Psychological Assessment
2012, Vol. 24, No. 1, 129–140

1040-3590/12/$12.00 DOI: 10.1037/a0024878

Examination of the Structural, Convergent, and Incremental Validity of the Reynolds Intellectual Assessment Scales (RIAS) With a Clinical Sample

Jason M. Nelson
University of Georgia

Gary L. Canivez
Eastern Illinois University

Empirical examination of the Reynolds Intellectual Assessment Scales (RIAS; C. R. Reynolds & R. W. Kamphaus, 2003a) has produced mixed results regarding its internal structure and convergent validity. Various aspects of validity of RIAS scores with a sample (N = 521) of adolescents and adults seeking psychological evaluations at a university-based clinic were examined. Results from exploratory factor analysis indicated only 1 factor, and confirmatory factor analysis (CFA) indicated that the 1-factor model was a good fit and a better fit than the 2-factor model. Hierarchical factor analysis indicated the higher order, general intelligence factor accounted for the largest amount of variance. Correlations with other measures of verbal/crystallized and nonverbal/fluid intelligence were supportive of the convergent validity of the Verbal Intelligence Index but not the Nonverbal Intelligence Index. Joint CFA with these additional measures resulted in a superior fit of the 2-factor model compared with the 1-factor model, although the Odd-Item-Out subtest was found to be a poor measure of nonverbal/fluid intelligence. Incremental validity analyses indicated that the Composite Intelligence Index explained a medium to large portion of academic achievement variance; the NIX and VIX explained a small amount of remaining variance. Implications regarding interpretation of the RIAS when assessing similar individuals are discussed.

Keywords: Reynolds Intellectual Assessment Scales (RIAS), intelligence, cognitive assessment

With the creation of the Reynolds Intellectual Assessment Scales (RIAS), Reynolds and Kamphaus (2003a) offered a unique alternative to other contemporary intelligence tests. Despite consisting of only four core intelligence subtests (two additional subtests purportedly measure memory), the RIAS is described as a "comprehensive measure of verbal and nonverbal intelligence and of general intelligence" (Reynolds & Kamphaus, 2003b, p. 12). Its average length of administration (20 to 25 min) is substantially less than that of other frequently used comprehensive intelligence tests (e.g., Wechsler Scales, Stanford-Binet, Woodcock-Johnson Tests of Cognitive Abilities). Other unique attributes include its general elimination of dependence on motor coordination, visual-motor speed, and reading ability in the measurement of intelligence (Reynolds & Kamphaus, 2005). Because the RIAS is both time- and cost-efficient, it will likely become an attractive alternative for psychologists in a variety of settings.

Carroll's (1993) hierarchical three-stratum theory of intelligence was a "primary theoretical guide" (Reynolds & Kamphaus, 2005, p. 462) when developing the RIAS, with the Composite Intelligence Index (CIX) serving as an indicator of stratum three (i.e., general intelligence [*g*]) and the Verbal Intelligence Index (VIX) and Nonverbal Intelligence Index (NIX) serving as indicators of the stratum two abilities of crystallized (*Gc*) and fluid (*Gf*) intelligence, respectively. Despite the influence of a hierarchical theory of intelligence, only first-order exploratory factor analysis (EFA) and first-order confirmatory factor analysis (CFA) were used to investigate the internal structure of the RIAS during its standardization, although statistical methods for examining hierarchical models are available (e.g., Schmid & Leiman, 1957, procedure). Notwithstanding this issue, Reynolds and Kamphaus (2003b) reported evidence from both EFA and CFA in support of a two-factor first-order model over a one-factor first-order model for the four-subtest configuration of the RIAS.

Published reviews of the RIAS have been largely positive (Andrews, 2007; Bracken, 2005; Dombrowski & Mrazik, 2008; Elliott, 2004; Schraw, 2005); however, concerns related to aspects of validity have been asserted. Dombrowski and Mrazik (2008) criticized the EFA procedures reported in the RIAS Professional Manual. Specifically, they questioned reliance on the examination of scree plots and eigenvalues in determining the number of factors to extract and recommended that future research incorporate more rigorous factor extraction methods such as Horn's parallel analysis (HPA; Horn, 1965) and minimum average partial (MAP; Velicer, 1976) tests. Furthermore, they judged the orthogonal factor rotation method (a method that constrains the correlations between the factors to zero) used on the standardization data to be insufficient and argued that an oblique rotation method (a method that permits the factors to be correlated) and a higher order factor analysis via the Schmid and Leiman (1957) procedure were warranted because of the substantial correlation between the first-order factors and the hierarchical theory of intelligence upon which the RIAS was based. The reviewers argued that without use of such methods the viability of interpreting the RIAS beyond a single factor is un-

This article was published Online First August 8, 2011.

Jason M. Nelson, Regents' Center for Learning Disorders, University of Georgia; Gary L. Canivez, Department of Psychology, Eastern Illinois University.

Correspondence concerning this article should be addressed to Jason M. Nelson, Regents' Center for Learning Disorders, University of Georgia, 337 Milledge Hall, Athens, GA 30602. E-mail: jmnelson@uga.edu

known. Additional concerns have been asserted regarding the convergent validity of RIAS scores, particularly the NIX. Both Bracken (2005) and Dombrowski and Mrazik (2008) highlighted the finding reported in the RIAS Professional Manual that the correlation between the RIAS NIX and the Wechsler Intelligence Scale for Children, Third Edition (WISC–III) Verbal IQ ($r = .60$, $p < .01$) was higher than that between the NIX and the WISC–III Performance IQ ($r = .33$, *ns*). As a result, Bracken (2005) concluded, "This finding shows that the RIAS Nonverbal Scale is not only not nonverbal, it is very much a verbal scale" (p. 213).

Validity Studies of the RIAS

Several validity studies of the RIAS have been conducted that address the concerns advanced in published reviews. Using a large sample ($N = 1{,}163$) of referred school-age students, Nelson, Canivez, Lindstrom, and Hatt (2007) examined the factor structure of the RIAS by conducting EFA with the factor extraction methods reported in the Professional Manual along with HPA and MAP, both orthogonal and oblique rotations, and a higher order analysis with the Schmid and Leiman (1957) procedure. All factor extraction methods indicated only one factor and the higher order analysis indicated the largest amount of variance was accounted for by the higher order *g* factor. On the basis of these findings, Nelson et al. (2007) recommended that the majority of interpretive weight be apportioned to the CIX. Using the same methods as Nelson et al. (2007), Dombrowski, Watkins, and Brogan (2009) re-analyzed the RIAS standardization data and also found evidence for only one factor and that the higher order factor accounted for the largest amount of variance. Dombrowski et al. (2009) made the same interpretive recommendations as Nelson et al. (2007).

Using CFA, Beaujean, McGlaughlin, and Margulies (2009) re-examined the Nelson et al. (2007) data and data presented in the RIAS Professional Manual for 6- to 11-year-olds along with their own data from a sample of approximately 700 school-age students referred for special education evaluations. In each of the data sets, they found that the first-order two-factor model using the four-subtest configuration was a good fit to the data and superior to the first-order one-factor model. A hierarchical model was not examined, however, because of the limitations of their statistical procedures. Beaujean et al. (2009) stated, "Given the number of subtests that make up the RIAS, this model is impossible to fit in a CFA context without either collecting additional variables and/or constraining certain parameters (e.g., factor loadings) within and/or between groups" (p. 938). Curiously, despite the limitations of their statistical procedures, Beaujean et al. (2009) recommended that the VIX and NIX be given more clinical attention than the CIX.

The divergence of findings from EFA and CFA procedures is not uncommon for tests of cognitive ability. In their examination of the major tests of cognitive ability dating back to their earliest versions, Frazier and Youngstrom (2007) found substantial divergence between EFA and CFA procedures and the possibility of overfactoring if only results from the latter are relied upon. When the results of EFA and CFA converge, confidence can be placed in the latent structure of a test, whereas divergence of results between these methods reduces such confidence (Gorsuch, 1983). Thus, one way of assessing validity associated with the internal structure of tests of cognitive ability is to use both EFA *and* CFA procedures and determine their level of convergence (DiStefano & Dombrowski, 2006). In doing so, using separate samples for each procedure is recommended to avoid capitalizing on chance solutions from one data set, and traditionally EFA is conducted prior to CFA (DiStefano & Dombrowski, 2006). It appears that conducting a study with both procedures on separate samples would help clarify the internal structure of the RIAS.

