School Psychology Quarterly
2009, Vol. 24, No. 4, 252–265

© 2009 American Psychological Association
1045-3830/09/$12.00 DOI: 10.1037/a0018030

Construct Validity of the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test: Convergent and Structural Validity

Gary L. Canivez
Eastern Illinois University

Timothy R. Konold
University of Virginia

Jason M. Collins and Greg Wilson
Eastern Illinois University

The Wechsler Abbreviated Scale of Intelligence (WASI; Psychological Corporation, 1999) and the Wide Range Intelligence Test (WRIT; Glutting, Adams, & Sheslow, 2000) are two well-normed brief measures of general intelligence with subtests purportedly assessing verbal–crystallized abilities and nonverbal–fluid–visual abilities. With a sample of 152 children, adolescents, and adults, the present study reports meaningful convergent validity coefficients and a latent factor structure consistent with the theoretical intellectual models both tests were constructed to reflect. Consideration of the hierarchical model of intelligence tests and issues regarding test interpretation are presented.

Keywords: construct validity, intelligence, WRIT, WASI, hierarchical structure

Brief assessment of general intelligence may serve a variety of clinical (e.g., screening and reevaluation) and research purposes. Although some have argued that clinical assessments require that a "comprehensive" battery of intelligence tests be administered to fully understand the nature of performance deficits in the context of individual profiles, research on various subtest analyses (i.e., subtest strengths and weaknesses or unique profiles) reveals them to lack sufficient reliability and validity (Canivez & Watkins, 1998, 1999, 2001; Glutting, McDermott, Konold, Snelbaker, & Watkins, 1998; Glutting, McDermott, Watkins, Kush, & Konold, 1997; Macmann & Barnett, 1997; McDermott, Fantuzzo, Glutting, Watkins, & Baggaley, 1992; Watkins & Canivez, 2004). The incremental validity (Haynes & Lench, 2003; Hunsley, 2003; Hunsley & Meyer, 2003) of factor-based scores of more "comprehensive" measures of intelligence has been questioned and found lacking (Glutting, Youngstrom, Ward, Ward, & Hale, 1997; Kahana, Youngstrom, & Glutting, 2002; Konold, 1999; Ree & Earles, 1991; Ree, Earles, & Treachout, 1994; Watkins & Glutting, 2000; Watkins, Glutting, & Lei, 2007; Youngstrom, Kogos, & Glutting, 1999). When estimating general intellectual functioning without regard to examining subtest performance, patterns, or profiles, intelligence tests with fewer subtests may provide more time- and cost-effective yet valid assessment.

Development of brief multidimensional (i.e., verbal *and* nonverbal estimates) measures of intelligence evolved out of the inadequacies of both single index intelligence screeners and short forms developed from comprehensive intelligence tests noted by A. S. Kaufman and Kaufman (1990) and Silverstein (1990). Among the problems identified were spuriously high correlations between short forms and the full-

Gary L. Canivez, Jason M. Collins, and Greg Wilson Department of Psychology, Eastern Illinois University; Timothy R. Konold, Curry School of Education, University of Virginia.

Jason M. Collins and Greg Wilson are now certified school psychologists in Illinois and Arizona, respectively.

Data from Jason M. Collins' and Greg Wilson's specialist in school psychology theses were combined for analyses in the present study.

We thank Peter W. Hartmann for helpful comments on a version of this article, and also thank Daniel Bialick for providing an electronic copy of his dissertation for review.

Correspondence concerning this article should be addressed to Gary L. Canivez, Department of Psychology, 600 Lincoln Avenue, Charleston, IL 61920-3099. E-mail: glcanivez@eiu.edu

length test, that short forms are developed and extracted from norms based on full test administration, and standardization and resulting scores might not correspond if only the short-form subtests are administered in isolation. J. C. Kaufman and Kaufman (2001) argued that development of short forms of comprehensive intelligence tests was unnecessary because multidimensional brief measures of intelligence are readily available and problems associated with short forms are avoided.

Although the Kaufman Brief Intelligence Test (KBIT; A. S. Kaufman & Kaufman, 1990) can be considered the first brief multidimensional test of intelligence, a number of other multidimensional brief intelligence tests have recently been developed and published. These include the Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003), the Wechsler Abbreviated Scale of Intelligence (WASI; Psychological Corporation, 1999), the Wide Range Intelligence Test (WRIT; Glutting et al., 2000), and a recent revision of the KBIT, the Kaufman Brief Intelligence Test—Second Edition (KBIT–2; A. S. Kaufman & Kaufman, 2004). All reflect similar measurement of intelligence based on Spearman's (1927) general intelligence construct, Wechsler's (1958) theory and measurement of intelligence, and hierarchical conceptualizations of intelligence articulated in theories of Carroll (1993, 2003), Cattell (1963), and Horn and Cattell (1966), now commonly referred to as Cattell–Horn–Carroll (CHC) theory.

The WASI (Psychological Corporation, 1999) is a brief measure of general intelligence designed for use with children and adults between the ages of 6 and 89 years. The WASI requires approximately 30 min to administer the full battery and 15 min to administer the abbreviated battery. The full battery consists of four subtests; Vocabulary (V) and Similarities (S) subtests combine to measure VIQ (verbal–crystallized abilities), and Block Design (BD) and Matrix Reasoning (MR) combine to measure PIQ (nonverbal–fluid abilities). The abbreviated battery consists of two subtests, Vocabulary and Matrix Reasoning, which combine to assess general intelligence. WASI subtests are scaled in *T* score units ($M = 50$, $SD = 10$), and the IQ scores are scaled in traditional IQ–standard score units ($M = 100$, $SD = 15$). Four IQ scores are provided for the WASI and include the Full-Scale IQ–Four Subtest (FSIQ-4), Full-Scale IQ–Two Subtest (FSIQ-2), Verbal IQ (VIQ), and Performance IQ (PIQ).

