Exploring the Multidimensional Structure of the WASI-II: Further Insights from Schmid-Leiman Higher-Order and Exploratory Bifactor Solutions

Brittany McGeehan, B.A., Nadine Ndip, B.A,, and Ryan J. McGill, Ph.D.

Abstract

The Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II; Wechsler, 2011a) is a versatile and widely utilized brief intelligence measure by assessment psychologists in a variety of clinical settings. Despite the fact that a hierarchical measurement model is implied in the WASI-II Manual (Wechsler, 2011b) and the scores provided by the instrument, a hierarchical measurement model was not specified in the structural validation studies conducted by the test publisher. Due to this shortcoming, two exploratory factor analytic methods (Schmid-Leiman [SL] and exploratory bifactor analysis [EBFA]) were utilized to disclose the higher-order structuring of WASI-II variables. Results from both solutions indicated that the WASI-II provides users with a strong measurement of general intelligence and suggest caution when interpreting beyond that dimension despite the provision of additional factor-level indices. Implications for clinical interpretation and appropriate use(s) of the WASI-II and other related brief measures are discussed.

Introduction

In contrast with previous screeners and short-forms (e.g., Silverstein, 1990) that were designed to provide users with a single estimate of general intelligence (i.e., FSIQ), contemporary brief intellectual measures now appraise cognitive functioning from a multidimensional perspective, yielding multiple first-order indexes and other related part scores in addition to a global FSIQ composite (Kaufman & Kaufman, 2001; Pierson, Kilmer, Rothlisberg, & McIntosh, 2012). These include the Kaufman Brief Intelligence Test-Second Edition (KBIT-2; Kaufman & Kaufman, 2004), the Reynolds Intellectual Assessment Scales-Second Edition (RIAS-2; Reynolds & Kamphaus, 2015), and the Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II; Wechsler, 2011a). Traditionally these measures have been utilized in clinical practice settings for screening and research purposes.

Although it is commonly argued that clinical assessments require that a "comprehensive" battery of intelligence be administered in order to best understand an individual's profile of cognitive strengths and weaknesses, recent structural validity investigations have revealed conflicting factor structures from those reported in the technical manuals of many contemporary cognitive measures (Canivez & McGill, 2016; Canivez, Watkins, & Dombrowski, 2016a, 2016b; Dombrowski, McGill, & Canivez, 2016; McGill, 2016; McGill & Canivez, 2016; McGill & Spurgin, 2015), calling into question their potential clinical/diagnostic utility. Due to these

limitations, practitioners have been encouraged more recently (Canivez, 2013; Frazier & Youngstrom, 2007; Yates & Taub, 2003; Youngstrom, 2008) to consider the potential cost/benefit advantages afforded by brief measures as a supplement to, or replacement for, more comprehensive measures of intellectual functioning.

However, before adopting brief measures for more widespread use in clinical practice, practitioners must have confidence in the underlying construct meaning for the scores provided on these measures. According to Kane (2013), "the validity of a proposed interpretation or use depends on how well the evidence supports the claims being made" (p. 1). At the most fundamental level, it must be demonstrated that test scores accurately reflect legitimate individual differences. Unfortunately, the validation procedures employed by test publishers for brief measures often do not match the implied theoretical structures for these instruments. For example, Dombrowski and Mrazik (2008) noted that although a hierarchical structure is implied for the RIAS, Reynolds and Kamphaus (2003) failed to investigate that structure in the Professional Manual. Subsequent hierarchical factor analytic investigations (Dombrowski, Watkins, & Brogan, 2009; Nelson & Canivez, 2012; Nelson, Canivez, Lindstrom, & Hatt, 2007) questioned the integrity of the lower-order part scores (Verbal/Nonverbal) provided on the RIAS and did not support confidant clinical interpretation beyond the global second-order FSIQ composite. Although frequently used by clinical, forensic, and school psychologists, the WASI-II provides another relevant example of a brief instrument which suffers from similar structural validation deficiencies.

Recently revised, the WASI-II (Wechsler, 2011a) is a brief measure of general intelligence designed for use with children and adults between the ages of 6-90 years. The WASI-II *Manual* (Wechsler, 2011b) indicates that the WASI-II was designed to measure a hierarchically ordered general intellectual ability factor (i.e., *g*) along with lower-order Verbal and Performance factors. Although users are advised to focus most of their interpretive weight on the FSIQ, additional consideration of performance on the Verbal (VC) and Performance (PR) dimensions is encouraged. For instance, the manual suggests that discrepant performance across the scales may be clinically noteworthy (p. 148) and base rates for observed differences are reported in a supplementary table (Table B.3, p. 194). Although extensive exploratory (EFA) and confirmatory (CFA) analyses supporting the presence of two factors are reported in the manual, considerable problems remain.