Along with studies of the factor structure of the RIAS, several convergent validity studies have been conducted. These studies found support for the convergent validity of the VIX (Beaujean, Firmin, Michonski, Berry, & Johnson, 2010; Edwards & Paulin, 2007; Krach, Loe, Jones, & Farrally, 2009; Smith, McChristian, Smith, & Meaux, 2009; Umphress, 2008); however, mixed results on the convergent validity of the NIX have been found. Whereas some researchers found support for the convergent validity of the NIX (Edwards & Paulin, 2007; Smith et al., 2009; Umphress, 2008), Beaujean et al. (2010) and Krach et al. (2009) found that the NIX was more highly correlated with measures of *Gc* than it was with measures of *Gf*. One way to clarify issues of convergent validity recommended in the literature is for future researchers to conduct joint CFA with the RIAS and other instruments purported to measure *Gf* and *Gc* (Beaujean et al., 2009, 2010).

Further research that may aid in clarifying the validity of RIAS scores but that has yet to be conducted is examination of incremental predictive validity (for a review of the concept of incremental validity, see Hunsley & Meyer, 2003). Incremental validity relates to the "extent to which a measure adds to the prediction of a criterion beyond what can be predicted with other data" (Hunsley, 2003, p. 443). In the case of the RIAS, research on the incremental validity of the VIX and NIX in predicting external criteria (e.g., academic achievement) beyond the CIX would help determine the interpretative weight that should be apportioned to the lower order factors relative to the higher order factor. Validity studies of the internal structure of intelligence tests, although necessary, are insufficient in informing higher order versus lower order interpretations (Canivez, Konold, Collins, & Wilson, 2009). Examining both structural *and* incremental validity appears particularly informative for determining the appropriate interpretation of RIAS CIX, VIX, and NIX scores.

Purpose of Study and Research Questions

The purpose of the current study was to investigate the internal structure of the RIAS four-subtest configuration using clinical samples of adolescents and adults. Additionally, we sought to examine the convergent validity of the VIX and NIX, along with the incremental validity of these scores in predicting various aspects of academic achievement beyond the CIX. To address this purpose, the following three research questions were investigated: (a) What is the factor structure of the RIAS and the level of convergence of EFA (including a higher order analysis via the Schmid & Leiman, 1957, procedure) and CFA procedures when independent samples are investigated? (b) Using joint CFA, how do the RIAS subtests align with other measures of crystallized and fluid intelligence? and (c) What is the incremental validity of the RIAS VIX and NIX beyond the CIX in predicting academic achievement, including basic academic skills, academic fluency, and higher level academic skills?

Method

Participants

Participants were 521 individuals who sought comprehensive psychological evaluations over the time period of 2005 to 2010 at a university-based clinic specializing in learning disorders. The sample ranged in age from 16 to 70 years ($M = 21.68$, $SD = 6.92$) and was nearly evenly split by sex ($n = 258$ male, $n = 263$ female). The majority of the sample consisted of college undergraduates (69.67%; $n = 363$); however, high school students preparing to attend college (19.58%; $n = 102$), and graduate students (5.95%; $n = 31$) also participated. Fifteen participants (2.88%) could not be classified according to these educational status categories (e.g., some had graduated college but had not yet begun graduate school). Ethnicity included 81.57% White ($n = 425$), 6.14% African American ($n = 32$), 2.11% Hispanic ($n = 11$), 1.34% Asian American ($n = 7$), 1.54% Multiracial ($n = 8$), .02% other ($n = 1$), and 7.1% missing data ($n = 37$). The sample consisted of individuals with a wide variety of diagnoses but was predominantly made up of individuals with learning disabilities (LD) and/or attention-deficit/hyperactivity disorder (ADHD). Diagnostic status included 33.59% LD ($n = 175$), 21.98% ADHD ($n = 152$), 12.48% comorbid LD/ADHD ($n = 65$), 11.52% other diagnoses (e.g., mood or anxiety disorders; $n = 60$), and 6.33% no diagnosis ($n = 33$). Participants were diagnosed by licensed psychologists using the University System of Georgia criteria (see http://rcld.uga.edu for a detailed description of these criteria).

The total sample was randomly split for separate use in the EFA ($n = 261$) and CFA ($n = 260$) analyses. The samples did not differ on demographic variables of age, $F(1, 519) = 0.15$, $p = .70$; sex, $\chi^2(1) = 0.002$, $p = .97$; ethnicity, $\chi^2(5) = 1.16$, $p = .95$; or diagnosis, $\chi^2(7) = 6.01$, $p = .54$. They also did not differ on the CIX, $F(1, 519) = 0.13$, $p = .72$; VIX, $F(1, 519) = 0.14$, $p = .70$; NIX, $F(1, 519) = 0.001$, $p = .98$; Guess-What subtest, $F(1, 519) = 0.48$, $p = .49$; Verbal Reasoning subtest, $F(1, 519) = 0.05$, $p = .82$; Odd-Item-Out subtest, $F(1, 519) = 0.24$, $p = .62$; or What's Missing subtest, $F(1, 519) = 0.16$, $p = .69$.

Instruments

RIAS. The RIAS is an individually administered intelligence test for persons between the ages of 3 and 94 years. It is composed of a two-subtest measure of verbal/crystallized intelligence (i.e., the VIX) and a two-subtest measure of nonverbal/fluid intelligence (i.e., the NIX). An overall score (i.e., the CIX) is calculated from the sum of the *T* scores of the four subtests. The two subtests that compose the VIX include Guess What (GWH) and Verbal Reasoning (VRZ). On the GWH subtest, examinees attempt to identify an object or concept through the use of two to four verbally presented clues. The VRZ subtest requires examinees to complete verbal analogies using one or two words. Odd-Item-Out (OIO) and What's Missing (WHM) compose the NIX. On the OIO subtest, examinees are shown a card with five to seven pictures and are asked to identify the one that does not go with the others. The WHM subtest requires examinees to identify the missing element within a presented picture.

Instruments for examining convergent validity. Form A of the Peabody Picture Vocabulary Test, Fourth Edition (PPVT–IV; Dunn & Dunn, 2007) and three subtests from the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition (WJ–III COG; Woodcock, McGrew, & Mather, 2001b) were used to examine the convergent validity of the VIX and NIX. These instruments were selected for use in the current study because of the availability of supportive evidence for score reliability and validity and because the norms of the instruments span the age range of the population seeking evaluations at the clinic.

Measures of crystallized intelligence. The PPVT–IV and WJ–III COG Verbal Comprehension subtest were used as indicators of crystallized intelligence. The PPVT–IV measures receptive vocabulary, a cognitive ability that is typically subsumed under the construct of crystallized intelligence (Carroll, 1993). Supportive evidence for reliability of PPVT–IV Form A scores included a mean alpha coefficient of .97 and a mean test–retest reliability coefficient of .92. The WJ–III COG Verbal Comprehension subtest, a test of vocabulary and reasoning using lexical knowledge, is part of the *Gc* Cluster and had a median split-half reliability coefficient of .92. Additional psychometric evidence for validity of scores from these measures and all other measures in the present study is available in the respective test manuals.

Measures of fluid intelligence. The WJ–III COG Concept Formation and Analysis-Synthesis subtests compose the *Gf* Cluster and were used as indicators of fluid intelligence. Concept Formation was designed to measure inductive reasoning and scores had a median split-half reliability coefficient of .94. Analysis-Synthesis is a deductive reasoning test with a median split-half reliability estimate of .90 for its scores.

Instruments for examining incremental validity. Various aspects of academic achievement were assessed and served as external criteria for the incremental validity analyses. These included measures of basic academic skills, academic fluency, and higher level academic skills. These instruments were selected on the basis of available psychometric data supportive of score reliability and validity and because the instrument norms span the age range of the population commonly seeking evaluations at the clinic.