Although exploratory factor analyses of the four WASI subtests are not reported in the manual (Psychological Corporation, 1999), joint exploratory factor analysis (EFA) of WASI subtests with Wechsler Intelligence Scale for Children—Third Edition (WISC–III) subtests ($N = 176$) and with Wechsler Adult Intelligence Scale—Third Edition (WAIS–III) subtests ($N = 248$) were reported. Factor pattern coefficients from oblique rotations provided evidence for the WASI V and S subtests' association with, and assignment to, the verbal comprehension dimension and the WASI BD and MR subtests' association with, and assignment to, the perceptual organization dimension in both samples. Confirmatory factor analyses (CFAs) of WASI subtests provided evidence for superiority of the two-factor over the one-factor model for the total standardization sample ($N = 2{,}245$) and for the various age subgroups (Psychological Corporation, 1999).

Ryan et al. (2003) examined the WASI factor structure using EFA procedures with the adult portion of the standardization sample ($n = 1{,}145$). The four-subtest correlation matrix published in the WASI manual and an independent clinical sample ($N = 201$) of referrals for neuropsychological or cognitive assessments was examined. Results supported subtest assignments of WASI V and S to the verbal dimension and WASI BD and MR to the perceptual dimension. Correlations between the verbal (Factor I) and perceptual (Factor II) dimensions were high (.75 for the standardization sample, .77 for the clinical sample), indicating a higher order dimension (*g*) and the need for consideration of a hierarchical structure. Hierarchical EFA, however, was not examined. Hays, Reas, and Shaw (2002) reported concurrent validity of the WASI and the KBIT. Statistically significant correlations were observed between the WASI Full-Scale IQ and KBIT IQ Composite ($r = .89$), WASI VIQ and KBIT Vocabulary ($r = .84$), and WASI PIQ and KBIT Matrices ($r = .84$) with their sample of 85 psychiatric inpatients.

The WRIT (Glutting et al., 2000) is a multidimensional measure of general intelligence designed for use with children and adults between the ages of 4 and 85 years. The technical manual presents the WRIT as an efficient measure of important multiple abilities that can be used in numerous clinical applications and is not portrayed as a brief or abbreviated measure. The WRIT can be administered in less than 30 min and consists of Verbal Analogies (VA) and Vocabulary (V) subtests that combine to measure Verbal IQ (verbal–crystallized abilities) and the Matrices (M) and Diamonds (D) subtests that combine to measure Visual IQ (nonverbal–fluid abilities). WRIT subtests and IQs are scaled in standard score units ($M = 100$, $SD = 15$). Three IQs are provided by the WRIT and include the General (GIQ), Verbal (VIQ), and Visual (VisIQ).

EFAs with the WRIT standardization sample supported the Vocabulary and Verbal Analogies subtests' associations with assignment to the verbal–crystallized dimension and the Diamonds and Matrices subtests' associations with assignment to the visual–fluid dimension. Moreover, these results replicated across race/ethnicity, education level, sex, and age (Glutting et al., 2000). Similar to results obtained with the WASI, the correlation between the two WRIT factors was .75, reflecting the hierarchical nature of intellectual abilities assessment. Nested model comparisons within a CFA framework supported a correlated two-factor model over a one-factor (*g*) model, and multigroup CFA revealed that the resulting structure replicated across race/ethnicity, education level, sex, and age subgroups (Glutting et al., 2000). Joint factor analyses of the WRIT with the WISC–III and with the WAIS–III reported in the WRIT manual also supported the factorial structure and association of WRIT subtests and, by extension, the construct validity of the WRIT. Shields, Konold, and Glutting (2004) found that the WRIT predicted academic achievement performance equally well across the demographic variables of race/ethnicity, sex, and socioeconomic status using Potthoff (1966) analyses to investigate predictive validity bias.

Investigation of the concurrent validity of the WRIT in a small college sample was also favorable. Bialik (2008) found strong correlations between the WRIT and WASI with a sample of 82 college students. Correlations between similar global scale IQ scores were statistically significant and shared substantial portions of variance: WASI FSIQ–WRIT GIQ ($r = .72$), WASI VIQ–WRIT VIQ ($r = .73$), and WASI PIQ–WRIT VisIQ ($r = .68$). WRIT–WASI subtest correlations ranged from .47 to .73. Mean score comparisons found the WASI to produce generally higher scores than the WRIT, with effect sizes ranging from small to large.

Both the WASI and WRIT have very similar subtests and global scores, are based on a hierarchical model of intelligence, and purport to measure verbal–crystallized and nonverbal–fluid–visual abilities (Stratum II factors). Accordingly, strong correlations would be expected between similar subtests and global composites, and joint factor analysis should reveal similar subtest associations with latent intellectual dimensions. The present study investigated relations between the WASI and WRIT using convergent validity comparisons, joint EFAs, and joint CFAs to investigate the construct validity of both tests. This is the first investigation to jointly consider these two measures within the framework of factor analytic methods.

Method

Participants

The sample ($N = 152$) included children and adolescents (ages 6–17 years; $n = 136$, 89.5%) and adults (ages 18 years and older; $n = 16$, 10.5%), with slightly more females ($n = 84$, 55.3%) than males ($n = 68$, 44.7%). The majority of participants were Caucasian ($n = 142$, 93.4%), with some Hispanic/Latino ($n = 8$, 5.3%) and Asian American ($n = 2$, 1.4%) participants, typical of the rural east-central Illinois geographic region where these data were collected. Participants ranged in age from 6 to 53 years ($M = 11.71$ years, $SD = 6.32$, $Mdn = 9.88$). The majority of the sample ($n = 113$, 74%) consisted of normal volunteers who were not referred for clinical or psychoeducational evaluations. The other 39 participants were referred for special education triennial reevaluations that resulted in the following disability classifications: learning disability ($n = 27$, 17.8%), mental retardation ($n = 6$, 3.9%), speech–language disability ($n = 2$, 1.3%), and one individual each (1%) from within developmental delay, other health impairment, and emotional disability and learning disability. Reevaluation results for one student resulted in the multidisciplinary evaluation team determining that the student was no longer disabled.

Procedure

Child and adolescent participants who were not referred for special education evaluations were obtained through solicitation of volunteers from schools within two independent school districts, and parent permission to participate was obtained. Those referred for psychoeducational reevaluation of suspected disabilities were administered the WASI and WRIT as part of triennial reevaluations following parent informed consent. Adult volunteers were college students enrolled in psychology courses, were administered the WASI and WRIT following their informed consent, and were not evaluated as part of a clinical assessment of educational difficulties. The WASI and WRIT were administered in counterbalanced order to control for potential order effects, and each participant was tested during a single test session. Multiple test administrators were used, and all test administrators were professionally trained in standardized intelligence testing and were either certified school psychologists or school psychologist interns. Data were coded anonymously with no personally identifiable information.

