock–Johnson III diagnostic supplement to the tests of cognitive abilities.* Itasca, IL: Riverside Publishing.