Measures of basic academic skills. The Woodcock-Johnson Tests of Achievement, Third Edition (WJ–III ACH; Woodcock, McGrew, & Mather, 2001a) Academic Skills Cluster was used as an indicator of overall basic academic skill. It is composed of the Letter-Word Identification, Calculation, and Spelling subtests. The median Academic Skills Cluster score reliability coefficient was .96.

Letter-Word Identification measures skill at reading words in isolation. On the Calculation subtest, examinees solve paper-and-pencil math computation problems. The Spelling subtest requires the examinee to spell progressively difficult words. Median split-half reliability estimates for scores from these three subtests were .94, .86, and .90, respectively.

Measures of academic fluency. The WJ–III ACH Academic Fluency Cluster was used as an indicator of overall academic fluency. It consists of the Reading Fluency, Math Fluency, and Writing Fluency subtests. The median reliability estimate of Academic Fluency Cluster scores was .93.

Reading Fluency measures the examinee's skill at quickly reading simple sentences and deciding if each statement is true. On the Math Fluency subtest, the examinee is required to solve simple arithmetic problems as quickly as possible. Writing Fluency requires the examinee to write simple, short sentences as quickly as possible. Both Reading Fluency and Math Fluency have time limits of 3 mins; Writing Fluency has a time limit of 7 mins. Median split-half reliability estimates were .90, .90., and .88 for Reading Fluency, Math Fluency, and Writing Fluency scores, respectively.

Measures of higher level academic skills. The WJ–III ACH Applied Problems subtest and Listening Comprehension Cluster and the Nelson-Denny Reading Test Form H (NDRT; Brown, Fishco, & Hanna, 1993) served as measures of higher level academic skills. Applied Problems measures math reasoning and scores had a median split-half reliability coefficient of .93. The Listening Comprehension Cluster measures the ability to understand oral language and scores had a median reliability estimate of .89. The NDRT is a timed test of reading comprehension. Kuder-Richardson 20 coefficients ranged from .85 to .89 for high school and college students.

Procedure

Data in the current study were archival and drawn from a database used to store deidentified demographic and assessment information for all students evaluated at the clinic. Each instrument was individually administered by a doctoral-level psychologist or a master's level clinician or doctoral student under the supervision of a licensed doctoral-level psychologist.

Data Analyses

Consistent with Reynolds and Kamphaus (2003b), principal axis EFAs were used to analyze reliable common variance from the four RIAS subtests using SPSS 19.0 for Macintosh OSX. As recommended by Gorsuch (1983), multiple criteria for determining the number of factors to retain were examined and included eigenvalues >1 (Guttman, 1954), the visual scree test (Cattell, 1966), standard error of scree (SE_{Scree}; Zoski & Jurs, 1996), HPA (Horn, 1965), MAP (Velicer, 1976), and theoretical expectation. The scree test was used to visually determine the optimum number of factors to retain but is subjective. The SE_{Scree}, reportedly the most accurate objective scree method (Nasser, Benson, & Wisenbaker, 2002), was used as programmed by Watkins (2007). HPA and MAP were included as they typically are more accurate than other factor extraction methods and therefore reduce overfactoring (Frazier & Youngstrom, 2007; Velicer, Eaton, & Fava, 2000). HPA indicated meaningful factors when eigenvalues from the present sample data were larger than eigenvalues produced by random data containing the same number of participants and factors (Lautenschlager, 1989). Random data and resulting eigenvalues for HPA were produced using the Monte Carlo PCA for Parallel Analysis computer program (Watkins, 2000) with 100 replications to provide stable eigenvalue estimates. The MAP criterion was computed using the SPSS code supplied by O'Connor (2000). Multiple factors were rotated with both oblique (promax) and orthogonal (varimax) methods to examine differences.

In hierarchical EFA, iterations in first-order principal axis factor extraction were limited to two in estimating final communality estimates (Gorsuch, 2003). EFA (principal axis extraction of two factors) was followed by promax (oblique) rotation ($k = 4$; Gorsuch, 2003) and the resulting first-order factors were orthogonalized using the Schmid and Leiman (1957) procedure as programmed in the MacOrtho computer program (Watkins, 2004). This transformed the "oblique factor analysis solution containing a hierarchy of higher order factors into an orthogonal solution which not only preserves the desired interpretation characteristics of the oblique solution but also discloses the hierarchical structuring of the variables" (Schmid & Leiman, 1957, p. 53). Although all but the theoretical expectation criterion for factor extraction suggested a one-factor solution, two first-order factors were extracted to compare results to other RIAS studies (Dombrowski et al., 2009; Nelson et al., 2007) as well as to examine proportions of variance attributed to RIAS first- and second-order factors.

Amos 18.0 (Arbuckle, 2009) was used to conduct CFAs using maximum likelihood (ML) estimation. ML estimation is the default method for estimating parameters in CFA. As recommended by Hu and Bentler (1999), a variety of fit indices were calculated to judge the fit of the models, including chi-square (χ^2), the Tucker Lewis Index (TLI), the Comparative Fit Index (CFI), the Goodness-of-Fit Index (GFI), the root-mean-square error of approximation (RMSEA), and the standardized root-mean-square residual (SRMR). Current criterion levels for determining good fit to the data include statistical insignificance for χ^2; TLI, CFI, and GFI values close to or greater than .95; RMSEA values close to or less than .06; and SRMR values close to or less than .08 (Hu & Bentler, 1999). Additional model fit criterion levels include RMSEA value cutoffs of .05 (good), .08 (adequate), and .10 (inadequate; Browne & Cudeck, 1993), and RMSEA values of .08 to .10 as indicating mediocre fit (MacCallum, Browne, & Sugawara, 1996). These fit indices were incorporated in the current study because they were utilized by Reynolds and Kamphaus (2003b) during the development of the RIAS and therefore aid in comparison.

We used the chi-square difference test ($\Delta\chi^2$) to determine whether statistically significant differences existed between CFA models. Akaike information criterion (AIC) and the Expected Cross-Validation Index (EVCI) were also used to assist in model comparison. Models with the smallest AIC and EVCI are usually considered the best fitting models (Loehlin, 2004).

Incremental validity was assessed through hierarchical multiple regression analyses (H–MRA) with WJ–III ACH Cluster and subtest scores and the NDRT serving as dependent variables. The RIAS CIX was entered into the first block and the RIAS VIX and NIX were entered into the second block. The change in predicted achievement variance (R^2) produced by the VIX and NIX in the second block indicated their combined incremental prediction of achievement beyond the RIAS CIX. Cohen's (1988) standards for R^2 effect size estimates were used for assessing the magnitude of incremental prediction ($R^2 = .03$ or 3% [small], $R^2 = .10$ or 10% [medium], $R^2 = .30$ or 30% [large]).

Results

Exploratory Factor Analysis

Pearson product–moment correlations and descriptive statistics for RIAS subtests are presented in Table 1. All skewness and kurtosis indices for all four subtests were less than 1 and therefore well within acceptable limits of normality (Curran, West, & Finch,

Table 1
Descriptive Statistics and Intercorrelations for RIAS Subtest Scores in the EFA Random Split Sample and CFA Random Split Sample

Subtest	*M*	*SD*	GWH	VRZ	OIO	WHM
		EFA sample (N = 261)				
Guess What (GWH)	49.29	7.30				
Verbal Reasoning (VRZ)	51.16	9.32	.62			
Odd-Item Out (OIO)	55.03	6.32	.24	.28		
What's Missing (WHM)	51.26	8.06	.38	.43	.25	
		CFA sample (N = 260)				
Guess What (GWH)	49.65	6.60				
Verbal Reasoning (VRZ)	51.35	9.61	.60			
Odd-Item Out (OIO)	54.76	6.37	.15	.25		
What's Missing (WHM)	51.54	8.22	.52	.46	.23	

Note. RIAS = Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003a); EFA = exploratory factor analysis; CFA = confirmatory factor analysis. All correlations were statistically significant ($p < .05$).