Data Analyses

Subtest scores of the WASI were converted from *T* scores ($M = 50$, $SD = 10$) to standard scores ($M = 100$, $SD = 15$) so that the subtest scores on both the WASI and the WRIT were in the same scale units. Pearson product–moment correlation coefficients were calculated to estimate levels of convergent validity between the various scales of the WRIT and the WASI. Differences between dependent correlation coefficients were tested using Hotelling's (1940) formula for a *t* test (Guilford & Fruchter, 1978, p. 164). Dependent *t* tests for differences between means were calculated between corresponding subtests and composite IQ scales of the two instruments, and the family-wise Type I error rate (i.e., $\alpha = .05$) was adjusted through use of the Bonferroni correction for testing multiple hypotheses ($.05/7 = .007$). Effect size estimates were calculated using Cohen's *d* (Cohen, 1988) and interpreted using his guidelines (0.20 = small, 0.50 = medium, and 0.80 = large effect sizes).

The Pearson product–moment correlation matrix of WASI and WRIT subtest scores was subjected to principal axis EFA with varimax rotation to investigate the orthogonal (uncorrelated) solution and promax rotation to investigate the oblique (correlated) solution using SPSS 13.0 for Macintosh OSX. Principal axis EFA was used to analyze reliable variance from the correlation matrix (Cudeck, 2000; Fabrigar, Wegener, MacCallum, & Strahan, 1999; Tabachnick & Fidel, 2007). Multiple criteria as recommended by Gorsuch (1983) were used to determine the number of factors to retain and included eigenvalues greater than 1 (Guttman, 1954), the scree test (Cattell, 1966), standard error of scree (Zoski & Jurs, 1996), parallel analysis (Horn, 1965), and minimum average partials (MAP; Velicer, 1976; O'Connor, 2000). Parallel analysis and MAP were included as Thompson and Daniel (1996) indicated that they are usually more accurate and are helpful so as not to overfactor (Frazier & Youngstrom, 2007; Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986). The scree test was used to visually determine the optimum number of factors to retain. The standard error of scree was used as programmed by Watkins (2007) as it was reported to be the most accurate objective scree method (Nasser, Benson, & Wisenbaker, 2002). Parallel analysis (see Figure 1) indicated meaningful factors when eigenvalues from the sample data were larger than those produced by random data containing the same number of participants and factors (Lautenschlager, 1989). Random data and resulting eigenvalues for parallel analyses were produced using the Monte Carlo PCA for Parallel Analysis computer program (Watkins, 2000) with 100 replications to provide stable eigenvalue estimates.

Because narrow abilities and broad ability factors are correlated, subtest performance on cognitive abilities tests reflects combinations of both first-order and second-order factors. Because of this, Carroll argued that variance from the higher order factor should be extracted first to residualize the lower order factors, leaving them orthogonal to each other and the higher order factor. Thus, variability associated with a higher order factor is accounted for prior to interpreting variability associated with lower order factors. Statistically, this is achieved through the use of the Schmid and Leiman (1957) procedure, which was recommended by Carroll (1993, 1995, 1997, 2003); Mc-

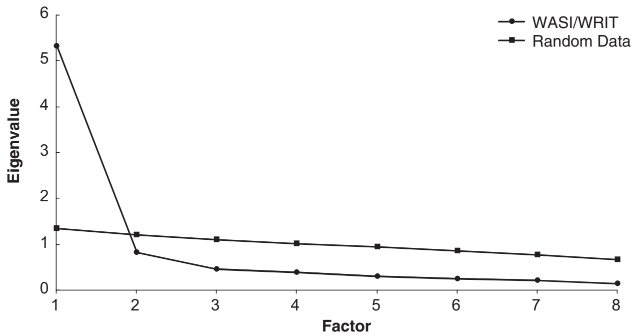

Figure 1. Horn's parallel analysis (HPA) scree plots.

Clain (1996); Gustafsson and Snow (1997); Carretta and Ree (2001); Ree, Carretta, and Green (2003); and Thompson (2004). Following others (Canivez, 2008; Watkins, 2006; Watkins, Wilson, Kotz, Carbone, & Babula, 2006), we further examined promax-rotated results, and resulting factors were orthogonalized using the Schmid and Leiman procedure as programmed in the MacOrtho computer program (Watkins, 2004).

A series of two hierarchically ordered CFAs were also investigated. In contrast to the methods of EFA described above, CFA provides a more restrictive test of the hypothesized factor structure by permitting imposed restrictions on relationships between observed variables and factors. One- and two-factor models were tested. Graphic representation of the one-factor model is illustrated in Figure 2 and the two-factor model is illustrated in Figure 3. The observed WASI and WRIT subtests are enclosed in boxes to differentiate them from the directly unobservable factors and uniqueness terms associated with each observed variable. Each observed variable was modeled to be directly influenced by a single factor as illustrated through the use of single-headed arrows. The curved double-headed arrows reflect the fact that factor correlations were estimated. Parameterization of the model included scaling the factors to one of the observed variables by fixing a single factor loading to one for each factor.

Numerous measures of model fit exist for evaluating the quality of measurement models, most developed under a somewhat different theoretical framework focusing on different components of fit (Browne & Cudeck, 1993; Hu & Bentler, 1995). For this reason, it is generally recommended that multiple measures be considered to highlight different aspects of fit (Tanaka, 1993). As a stand-alone measure of fit, chi square (χ^2) is known to reject trivially misspecified models estimated on large sample sizes (Hu & Bentler, 1995; Kaplan, 1990; Kline, 2005). The chi-square ratio (χ^2/df), however, was used to evaluate stand-alone models. This index tends to be less sensitive to sample size, and values less than 3 (or in some instances 5) are often taken to indicate acceptable models (Kline, 2005). Several additional measures of fit were considered in evaluating model quality. These included the Bentler–Bonett normed fit index (NFI), Tucker–Lewis index (TLI), comparative fit index (CFI), root-mean-square error of approximation (RMSEA), Akaike's information criterion (AIC), and Bayes's information criterion (BIC). These first three measures generally range between 0 and 1.0. Traditionally, values of 0.90 or greater were taken as evidence of good-fitting models (Bentler & Bonett, 1980). However, more recent research suggests that better fitting models produce values around 0.95 (Hu & Bentler, 1999). By contrast, smaller RMSEA, AIC, and BIC values support better fitting models. All models were estimated with the Analysis of Moment Structures (AMOS; Arbuckle & Wothke, 1999) program.