1996). Exploratory factor analysis results produced a Kaiser–Meyer–Olkin Measure of Sampling Adequacy coefficient of .69 (≥.6 required for good EFA; Tabachnick & Fidell, 2007) and Bartlett's Test of Sphericity $\chi^2 = 212.61$, $p < .0001$ (indicating correlation matrix was not random). Communality estimates ranged from .13 (OIO) to .68 (VRZ; *Mdn* = .41). Given the present communality estimates and sample size, it was judged the present sample size was adequate for factor analysis procedures (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Multiple criteria for determining the number of factors to extract and retain were in agreement as eigenvalues > 1, visual scree, SE_{Scree}, HPA, and MAP all indicated the extraction of one factor. Only theoretical consideration suggested that two factors be extracted. Two factors were extracted for examination because it is argued that it is better to overfactor than to underfactor (Wood, Tataryn, & Gorsuch, 1996) and allowed for comparison to other similar studies.

Table 2 presents results from the EFAs including varimax factor structure coefficients, promax factor pattern coefficients, promax factor structure coefficients, eigenvalues, and percents of variance. RIAS subtest *g*-loadings (structure coefficients from the first unrotated factor) ranged from good (VRZ and GWH) to fair (WHM) to poor (OIO) based on Kaufman's (1994) criteria (≥.70 = good, .50–.69 = fair, <.50 = poor). Varimax factor structure coefficients and promax factor pattern coefficients provided support for the theoretically consistent assignment of RIAS subtests to the latent factors they represent (Verbal and Nonverbal), but promax factor pattern and structure coefficients illustrated the correlated nature of the two factors. In the oblique solution, the correlation between Factor I (Verbal) and Factor II (Nonverbal) of .77 was high and indicated the presence of a higher order factor (Gorsuch, 1983; Tabachnick & Fidell, 2007). Given this large factor correlation, it is best to examine the hierarchical structure of the RIAS that is consistent with its development and theoretical representation.

Hierarchical Exploratory Factor Analysis

Results from the Schmid and Leiman (1957) procedure are presented in Table 3 and illustrate the proportions of variance apportioned to the higher order *g* and lower order Verbal and Nonverbal factors. The higher order *g* factor accounted for 34.39% of the total variance and 83.96% of the common variance. The *g* factor also accounted for between 14% and 51% of individual subtest variability. At the first-order level, Factor I (Verbal) accounted for an additional 5.16% of the total variance and 12.60% of the common variance, while Factor II (Nonverbal) accounted

Table 2
Two-Factor Principal Axis Exploratory Factor Analysis of the RIAS Four-Subtest Configuration With Varimax and Promax Rotations

		Varimax structure coefficients		Promax pattern coefficients		Promax structure coefficients	
Subtest	*g*-loading	I Verbal	II Nonverbal	I Verbal	II Nonverbal	I Verbal	II Nonverbal
GWH	.75	**.70**	.30	**.80**	−.05	**.76**	**.56**
VRZ	.81	**.71**	.40	**.74**	.10	**.81**	**.66**
OIO	.37	.16	.41	−.03	**.46**	.32	.44
WHM	.56	.34	**.49**	.16	**.46**	**.51**	**.58**
Eigenvalues		2.14	.83				
% Variance		41.45	3.22				

Note. RIAS = Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003a); GWH = Guess What; OIO = Odd-Item Out; VRZ = Verbal Reasoning; WHM = What's Missing. *g*-loadings are factor structure coefficients from the first unrotated factor (two-factor extraction). Salient factor structure coefficients (>.44) based on Comrey and Lee's (1992) classifications are presented in bold. Promax rotated Factor 1 and Factor 2 (r = .77).

Table 3
Factor Structure Coefficients and Variance Sources for the RIAS Four-Subtest Configuration Based on the Orthogonalized Higher Order Factor Model in the EFA Split Sample

	General		Verbal		Nonverbal				
Subtest	b	$\%S^2$	b	$\%S^2$	b	$\%S^2$	h^2	u^2	s^2
GWH	0.66	43	0.32	10	−0.01	0	0.54	0.46	0.38
VRZ	0.37	14	0.01	0	0.17	3	0.17	0.83	0.73
OIO	0.72	51	0.30	9	0.04	0	0.61	0.39	0.33
WHM	0.54	29	0.10	1	0.16	3	0.33	0.67	0.59
% Total S^2		34.39		5.16		1.41	40.96	59.04	51.05
% Common S^2		83.96		12.60		3.43			

Note. RIAS = (Reynolds & Kamphaus, 2003a); EFA = exploratory factor analysis; GWH = Guess What; VRZ = Verbal Reasoning; OIO = Odd-Item Out; WHM = What's Missing; b = factor structure coefficient (loading); h^2 = communality; u^2 = uniqueness; s^2 = specific (uniqueness − error).

for an additional 1.41% of the total variance and 3.43% of the common variance.

Confirmatory Factor Analyses

Pearson product-moment correlations and descriptive statistics for RIAS subtests in the CFA sample are presented in Table 1. All skewness and kurtosis indices for all RIAS subtests were less than 1 and therefore well within acceptable limits of normality (Curran et al., 1996). One- and two-factor model fit statistics are as follows: One-factor $\chi^2(2) = 6.84$, $p < .05$, CFI = .98, GFI = .99, TLI = .94, RMSEA = .10, SRMR = .03, AIC = 22.84, ECVI = .09; Two-factor $\chi^2(1) = 6.84$, $p < .05$, CFI = .98, GFI = .99, TLI = .89, RMSEA = .13, SRMR = .03, AIC = 23.25, ECVI = .09. The majority of fit index values were within or close to the cutoff values discussed above and therefore indicate a generally good model fit for the one-factor model. The fit indices for the two-factor model were slightly more variable. Whereas the CFI, GFI, and SRMR indicated good model fit, the χ^2, TLI, and RMSEA indicated inadequate fit.

Comparison of the two models with $\Delta\chi^2$ indicated the models were not statistically different, $\Delta\chi^2(1) = 1.59$, $p = .21$. AIC was slightly lower for the one-factor model and EVCI was identical for the two models. As shown in Figure 1, the GWH and VRZ subtests had good loadings on both the general and verbal/crystallized intelligence factors. The WHM had a fair loading on the general intelligence factor and a good loading on the nonverbal/fluid intelligence factor. The OIO subtest had poor loadings on both the general intelligence and nonverbal/fluid intelligence factors.

Joint Confirmatory Factor Analysis

Data for these analyses were obtained from the overall sample ($N = 521$) and included all participants from this sample that were administered the convergent validity measures ($N = 336$). Table 4 presents the Pearson product-moment correlations for the RIAS composite and subtest scores and the WJ–III COG and PPVT–IV scores. Correlations may be attenuated due to restricted range so correlations were corrected for variability in RIAS scores (Guilford & Fruchter, 1978) where the unrestricted RIAS *SD* = 15;

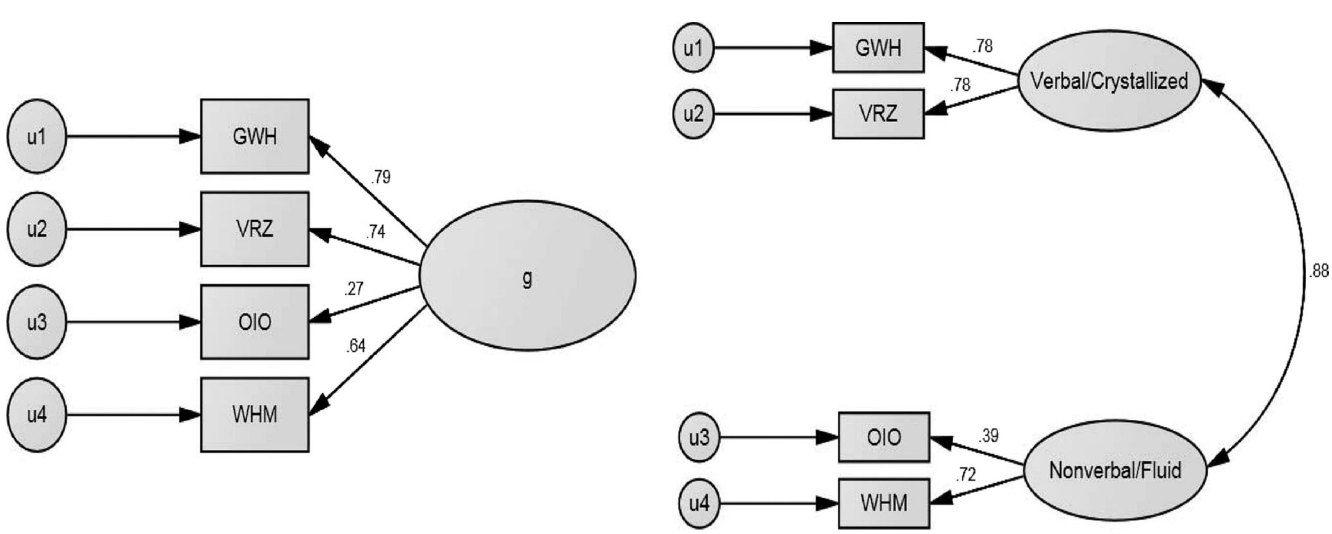

Figure 1. Diagram of confirmatory factor analysis models. GWH = Reynolds Intellectual Assessment Scales (RIAS; Reynolds & Kamphaus, 2003a) Guess What; VRZ = RIAS Verbal Reasoning; OIO = RIAS Odd-Item-Out; WHM = RIAS What's Missing.

Table 4
Descriptive Statistics and Intercorrelations in the Joint CFA Sample (N = 336)

Test	*M*	*SD*	CIX (r_{adj})	VIX (r_{adj})	NIX (r_{adj})
RIAS Composite Intelligence Index (CIX)	103.91	10.46			
RIAS Verbal Intelligence Index (VIX)	101.81	11.19	.91 (.95)		
RIAS Nonverbal Intelligence Index (NIX)	106.68	10.05	.85 (.92)	.55 (.66)	
RIAS Guess What	49.26	6.97	.79 (.88)	.86 (.91)	.48 (.63)
RIAS Verbal Reasoning	51.01	9.44	.84 (.91)	.92 (.95)	.51 (.66)
RIAS Odd-Item-Out	54.69	6.21	.56 (.70)	.30 (.39)	.73 (.85)
RIAS What's Missing	51.75	8.17	.77 (.87)	.54 (.65)	.85 (.92)
WJ–III COG Verbal Comprehension	95.16	10.87	.74 (.84)	.78 (.86)	.50 (.65)
PPVT	101.70	12.11	.73 (.84)	.75 (.84)	.52 (.67)
WJ III–COG Fluid Reasoning Cluster	99.16	12.36	.63 (.76)	.59 (.70)	.53 (.68)
WJ III–COG Concept Formation	98.11	12.28	.62 (.75)	.58 (.69)	.51 (.66)
WJ III–COG Analysis-Synthesis	100.60	12.48	.48 (.62)	.45 (.56)	.41 (.56)

Note. CFA = confirmatory factor analysis; RIAS = Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003a); WJ–III COG = Woodcock-Johnson Tests of Cognitive Abilities, Third Edition (Woodcock, McGrew, & Mather, 2001b); PPVT = Peabody Picture Vocabulary Test, Fourth Edition (Dunn & Dunn, 2007). Correlations in parentheses (r_{adj}) are adjusted Pearson product-moment correlations corrected for variability in RIAS scores (Guilford & Fruchter, 1978) where the unrestricted RIAS *SD* = 15. Full correlation matrix is available upon request from the first author.

these corrected correlations (r_{adj}) are presented in parentheses in Table 4. Correlations between the VIX and the additional verbal measures were compared with the correlations between the VIX and the additional nonverbal measures. These comparisons were also made with the NIX and the additional measures. To reduce Type I error attributable to multiple comparisons, the alpha level for *t* tests was adjusted using the Bonferroni correction. The corrected alpha level was set at .006 (.05/8) to determine statistical significance. The correlation between the VIX and WJ III COG Verbal Comprehension was larger than the correlations between the VIX and WJ III COG Concept Formation, $t(333) = 4.83$, $p <$ .0001, and the VIX and WJ III COG Analysis-Synthesis, $t(333) = 6.93$, $p < .0001$. Likewise, the correlation between the VIX and PPVT IV was larger than the correlations between the VIX and WJ III COG Concept Formation, $t(333) = 3.87$, $p < .001$, and the VIX and WJ III COG Analysis-Synthesis, $t(333) = 6.02$, $p < .0001$. The correlation between the NIX and WJ III COG Concept Formation was not significantly different than the correlations between the NIX and WJ III COG Verbal Comprehension, $t(333) = 0.16$, $p = .87$, and the NIX and PPVT IV, $t(333) = -0.16$, $p = .87$. Additionally, the correlation between the NIX and WJ III Analysis-Synthesis was not significantly different than the correlations between the NIX and WJ III COG Verbal Comprehension, $t(333) = -1.40$, $p = .16$, and the NIX and PPVT IV, $t(333) = -1.74$, $p = .08$. Results were the same for these comparisons using r_{adj}. In summary, the correlations between the RIAS VIX and measures of crystallized intelligence were strong and higher than were the correlations between the RIAS VIX and measures of fluid intelligence. In contrast, such divergence was not apparent for the RIAS NIX, because its correlations were moderate and similar for both measures of fluid *and* crystallized intelligence.

Joint CFA indicated superior fit for the two-factor model (see Figure 2) compared with the one-factor model, $\Delta\chi^2(1) = 42.11$, $p < .01$. Further evidence for the superior fit of the two-factor model included AIC and ECVI of 83.94 and .25, respectively, for the two-factor model compared with AIC and ECVI of 124.05 and .37, respectively, for the one-factor model. All fit indices for the two-factor model except for χ^2 (a fit index that is sample-size dependent; Tanaka, 1993) indicated it was a good fit to the data, CFI = .99, GFI = .96, TLI = .96, RMSEA = .07, and SRMR = .04. Despite good overall model fit, Figure 2 shows that, whereas the GWH and VRZ subtests had good loadings on the crystallized/verbal intelligence factor, the factor loadings for the OIO and WHM on the fluid/nonverbal intelligence factor were poor and good, respectively.

Predictive and Incremental Validity

Pearson product-moment correlations and r_{adj} for RIAS CIX, VIX, and NIX and academic achievement scores from the WJ–III ACH and NDRT are presented in Table 5. As expected, ability–achievement correlations were biased downward due to restricted range in the present sample. Correlations for RIAS CIX, VIX, and NIX with WJ–III ACH fluency measures and the spelling subtest were all lower than correlations for the other WJ–III ACH academic skills measures and Applied Problems. Also of note was that the CIX and VIX produced higher correlations with achievement measures than did the NIX.