Results

Convergent Validity

Pearson product–moment correlations and descriptive statistics for WASI and WRIT subtests are presented in Table 1 and show the statistically

Table 1
Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test Descriptive Statistics and Correlations

			WASI				WRIT			
Subtests	*M*	*SD*	1	2	3	4	5	6	7	8
WASI										
1. Vocabulary	97.75	17.37								
2. Block Design	97.87	14.49	.57							
3. Similarities	103.81	17.26	.79	.57						
4. Matrix Reasoning	99.81	16.61	.62	.65	.60					
WRIT										
5. Verbal Analogies	101.51	14.77	.69	.51	.70	.53				
6. Vocabulary	100.63	16.42	.83	.54	.74	.57	.71			
7. Matrices	101.45	16.17	.56	.59	.58	.71	.65	.58		
8. Diamonds	100.64	13.92	.51	.66	.55	.62	.51	.53	.62	

Note. $N = 152$. WASI = Wechsler Abbreviated Scale of Intelligence; WRIT = Wide Range Intelligence Test; V = Verbal; NV = Nonverbal. All correlations statistically significant, $p < .0001$.

significant correlations between subtests. Generally higher correlations were observed between similar or identical subtests, illustrating convergent validity for the WASI and WRIT subtests. Correlations between similar global scale IQs were also statistically significant: WASI FSIQ–WRIT GIQ ($r = .86$), WASI VIQ–WRIT VIQ ($r = .84$), and WASI PIQ–WRIT VisIQ ($r = .79$). The WASI VIQ–WRIT VIQ correlation ($r = .84$) was significantly larger than the WASI VIQ–WRIT VisIQ correlation ($r = .64$), $t(149) = 4.80$, $p < .0001$, and significantly larger than the WRIT VIQ–WASI PIQ correlation ($r = .63$), $t(149) = 4.96$, $p < .0001$. The WASI PIQ–WRIT VisIQ correlation ($r = .79$) was significantly larger than the WASI PIQ–WRIT VIQ correlation ($r = .63$), $t(149) = 3.37$, $p < .001$, and significantly larger than the WRIT VisIQ–WASI VIQ correlation ($r = .64$), $t(149) = 3.22$, $p < .001$.

Dependent t tests for differences between means for similar subtests and global composite scores produced three statistically significant differences out of seven comparisons. The WRIT V ($M = 100.63$, $SD = 16.42$) was higher than the WASI V ($M = 97.75$, $SD = 17.37$), $t(151) = 3.54$, $p < .001$, $d = 0.17$. The WRIT D ($M = 100.64$, $SD = 13.92$) was higher than the WASI BD ($M = 97.87$, $SD = 14.49$), $t(151) = 2.93$, $p < .004$, $d = 0.20$. Finally, the WRIT VisIQ ($M = 101.47$, $SD = 16.20$) was higher than the WASI PIQ ($M = 98.95$, $SD = 14.99$), $t(151) = 3.05$, $p < .003$, $d = 0.17$. No other pairwise comparisons were statistically significant.

Exploratory Factor Analysis

EFA results produced a Kaiser–Meyer–Olkin measure of sampling adequacy coefficient of .90, and Bartlett's test of sphericity was 887.52, $p < .0001$. Communality estimates ranged from .55 to .77 ($Mdn = .67$). Given the high communality estimates, the present sample size was judged adequate for factor analysis procedures (Fabrigar et al., 1999; Floyd & Widaman, 1995; MacCallum, Widaman, Zhang, & Hong, 1999). Multiple criteria for determining the number of factors to extract and retain were not in agreement. The criteria of eigenvalues greater than 1, standard error of scree, and Horn's parallel analysis (see Figure 1) indicated that one factor should be extracted, and the scree test (see Figure 1), MAP, and theoretical consideration indicated that two factors might be extracted. One *and* two factors were extracted (Wood, Tataryn, & Gorsuch, 1996) and examined through principal axis factor analysis, and when two factors were extracted, they were rotated using the varimax procedure to achieve an orthogonal (uncorrelated factors) solution and rotated using the promax procedure to achieve an oblique (correlated factors) solution.

Table 2 presents results from EFAs including varimax factor structure coefficients, promax factor structure coefficients, promax factor pattern coefficients, eigenvalues, and the percentage of variance accounted for. All WRIT and WASI subtests produced g loadings exceeding .70, and all were considered "good" according to criteria by A. S. Kaufman (1994). Factor

Table 2
Exploratory Factor Analysis Results for the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test

		Varimax structure coefficient[a]		Promax pattern coefficient[a]		Promax structure coefficient[a]	
Subtest	*g*	Factor I (V)	Factor II (NV)	Factor I (V)	Factor II (NV)	Factor I (V)	Factor II (NV)
WASI							
Vocabulary	.86	**.84**	.37	**.93**	−.02	**.91**	**.68**
Block Design	.74	.36	**.70**	.06	**.74**	**.61**	**.78**
Similarities	.83	**.74**	.43	**.76**	.13	**.85**	**.70**
Matrix Reasoning	.78	.38	**.73**	.08	**.77**	**.65**	**.83**
WRIT							
Verbal Analogies	.78	**.67**	.43	**.66**	.17	**.79**	**.67**
Vocabulary	.84	**.81**	.36	**.90**	−.01	**.89**	**.66**
Matrices	.78	.41	**.70**	.13	**.71**	**.66**	**.81**
Diamonds	.73	.31	**.73**	−.02	**.81**	**.58**	**.79**
Eigenvalue	5.33	0.84					
Percentage of variance							
Common		62.91	6.73				
Cumulative		62.91	69.64				