H–MRA results are presented in Table 6 and illustrate the portions of achievement variance accounted for by the RIAS CIX in the first block followed by portions of achievement variance incrementally accounted for by the VIX and NIX in the second block (after CIX variance was partialled out). The change in R^2 from the second block provided the estimate of the combined incremental prediction of VIX and NIX after accounting for achievement predicted by the CIX. Using Cohen's (1988) effect size descriptors for R^2, the CIX provided medium predictive effects for the WJ–III ACH Academic Skills Cluster, Letter–Word Identification, Calculation, and Spelling, and large predictive effects for Applied Problems and Listening Comprehension. CIX predictions of WJ–III ACH Academic Fluency Cluster, Reading Fluency, Math Fluency, and Writing Fluency were small. The CIX provided a medium predictive effect for the NDRT. However, although the incremental predictive contribution of VIX and NIX scores across all achievement measures (except Writing Fluency) were statistically significant ($p < .005$) they all represented small effect sizes. Additional achievement variance accounted for by the

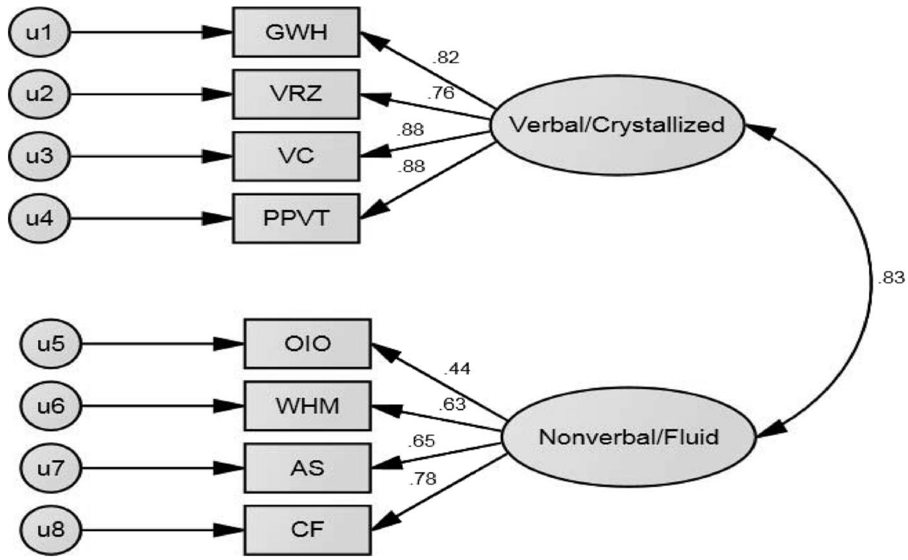

Figure 2. Diagram of joint confirmatory factor analysis. GWH = Reynolds Intellectual Assessment Scales (RIAS; Reynolds & Kamphaus, 2003a) Guess-What; VRZ = RIAS Verbal Reasoning; VC = Woodcock-Johnson Tests of Cognitive Abilities, Third Edition (WJ–III COG; Woodcock, McGrew, & Mather, 2001b) Verbal Comprehension; PPVT = Peabody Picture Vocabulary Test, Fourth Edition (Dunn & Dunn, 2007); OIO = RIAS Odd-Item-Out; WHM = RIAS What's Missing; AS = WJ–III COG Analysis-Synthesis; CF = WJ–III COG Concept Formation.

VIX and NIX beyond the CIX ranged from 0.9% to 7.6% (*Mdn* = 3.2%).

Discussion

The aim of the current study was to examine various aspects of validity of RIAS scores using a sample of adolescents and adults referred for psychological evaluations at a university-based clinic. We sought to clarify the structural and convergent validity of RIAS scores and examine their ability to predict various academic outcomes. In so doing, we endeavored to better inform appropriate interpretive practices with the RIAS, particularly the interpretative weight that should be apportioned to the first-order and second-order factors.

Results from the EFA, using multiple factor extraction criteria, suggested the extraction of only one factor using the four-subtest

Table 5
Descriptive Statistics and Pearson Product-Moment Correlations Between RIAS Composite Scores and Tests of Academic Achievement

				RIAS		
Achievement tests	*n*	*M*	*SD*	CIX *r* (r_{adj})	VIX *r* (r_{adj})	NIX *r* (r_{adj})
WJ–III Achievement Skills Cluster	489	93.39	12.84	.51 (.65)	.57 (.68)	.27 (.39)
WJ–III Letter-Word Identification	496	92.52	11.83	.46 (.60)	.53 (.64)	.23 (.34)
WJ–III Calculation	491	94.02	15.20	.48 (.63)	.49 (.60)	.33 (.47)
WJ–III Spelling	493	93.84	13.69	.33 (.45)	.41 (.51)	.12 (.18)
WJ–III Academic Fluency Cluster	512	89.75	12.85	.31 (.42)	.35 (.45)	.16 (.23)
WJ–III Reading Fluency	516	90.63	12.30	.27 (.38)	.32 (.41)	.13 (.19)
WJ–III Math Fluency	517	88.37	13.29	.17 (.24)	.22 (.28)	.05 (.08)
WJ–III Writing Fluency	513	97.05	11.62	.31 (.43)	.32 (.42)	.21 (.31)
WJ–III Applied Problems	416	92.88	10.69	.61 (.75)	.61 (.72)	.44 (.59)
WJ–III Listening Comprehension	507	96.89	10.07	.65 (.78)	.67 (.77)	.45 (.60)
Nelson–Denny Reading	515	191.55	26.19	.47 (.61)	.51 (.62)	.27 (.38)

Note. RIAS = Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003a); CIX = Composite Intelligence Index; VIX = Verbal Intelligence Index; NIX = Nonverbal Intelligence Index; WJ–III = Woodcock-Johnson Tests of Achievement, Third Edition (Woodcock, McGrew, & Mather, 2001b). Correlations in parentheses (r_{adj}) are adjusted Pearson product-moment correlations corrected for variability in RIAS scores (Guilford & Fruchter, 1978) where the unrestricted RIAS *SD* = 15. WJ–III ACH scores *M* = 100, *SD* = 15. Nelson–Denny Reading normalized scale score *M* = 200, *SD* = 25.

Table 6
Incremental Contribution of Observed RIAS VIX and NIX in Predicting Academic Achievement Composite and Subtest Scores

	Academic Skills Cluster		Academic Fluency Cluster	
Predictor	Variance (%)	Increment[a] (%)	Variance (%)	Increment[a] (%)
RIAS CIX	25.8	25.8*	9.3	9.3*
RIAS Factors (df = 2)[b]	33.4	7.6*	12.4	3.1*
VIX		0.0		0.0
NIX		0.3		0.2

	Letter-Word Identification		Calculation		Spelling	
Predictor	Variance (%)	Increment[a] (%)	Variance (%)	Increment[a] (%)	Variance (%)	Increment[a] (%)
RIAS CIX	21.2	21.2*	23.3	23.3*	10.7	10.7*
RIAS Factors (df = 2)[b]	28.9	7.6*	25.1	1.9*	18.2	7.5*
VIX		0.0		0.2		0.2
NIX		0.2		0.0		0.1

	Reading Fluency		Math Fluency		Writing Fluency	
Predictor	Variance (%)	Increment[a] (%)	Variance (%)	Increment[a] (%)	Variance (%)	Increment[a] (%)
RIAS CIX	7.3	7.3*	2.7	2.7*	9.9	9.9*
RIAS Factors (df = 2)[b]	10.5	3.2*	5.1	2.4*	10.8	0.9
VIX		0.1		0.1		0.0
NIX		0.5		0.0		0.0

	Applied Problems		Listening Comprehension		Nelson Denny Reading Test	
Predictor	Variance (%)	Increment[a] (%)	Variance (%)	Increment[a] (%)	Variance (%)	Increment[a] (%)
RIAS CIX	37.2	37.2*	42.8	42.8*	21.6	21.6*
RIAS Factors (df = 2)[b]	39.3	2.1*	46.8	4.0*	26.1	4.5*
VIX		0.4		0.8		0.0
NIX		0.1		0.3		0.4

Note. RIAS = Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003a); CIX = Composite Intelligence Index; VIX = Verbal Intelligence Index; NIX = Nonverbal Intelligence Index; DSS = .
[a] Unless otherwise indicated, all unique contributions are squared part correlations equivalent to changes in R^2 if this variable was entered last in block entry regression procedure. [b] Partialing out RIAS CIX.
* $p \leq .005$.

RIAS configuration. These results replicated those of Nelson et al. (2007) with a school-age referred sample and Dombrowski et al. (2009) with the standardization sample of the RIAS. The extraction of two factors resulted in highly correlated factors, suggesting the likely presence of a higher order factor. Results from the examination of the hierarchical structure of the RIAS using the Schmid and Leiman (1957) procedure indicated that the higher order, *g* factor accounted for the largest portions of total and common variance. These results also replicated those of Nelson et al. (2007) and Dombrowski et al. (2009).