Note. $N = 152$. WASI = Wechsler Abbreviated Scale of Intelligence; WRIT = Wide Range Intelligence Test; g = general intelligence; V = Verbal; NV = Nonverbal.
[a] Salient factor structure coefficients (> .44) based on Comrey and Lee's (1992) classifications are presented in bold type. Promax-rotated Factor I and Factor II $r = .75$.

structure coefficients for the varimax and promax rotations provided support for the theoretically consistent assignment of WASI and WRIT subtests to the latent factors they represent (Verbal–Crystallized and Nonverbal–Fluid–Visual). In examining the oblique solution, the correlation between Factor I (Verbal–Crystallized Ability) and Factor II (Nonverbal–Fluid Ability) of .75 was very high and indicated the presence of a higher order factor (*g* [general intelligence]). Given this very high factor correlation, it is best to consider the hierarchical structure of the WASI and WRIT that is consistent with their development and theoretical representation. Salient promax factor structure coefficients for all subtests on both factors also indicated the need to consider the oblique nature of the first-order dimensions and need for consideration of higher order relations.

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–Crystallized and Nonverbal–Fluid–Visual) factors. The second-order (general) factor accounted for 53.83% of the total variance and 77.75% of the common variance. The general factor also accounted for between 46% and 62% of individual subtest variability. At the first-order level, Factor I (Verbal–Crystallized) accounted for an additional 8.24% of the total variance and 11.91% of the common variance, and Factor II (Nonverbal–Fluid–Visual) accounted for an additional 7.16% of the total variance and 10.34% of the common variance. The first- and second-order factors combined to measure 69.24% of the variance in WASI and WRIT scores, resulting in 30.76% unique variance (combination of specific and error variance).

Confirmatory Factor Analysis

Results of CFA revealed that the one-factor model (see Figure 2) did not provide a good representation of these data when gauged in relation to either historical or contemporary standards for good-fitting models (see Table 4). By contrast, all measures of fit for the two-factor model were found to exceed recently recommended thresholds for good fit (i.e., > 0.95; NFI = 0.96, TLI = 0.97, and CFI = 0.98), with information-based indices

Table 3
Factor Structure Coefficients and Variance Sources for the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test Subtests Based on the Orthogonalized Higher Order Factor Model

	General intelligence		Verbal		Nonverbal			
Subtest	b	$\%S^2$	b	$\%S^2$	b	$\%S^2$	h^2	u^2
WASI								
Vocabulary	.79	62	.46	21	−.01	0	.82	.18
Block Design	.69	48	.03	0	.37	13	.61	.39
Similarities	.77	59	.38	14	.06	0	.73	.27
Matrix Reasoning	.73	53	.04	0	.38	15	.68	.32
WRIT								
Verbal Analogies	.72	52	.33	11	.08	1	.64	.36
Vocabulary	.77	59	.44	19	.00	0	.78	.22
Matrices	.73	53	.06	0	.36	13	.66	.34
Diamonds	.68	46	−.01	0	.39	15	.61	.39
Total $\%S^2$	53.83		8.24		7.16		69.24	30.76
Common $\%S^2$	77.75		11.91		10.34		—	—

Note. WASI = Wechsler Abbreviated Scale of Intelligence; WRIT = Wide Range Intelligence Test; b = factor structure coefficient (loading); $\%S^2$ = percentage variance; h^2 = communality; u^2 = uniqueness.

(i.e., AIC and BIC) further demonstrating appreciably better fit for the two-factor model (see Table 4). Squared multiple correlations are presented above the observed variable boxes in Figure 2.

Standardized model parameter estimates (i.e., factor loadings, factor correlations, and squared multiple correlations) for the two-factor model are shown in Figure 3. All factor loadings and correlations were large and statistically significant. In addition, the squared multiple correlations (shown above the observed variable boxes in Figure 3) were appreciable and indicated that the factors accounted for meaningful portions of observed score variance. The correlation between factors of .80 revealed that the two factors share meaningful portions of variance.

Discussion

The present study examined relations between the WASI and WRIT through several construct validity methods. Results indicated that both the WASI and WRIT appeared to be measuring the same constructs, with considerable shared variance for the global IQ measures (74%), verbal–crystallized measures (71%), and nonverbal–fluid–visual measures (62%). These are similar to results found by Bialik (2008) with the WASI and WRIT and similar to coefficients found by Hays et al. (2002) with the WASI and KBIT. Correlations between the WASI and WRIT in the present study were higher than those obtained by Bialik. Correlations between the WASI VIQ and WRIT VIQ and between the WASI PIQ and WRIT VisIQ were significantly higher than correlations across domains (i.e., verbal–crystallized and nonverbal–visual–fluid correlations). This was also observed in correlations reported by Bialik. This provides additional evidence of convergent validity for the WASI and WRIT.

Mean scores for global and subtest scores in the present study were quite similar and did not meaningfully differ. Whereas Bialik (2008) found WASI scores to be generally higher than WRIT scores with small to large effect sizes, the present study found the opposite with WRIT scores slightly higher than WASI scores when statistically significant differences were observed. Effect sizes for these mean differences in the present study, however, were small and indicated that they were not practically or clinically meaningful. It is likely that such mean differences are due to sampling error. As such, it is likely that the WRIT and WASI may be used interchangeably in estimating general intelligence. Convergent validity for both measures was strongly supported.

Joint EFAs also provided strong construct validity evidence that the WRIT V and VR and WASI V and S subtests are measuring a similar

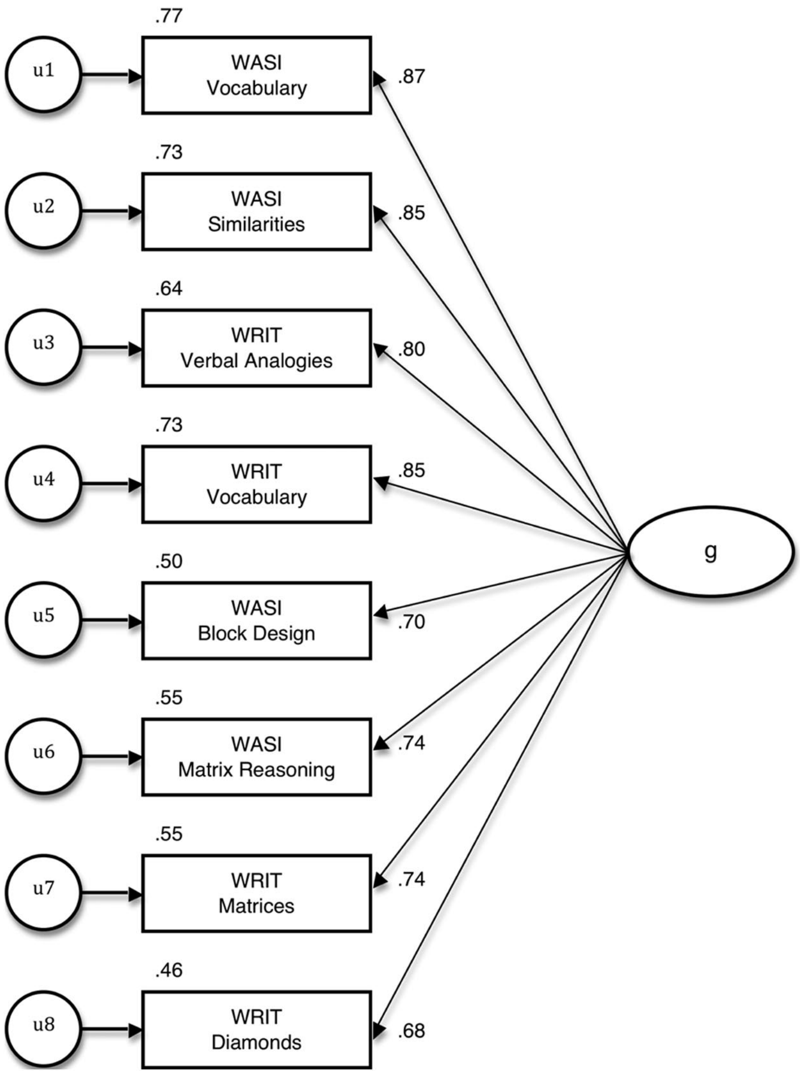

Figure 2. One-factor model with confirmatory factor analysis (CFA) estimated standardized coefficients (squared multiple correlations shown above the observed variable boxes). WASI = Wechsler Abbreviated Scale of Intelligence; WRIT = Wide Range Intelligence Test.

verbal–crystallized dimension. Strong evidence was present for the WRIT D and M and WASI BD and MR subtests measuring a similar nonverbal–fluid–visual dimension. These results are also similar to those reported in the respective test manuals (Glutting et al., 2000; Psychological Corporation, 1999). The verbal–crystallized and nonverbal–fluid–visual factor correlatiosn of .75 from EFA and .80 from CFA in the present study are very close or identical to the factor correlations obtained for the WRIT standardization sample (r = .75; Glutting et al., 2000) and WASI (r = .75 [WASI adult standardization subsample] and .77 [adult clinical sample]; Ryan et al., 2003). These high factor correlations reflect the redundancy of measurement between the verbal–crystallized and nonverbal–fluid–visual factors that defines the higher order general intelligence dimension (g). This redundancy is common among intellectual tests (Canivez, 2008; Carroll, 1993,

Table 4
Confirmatory Factor Model Fit Statistics

Model	χ^2	*df*	*p*	NFI	TLI	CFI	RMSEA	AIC	BIC
One-factor	119.61	20	$<.001$	0.87	0.84	0.89	0.18	151.61	199.99
Two-factor	38.45	19	.005	0.96	0.98	0.98	0.08	72.45	74.60

Note. NFI = normed fit index; TLI = Tucker–Lewis index; CFI = comparative fit index; RMSEA = root-mean-square error of approximation; AIC = Akaike information criterion; BIC = Bayes's information criterion.

2003; Macmann & Barnett, 1997; Watkins, 2006; Watkins et al., 2006).

Consistent with Jensen's (1998) observations, as well as recent investigations of other measures of intelligence such as the Wechsler Intelligence Scale for Children—Fourth Edition (WISC–IV; Wechsler, 2003), Stanford–Binet Intelligence Scales (Roid, 2003a, 2003b, 2003c), and Reynolds Intellectual Assessment Scales (Reynolds & Kamphaus, 2003), most variability measured by the WASI and WRIT corresponds to the second-order general intelli-

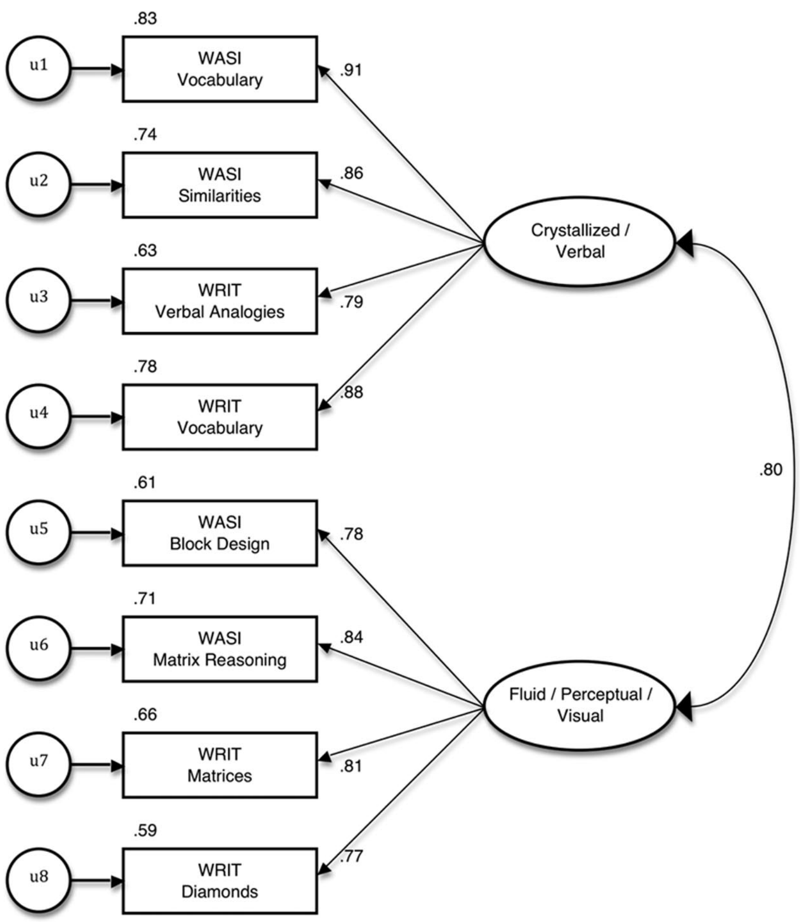

Figure 3. Two-factor model with confirmatory factor analysis (CFA) estimated standardized coefficients (squared multiple correlations shown above the observed variable boxes). WASI = Wechsler Abbreviated Scale of Intelligence; WRIT = Wide Range Intelligence Test.