We further sought to clarify the internal structure of the RIAS by using CFA on a separate data set to avoid capitalization on chance solutions. Examination of the one-factor and two-factor models indicated that the former was a good fit to the data across the majority of fit indices, whereas the latter's fit was slightly more variable across the fit indices. Comparisons of the models indicated no significant differences and a slightly lower AIC for the one-factor model compared with the two-factor model. These results indicate that the more parsimonious one-factor model was a more appropriate depiction of the internal structure of the RIAS than was the two-factor model. The convergence of these results with the EFA results derived from a separate sample provides strong evidence for the interpretation of the RIAS as a one-factor test.

Results were supportive of the convergent validity of the VIX. Correlations of the VIX with scores from other verbal ability measures were strong. Although the VIX was moderately correlated with measures of nonverbal ability, it was less correlated with these measures than with measures of verbal ability. In contrast, the NIX was moderately correlated with measures of nonverbal *and* verbal abilities; these correlations were similar and therefore did not demonstrate divergence. These results are consistent with those of Krach et al. (2009), who found that the NIX was as correlated with measures of *Gc* as it was with measures of *Gf* but that the VIX was more highly correlated with measures of *Gc* than it was with measures of *Gf*. Similarly, Beaujean et al. (2010) found the VIX to be highly correlated with other measures of *Gc* but the NIX to have higher correlations with measures of *Gc* than with measures of *Gf*. Both Krach et al. (2009) and Beaujean et al. (2010) used samples of college students in their studies, with the former examining correlations between the NIX and the WJ–III COG Fluid Reasoning Cluster and the latter between the NIX and the Shipley Institute of Living Scales Abstraction subtest. Taken together, these results support the convergent validity of the VIX but not the NIX. In contrast, Smith et al. (2009) found the NIX to be moderately correlated with the Wechsler Adult Intelligence Scale—Third Edition (WAIS–III) Performance IQ and more highly correlated with the Performance IQ than with the Verbal IQ in a sample of college students diagnosed with LD and/or ADHD. Umphress (2008) and Edwards and Paulin (2007) found similar correlational patterns between the RIAS NIX and the Wechsler Scales Performance IQ in samples of adults with intellectual disabilities and school-age children referred for psychoeducational testing, respectively. These mixed results may be due to the variation in instruments used to investigate the NIX as a measure of fluid intelligence across studies. It should be noted that, thus far, a gold standard for measuring fluid intelligence has yet to be identified.

As suggested by Beaujean et al. (2010), we also conducted a joint CFA of the RIAS with other measures of crystallized and fluid intelligence to better examine the abilities underlying RIAS performance. A common recommendation in CFA is to include at least three observed indicators per hypothesized factor to adequately indentify factors (Kline, 2005). Therefore, including additional measures of crystallized/verbal and fluid/nonverbal intelligence to the two verbal and two nonverbal RIAS subtests should aid in better identifying its underlying constructs. Results indicated that the two-factor model fit better than did the one-factor model. The GWH and VRZ subtests had good loadings on the crystallized intelligence factor, whereas the loadings for the OIO and WHM subtests on the fluid intelligence factor were poor and fair, respectively. Therefore, although the two-factor model was a good fit to the data and a better fit than the one-factor model, the poor factor loading of the OIO subtest indicates that it was not a good measure of fluid intelligence in this sample.

Across the analyses of the current study, the OIO subtest had unexpectedly poor associations on both the general and fluid intelligence factors. OIO loadings were higher and in the fair to good range in analyses of the RIAS standardization data (Dombrowski et al., 2009; Reynolds & Kamphaus, 2003b) and in analyses of independent school-age samples (Beaujean et al., 2009; Nelson et al., 2007). Examination of the mean scores of the current samples indicates higher scores (and lower *SD*) on the OIO subtest compared with the other RIAS subtests (see Table 1). Whereas mean scores on the GWH, VRZ, and WHM were consistent with the standardization sample means, the mean score on the OIO subtest was approximately 0.5 *SD* higher. Reasons for this pattern of scores are unclear. Although speculative, the multiple-choice format of the OIO subtest may have been advantageous for the current sample, one that may have possessed better test taking skills than the general population given their level of educational attainment. On the OIO subtest, examinees are shown five to seven stimuli per page and required to point to the stimulus that does not go with the others. If they respond incorrectly, they are permitted a second chance to provide another response within a 20-s time limit. The current sample may have been better than the general population at using process of elimination to reach a correct response, a strategy that may be more related to test taking skill than it is to general intelligence.

Combined with the EFA and CFA findings, results from the incremental validity analyses provide further evidence for interpretive emphasis of the CIX over the NIX and VIX. The CIX had medium to large predictive effects for higher level academic achievement and medium predictive effects for basic academic achievement. Although the VIX and NIX accounted for additional variance in both basic and higher level achievement once variance attributed to the CIX was accounted for, the magnitude of the predictive variance of the first-order factors was small. These findings are consistent with incremental validity studies of other, and significantly longer, intelligence tests in which the general intelligence factor (full scale score) predicted the largest portion of academic achievement variance and the first-order factors explained minimal to no additional variance (Glutting, Watkins, Konold, & McDermott, 2006; Glutting, Youngstrom, Ward, Ward, & Hale, 1997; Watkins & Glutting, 2000; Watkins, Glutting, & Lei, 2007).

Limitations

The results of this study should be interpreted in light of the following limitations. Several characteristics of the present sample limit generalizability of these results. The ethnic distribution was not representative of the country at large, and the sample was drawn from only one geographic region of the country and therefore not representative of the entire United States. Furthermore, the sample was a specific subset of adolescents and adults with suspected or documented psychological disorders; thus, the generalizability of these findings to the general population of adolescents and adults is unknown. Including a nonclinical control group and examining measurement equivalence across clinical and nonclinical groups would have provided a more rigorous examination of the internal structure of the RIAS because this would directly test whether the same latent variables underlie test scores derived from the different groups, as well as testing whether the metric relationships between the test scores and their corresponding latent variable are equivalent for each group (Bowden et al., 2008). Additionally, comparing the incremental predictive validity of RIAS scores across clinical and nonclinical groups would have been informative because of the possibility that inclusion of a nonclinical group, particularly those without learning problems, may have resulted in different incremental validity of RIAS scores when predicting academic achievement.

Implications for Practice

Results of the current study suggest two major implications for practice. First, those choosing to use the RIAS with similar individuals are urged to place the greatest interpretive weight on the second-order factor (the CIX) rather than the first-order factors (the VIX and NIX). A one-factor model was supported in independent samples with both EFA and CFA procedures and hierarchical analysis indicated that the largest portion of variance was captured by the *g* factor. Additionally, little variance in academic achievement was attributed to the VIX and NIX once the medium to large amounts of achievement variance was accounted for by the CIX. Second, we encourage practitioners to be mindful of the variability of the four RIAS subtests in terms of measurement of *g*. Whereas the GWH and VRZ subtests were found to have good *g*-loadings, the WHM and OIO subtests had fair and poor *g*-loadings, respectively. When considering the results of the current study, it is important to keep in mind that validity is not a quality of tests but of the interpretations and uses of tests. The current results suggest that the RIAS is perhaps most useful as a screener of general intellectual functioning with individuals similar to those in the current study. Future research is needed to determine other valid uses and interpretations of the RIAS.

References

Andrews, J. W. (2007). Test review of Reynolds Intellectual Assessment Scales. *Journal of Psychoeducational Assessment, 25,* 402–408. doi:10.1177/0734282907300381

Arbuckle, J. L. (2009). *Amos 18.0.* Chicago, IL: Smallwaters.

Beaujean, A. A., Firmin, M. W., Michonski, J. D., Berry, T., & Johnson, C. (2010). A multitrait–multimethod examination of the Reynolds Intellectual Assessment Scales in a college sample. *Assessment, 17,* 347–360. doi:10.1177/1073191109356865

Beaujean, A. A., McGlaughlin, S. M., & Margulies, A. S. (2009). Factorial validity of the Reynolds Intellectual Assessment Scales for referred students. *Psychology in the Schools, 46,* 932–950. doi:10.1002/pits.20435

Bowden, S. C., Gregg, N., Bandalos, D., Davis, M., Coleman, C., Holdnack, J. A., & Weiss, L. G. (2008). Latent mean and covariance differences with measurement equivalence in college students with developmental difficulties versus the Wechsler Adult Intelligence Scale III/Wechsler Memory Scale III normative sample. *Educational and Psychological Measurement, 68,* 621–642. doi:10.1177/0013164407310126

Bracken, B. A. (2005). Reynolds Intellectual Assessment Scales and the Reynolds Intellectual Screening Test. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurement yearbook* (pp. 892–894). Lincoln, NE: Buros Institute of Mental Measurements.