gence factor (*g*) or Stratum III of the CHC model, and smaller portions of variance are attributed to the first-order factors or CHC Stratum II (Canivez, 2008; Carroll, 1993, 2003; Nelson, Canivez, Lindstrom, & Hatt, 2007; Watkins, 2006; Watkins et al., 2006). As such, interpretation of the WASI and WRIT should primarily reside at the FSIQ–GIQ levels. This is not a surprising finding as both measures were constructed to measure general intellectual abilities.

These results present a bit of a dilemma for overall consideration of construct validity. While the WRIT and WASI subtests appear to measure theoretically consistent verbal–crystallized and nonverbal–fluid–visual dimensions *and* the two-factor model fits these data better than a one-factor model (also observed in the respective standardization samples), the correlated nature of these factors *and* the greatest proportion of variance of WASI and WRIT scores at the CHC Stratum III (general intelligence) level suggest interpretation should primarily reside at the WRIT GIQ and WASI FSIQ level. This glass half-full versus glass half-empty situation cannot be adequately answered or resolved from a structural validity or internal structure perspective (EFA or CFA).

What may be of greater importance in resolving the relative importance of higher versus lower order interpretations will be investigation of WRIT and WASI scores relative to external criteria. Such is the case in investigations of incremental validity (Haynes & Lench, 2003; Hunsley, 2003; Hunsley & Meyer, 2003). If external criteria such as academic achievement in reading, mathematics, and writing are significantly predicted by the first-order or CHC Stratum II factors (verbal–crystallized and nonverbal–fluid–visual dimensions) *after* accounting for prediction by the higher order CHC Stratum III factor (general intelligence), then those lower order dimensions may be of importance for predictive and interpretive purposes. Incremental validity studies of more "comprehensive" measures of intelligence, however, have not supported many first-order factors or CHC Stratum II factor-based scores (Glutting, Youngstrom, et al., 1997; Kahana et al., 2002; Konold, 1999; Ree & Earles, 1991; Ree et al., 1994; Watkins & Glutting, 2000; Watkins et al., 2007; Youngstrom et al., 1999). It may be that even the longer, more "comprehensive" measures of intelligence will require additional subtests to capture sufficient variance at the first-order or CHC Stratum II level to provide additional meaningful prediction of academic achievement beyond the second-order CHC Stratum III general intelligence factor. Unfortunately, this will further extend assessment time because of longer tests of cognitive abilities, so such procedures may not be cost or time effective.

As with all studies, limitations of the sample or design serve to qualify results. The present study does not have a large or ethnically diverse sample and should not be considered representative of the overall population. Geographic location also limits generalization. The present sample also includes proportionally more students with learning disability than typically found in the population. Accordingly, generalizations should be tempered by this small, rural, regional sample. However, results with this restricted sample were consistent with larger, representative samples obtained in the standardization of the WRIT and WASI and supportive of the construct validity of both the WRIT and WASI.

References

Arbuckle, J. L., & Wothke, W. (1999). *AMOS 4.0 user's guide.* Chicago: Small Waters.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Bialik, D. (2008). Brief measures of cognitive abilities: A concurrent validity investigation of the WRIT and WASI. *Dissertation Abstracts International: Section B. Sciences and Engineering, 68*(12), 8443.

Browne, M. W., & Cudeck, R. (1993). Alternative 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. (2008). Hierarchical factor structure of the Stanford–Binet Intelligence Scales—Fifth Edition. *School Psychology Quarterly, 23,* 533–541.

Canivez, G. L., & Watkins, M. W. (1998). Long-term stability of the Wechsler Intelligence Scale for Children—Third Edition. *Psychological Assessment, 10,* 285–291.

Canivez, G. L., & Watkins, M. W. (1999). Long-term stability of the Wechsler Intelligence Scale for Children—Third Edition among demographic subgroups: Gender, race, and age. *Journal of Psychoeducational Assessment, 17,* 300–313.

Canivez, G. L., & Watkins, M. W. (2001). Long-term stability of the Wechsler Intelligence Scale for Children—Third Edition among students with disabilities. *School Psychology Review, 30,* 361–376.

Carretta, T. R., & Ree, J. J. (2001). Pitfalls of ability research. *International Journal of Selection and Assessment, 9,* 325–335.

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

Carroll, J. B. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research, 30,* 429–452.

Carroll, J. B. (1997). Theoretical and technical issues in identifying a factor of general intelligence. In B. Devlin, S. E. Fienberg, D. P. Resnick, & K. Roeder (Eds.), *Intelligence, genes, and success: Scientists respond to the bell curve* (pp. 125–156). New York: Springer-Verlag.

Carroll, J. B. (2003). The higher stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 5–21). New York: Pergamon Press.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54,* 1–22.

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

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.

Cudeck, R. (2000). Exploratory factor analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of multivariate statistics and mathematical modeling* (pp. 265–296). New York: Academic Press.

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.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7,* 286–299.

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.

Glutting, J., Adams, W., & Sheslow, D. (2000). *Wide Range Intelligence Test.* Wilmington, DE: Wide Range.

Glutting, J. J., McDermott, P. A., Konold, T. R., Snelbaker, A. J., & Watkins, M. W. (1998). More ups and downs of subtest analysis: Criterion validity of the DAS with an unselected cohort. *School Psychology Review, 27,* 599–612.

Glutting, J. J., McDermott, P. A., Watkins, M. W., Kush, J. C., & Konold, T. R. (1997). The base rate problem and its consequences for interpreting children's ability profiles. *School Psychology Review, 26,* 176–188.

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.

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

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

Gustafsson, J.-E., & Snow, R. E. (1997). Ability profiles. In R. F. Dillon (Ed.), *Handbook on testing* (pp. 107–135). Westport, CT: Greenwood Press.

Guttman, L. (1954). Some necessary and sufficient conditions for common factor analysis. *Psychometrica, 19,* 149–161.

Haynes, S. N., & Lench, H. C. (2003). Incremental validity of new clinical assessment methods. *Psychological Assessment, 15,* 456–466.

Hays, J., Reas, D., & Shaw, J. (2002). Concurrent validity of the Wechsler Abbreviated Scale of Intelligence and the Kaufman Brief Intelligence Test among psychiatric inpatients. *Psychological Reports, 90,* 355–359.

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

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligence. *Journal of Educational Psychology, 57,* 253–270.

Hotelling, H. (1940). The selection of variates for use in prediction, with some comments on the general problem of nuisance parameters. *Annals of Mathematical Statistics, 11,* 271–283.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

Hunsley, J. (2003). Introduction to the special section on incremental validity and utility in clinical assessment. *Psychological Assessment, 15,* 443–445.

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

Jensen, A. R. (1998). *The* g *factor.* Westport, CT: Praeger.

Kahana, S. Y., Youngstrom, E. A., & Glutting, J. J. (2002). Factor and subtest discrepancies on the Differential Ability Scales: Examining prevalence and validity in predicting academic achievement. *Assessment, 9,* 82–93.

Kaplan, D. (1990). Evaluating and modifying covariance structure models: A review and recommendation. *Multivariate Behavioral Research, 25,* 137–155.

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

Kaufman, A. S., & Kaufman, N. L. (1990). *Manual for the Kaufman Brief Intelligence Test.* Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test—Second Edition.* Circle Pines, MN: American Guidance Service.

Kaufman, J. C., & Kaufman, A. S. (2001). Time for the changing of the guard: A farewell to short forms of intelligence tests. *Journal of Psychoeducational Assessment, 19,* 245–267.

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

Konold, T. R. (1999). Evaluating discrepancy analyses with the WISC–III and WIAT. *Journal of Psychoeducational Assessment, 17,* 24–35.

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

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4,* 84–99.

Macmann, G. M., & Barnett, D. W. (1997). Myth of the master detective: Reliability of interpretations for Kaufman's "Intelligent Testing" approach to the WISC–III. *School Psychology Quarterly, 12,* 197–234.

McClain, A. L. (1996). Hierarchical analytic methods that yield different perspectives on dynamics: Aids to interpretation. *Advances in Social Science Methodology, 4,* 229–240.

McDermott, P. A., Fantuzzo, J. W., Glutting, J. J., Watkins, M. W., & Baggaley, A. R. (1992). Illusions of meaning in the ipsative assessment of children's ability. *The Journal of Special Education, 25,* 504–526.

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.

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

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.

Potthoff, F. R. (1966). *Statistical aspects of the problem of bias in psychological tests* (Institute of Statistics Mimeo Series No. 479). Chapel Hill: University of North Carolina Department of Statistics.

Psychological Corporation. (1999). *Wechsler Abbreviated Scale of Intelligence.* San Antonio, TX: Author.

Ree, M. J., Carretta, T. R., & Green, M. T. (2003). The ubiquitous role of *g* in training. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 262–274). New York: Pergamon Press.

Ree, M. J., & Earles, J. A. (1991). Predicting training success: Not much more than *g*. *Personnel Psychology, 44,* 321–331.

Ree, M. J., Earles, J. A., & Treachout, M. S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology, 79,* 518–524.

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

Roid, G. (2003a). *Stanford–Binet Intelligence Scales: Fifth edition.* Itasca, IL: Riverside Publishing.

Roid, G. (2003b). *Stanford–Binet Intelligence Scales: Fifth Edition. Examiner's manual.* Itasca, IL: Riverside Publishing.

Roid, G. (2003c). *Stanford–Binet Intelligence Scales: Fifth Edition. Technical manual.* Itasca, IL: Riverside Publishing.

Ryan, J. J., Carruthers, C. A., Miller, L. J., Souheaver, G. T., Gontkovsky, S. T., & Zehr, M. D. (2003). Exploratory factor analysis of the Wechsler Abbreviated Scale of Intelligence (WASI) in adult standardization and clinical samples. *Applied Neuropsychology, 10,* 252–256.

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

Shields, J., Konold, T. R., & Glutting, J. J. (2004). Validity of the Wide Range Intelligence Test: Differential effects across race/ethnicity, gender, and educational level. *Journal of Psychoeducational Assessment, 22,* 287–303.

Silverstein, A. B. (1990). Short forms of individual intelligence tests. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 3–11.

Spearman, C. (1927). *The abilities of man.* New York: Cambridge University Press.

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

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.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, DC: American Psychological Association.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement, 56,* 197–208.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 31,* 321–327.

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: Author.

Watkins, M. W. (2004). *MacOrtho* [Computer software]. State College, PA: Author.

Watkins, M. W. (2006). Orthogonal higher order structure of the Weschsler Intelligence Scale for Children—Fourth Edition. *Psychological Assessment, 18,* 123–125.

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

Watkins, M. W., & Canivez, G. L. (2004). Temporal stability of WISC–III subtest composite: Strengths and weaknesses. *Psychological Assessment, 16,* 133–138.

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.

Watkins, M. W., Glutting, J. J., & Lei, 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.

Watkins, M. W., Wilson, S. M., Kotz, K. M., Carbone, M. C., & Babula, T. (2006). Factor structure of the Wechsler Intelligence Scale for Children—Fourth Edition among referred students. *Educational and Psychological Measurement, 66,* 975–983.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Williams & Wilkins.

Wechsler, D. (2003). *Manual for the Wechsler Intelligence Scale for Children—Fourth Edition.* San Antonio, TX: Psychological Corporation.

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.

Youngstrom, E. A., Kogos, J. L., & Glutting, J. J. (1999). Incremental efficacy of Differential Ability Scales factor scores in predicting individual achievement criteria. *School Psychology Quarterly, 14,* 26–39.

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.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 117,* 253–269.