Brown, J. I., Fishco, V. V., & Hanna, G. (1993). *Nelson-Denny Reading Test.* Chicago, IL: Riverside.

Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Canivez, G. L., Konold, T. R., Collins, J. M., & Wilson, G. (2009). Construct validity of the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test: Convergent and structural validity. *School Psychology Quarterly, 24,* 252–265. doi:10.1037/a0018030

Carroll, J. B. (1993). *Human cognitive abilities.* Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511571312

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245–276. doi:10.1207/s15327906mbr0102_10

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Curran, P. J., West, S. G., & Finch, S. G. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1,* 16–29. doi:10.1037/1082-989X.1.1.16

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

Dombrowski, S. C., & Mrazik, M. (2008). Test review: Reynolds Intellectual Assessment Scales. *Canadian Journal of School Psychology, 23,* 223–230. doi:10.1177/0829573508324458

Dombrowski, S. C., Watkins, M. W., & Brogan, M. J. (2009). An exploratory investigation of the factor structure of the Reynolds Intellectual Assessment Scales (RIAS). *Journal of Psychoeducational Assessment, 27,* 494–507. doi:10.1177/0734282909333179

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition.* San Antonio, TX: Pearson.

Edwards, O. W., & Paulin, R. V. (2007). Referred students' performance on the Reynolds Intellectual Assessment Scales and the Wechsler Intelligence Scale for Children—Fourth Edition. *Journal of Psychoeducational Assessment, 25,* 334–340. doi:10.1177/0734282907300453

Elliott, R. W. (2004). Reynolds Intellectual Assessment Scales. *Archives of Clinical Neuropsychology, 19,* 325–328. doi:10.1016/j.acn.2003.10.001

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4,* 272–299. doi:10.1037/1082-989X.4.3.272

Frazier, T. W., & Youngstrom, E. A. (2007). Historical increase in the number of factors measured by commercial tests of cognitive ability: Are we overfactoring? *Intelligence, 35,* 169–182. doi:10.1016/j.intell.2006.07.002

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

Glutting, J. J., Youngstrom, E. A., Ward, T., Ward, S., & Hale, R. L. (1997). Incremental efficacy of WISC–III factor scores in predicting achievement: What do they tell us? *Psychological Assessment, 9,* 295–301. doi:10.1037/1040-3590.9.3.295

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Gorsuch, R. L. (2003). Factor analysis. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Vol. 2. Research methods in psychology* (pp. 143–164). Hoboken, NJ: Wiley.

Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). New York, NY: McGraw–Hill.

Guttman, L. (1954). Some necessary and sufficient conditions for common factor analysis. *Psychometrika, 19,* 149–161. doi:10.1007/BF02289162

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185. doi:10.1007/BF02289447

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. doi:10.1080/10705519909540118

Hunsley, J. (2003). Introduction to the special section on incremental validity and utility in clinical assessment. *Psychological Assessment, 15,* 443–445. doi:10.1037/1040-3590.15.4.443

Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment, 15,* 446–455. doi:10.1037/1040-3590.15.4.446

Kaufman, A. S. (1994). *Intelligent testing with the WISC–III.* New York, NY: Wiley.

Kline, R. B. (2005). *Principles and practice of structural equation modeling.* New York, NY: Guilford Press.

Krach, S. K., Loe, S. A., Jones, W. P., & Farrally, A. (2009). Convergent validity of the Reynolds Intellectual Assessment Scales (RIAS) using the Woodcock-Johnson Tests of Cognitive Ability, Third Edition (WJ–III) with university students. *Journal of Psychoeducational Assessment, 27,* 355–365. doi:10.1177/0734282909331749

Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research, 24,* 365–395. doi:10.1207/s15327906mbr2403_6

Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural analysis* (4th ed.). Hillsdale, NJ: Erlbaum.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of samples size for covariance structure modeling. *Psychological Methods, 1,* 130–149. doi:10.1037/1082-989X.1.2.130

Nasser, F., Benson, J., & Wisenbaker, J. (2002). The performance of regression-based variations of the visual scree for determining the number of common factors. *Educational and Psychological Measurement, 62,* 397–419. doi:10.1177/00164402062003001

Nelson, J. M., Canivez, G. L., Lindstrom, W., & Hatt, C. V. (2007). Higher-order exploratory factor analysis of the Reynolds Intellectual Assessment Scales with a referred sample. *Journal of School Psychology, 45,* 439–456. doi:10.1016/j.jsp.2007.03.003

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments & Computers, 32,* 396–402. doi:10.3758/BF03200807

Reynolds, C. R., & Kamphaus, R. W. (2003a). *Reynolds Intellectual Assessment Scales.* Lutz, FL: Psychological Assessment Resources.

Reynolds, C. R., & Kamphaus, R. W. (2003b). *Reynolds Intellectual Assessment Scales professional manual.* Lutz, FL: Psychological Assessment Resources.

Reynolds, C. R., & Kamphaus, R. W. (2005). Introduction to the Reynolds Intellectual Assessment Scales and the Reynolds Intellectual Screening Test. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 461–483). New York, NY: Guilford Press.

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

Schraw, G. (2005). Review of the Reynolds Intellectual Assessment Scales and the Reynolds Intellectual Screening Test. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurement yearbook* (pp. 894–895). Lincoln, NE: Buros Institute of Mental Measurements.

Smith, B. L., McChristian, C. L., Smith, T. D., & Meaux, J. (2009). The relationship of the Reynolds Intellectual Assessment Scales and the Wechsler Adult Intelligence Scale—Third Edition. *Perceptual and Motor Skills, 109,* 30–40. doi:10.2466/pms.109.1.30-40

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. S. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10–39). Newbury Park, CA: Sage.

Umphress, T. B. (2008). A comparison of low IQ scores from the Reynolds Intellectual Assessment Scales and the Wechsler Adult Intelligence Scale—Third Edition. *Intellectual and Developmental Disabilities, 46,* 229–233. doi:10.1352/2008.46:229-233

Velicer, W. F. (1976). Determining the number of components form the matrix of partial correlations. *Psychometrika, 41,* 321–327. doi:10.1007/BF02293557

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A view and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: A festschrift to Douglas Jackson at seventy* (pp. 41–71). Norwell, MA: Kluwer Academic.

Watkins, M. W. (2000). *Monte Carlo PCA for parallel analysis* [Computer software]. State College, PA: Ed & Psych Associates.

Watkins, M. W. (2004). *MacOrtho.* [Computer software]. State College, PA: Ed & Psych Associates.

Watkins, M. W. (2007). *Sescree* [Computer software]. State College, PA: Ed & Psych Associates.

Watkins, M. W., & Glutting, J. J. (2000). Incremental validity of WISC–III profile elevation, scatter, and shape information for predicting reading and math achievement. *Psychological Assessment, 12,* 402–408. doi:10.1037/1040-3590.12.4.402

Watkins, M. W., Glutting, J. J., & Lie, P.-W. (2007). Validity of the Full-Scale IQ when there is significant variability among WISC–III and WISC–IV factor scores. *Applied Neuropsychology, 14,* 13–20. doi:10.1080/09084280701280353

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and over-extraction on principal axis factor analysis with varimax rotation. *Psychological Methods, 1,* 354–365. doi:10.1037/1082-989X.1.4.354

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

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

Zoski, K. W., & Jurs, S. (1996). An objective counterpart to the visual scree test for factor analysis: The standard error scree. *Educational and Psychological Measurement, 56,* 443–451. doi:10.1177/0013164496056003006

Received March 16, 2011
Revision received June 17, 2011
Accepted June 20, 2011 ■