Issues with WASI-II Validation

Inexplicably, the publisher failed to use the data obtained from the WASI-II normative sample in their EFA analyses. Instead, the EFA results reported in the manual were extrapolated from analyses conducted with participants in the WASI-II/WISC-IV (*N* = 201) and WASI-II/WAIS-IV (*N* = 182) correlation studies. As described in the manual, the four WASI-II subtests were substituted for corresponding measures in the WISC-IV and WAIS-IV; EFA then commenced using the 10 subtest core battery configurations for those instruments with a forced four-factor extraction consistent with Wechsler Scale theory at that time (Beal, Willis, & Dumont, 2013). The salient and theoretically consistent loading coefficients that were observed for the four WASI-II measures on the VC and PR factors were interpreted as supporting the two-factor model hypothesized for the WASI-II. Although similar procedures were used to validate the previous version of the instrument, they are a clear departure from factor analytic best practice (Fabrigar, Wegener, & MacCallum, 1999; Preacher & MaCallum, 2003; Thompson & Daniel, 1996). It was argued that these procedures were needed because the data obtained from

the WASI-II normative sample "may not be sufficient to define the potential factors and may result in an *under-estimation* [emphasis added] of the number of factors composing the scale" (Wechsler, 2011b, p. 124). However, this justification is questionable given the fact that various iterations of the Wechsler Scales and other related cognitive batteries have long contained factors produced from only two indicators. Structural validity studies for these instruments commenced without resorting to such procedures to promote identification of hypothesized latent dimensions. Furthermore, even if one were to accept the veracity of simply extrapolating the WASI-II structure from that of the full WISC-IV/WAIS-IV test batteries, additional problems would remain. One of these concerns relates to the models that were employed in the EFA and CFA analyses.

When conducting exploratory analyses, the publisher relied solely on the principal axis factoring (PAF) method with promax rotation to elucidate the structure of the WISC-IV/WAIS-IV. The test author correctly employed an oblique rotation under the assumption of correlated factors. Although an oblique rotation is necessary, it may not be singularly sufficient and an additional step is required. According to Gorsuch (1983), higher-order factors are implicit in all oblique rotations; consequently, he recommended that second-order factors be extracted and examined. Unfortunately, the publisher elected to forego second-order factor analysis in the EFA procedures they employed.

According to Carroll (2003), all cognitive measures are composed of reliable variance that is attributable to a higher-order general factor, reliable variance that is attributable to first-order group factors, and error variance. Due to this multidimensionality, Carroll argued that variance from the higher-order factor must be extracted first to residualize the lower-order factors, leaving them orthogonal to the higher-order dimension. Thus, variability associated with a higher-order factor is accounted for before interpreting variability associated with lower-order factors, resulting in variance being apportioned to higher-order and lower-order dimensions. To accomplish this task, Carroll (1993; 1995) recommended second-order factor analysis of first-order factor correlations followed by a Schmid-Leiman transformation (SL; Schmid & Leiman, 1957). Also missing from the WASI-II Manual were empirical support for EFA factor extraction, proportions of variance accounted for by the hypothesized second-order factor and the two first-order factors, second-order subtest loadings, subtest specificity estimates, and model-based reliability estimates including omega coefficients (ω; Canivez, 2016; Reise, 2012; Rodriguez, Reise, & Haviland, 2016). The body of literature on factor analysis methodology (e.g., Carroll, 1993, 1995, 2003; Gorsuch, 1983; McClain, 1996; Carretta & Ree, 2001; Thompson, 2004) and model-based reliability (e.g., Reise, 2012; Reise, Bonifay, & Haviland, 2013) recommends the inclusion of this information because it assists test users in determining how the instrument should be interpreted.

EFA results in the manual were additionally supported by CFA results that compared competing one and two-correlated factors models with the four core subtests. Of the two first-order models, the oblique two-factor model was the best fitting, though model loading coefficients and other important information such as the correlation between the latent factors were not disclosed. Ironically, dimension identification did not appear to be as much of a concern (as per the EFA analyses) to the publisher when specifying an under-identified model (oblique two-factors) in CFA. Again, higher-order structure was not explicated in these analyses though it should be noted that specification of a second-order factor in a measurement model that is empirically under-identified is only possible by constraining the second-order loadings to be equal, a procedure that some researchers find to be problematic (Beaujean, McGlaughlin, &

Margulies, 2009; Brown, 2016). Furthermore, even if the necessary equality constraint is induced, model fit will remain equivalent to the previously specified model (correlated factors), further degrading its empirical utility.

Nevertheless, exclusive reliance on EFA/CFA correlated factors models does not substantiate the implied structure of the measurement instrument. Because the WASI-II is organized into three levels with subtest scores being combined to form the VC and PR factor scores and the FSIQ composite, its implied structures is hierarchical (Kranzler & Keith, 1999). According to McClain (1996), it is a mistake to interpret a second-order factor on the basis of first-order dimensions. Due to these limitations, a case can be made that primary interpretive emphasis in understanding the latent structure of the WASI-II should be placed on EFA procedures that permit higher-order factor analysis and a more thorough investigation of the implied WASI-II theoretical structure (Dombrowski, Watkins, & Brogan, 2009).

Limitations of Previous Research

Since its publication over 15 years ago, the WASI has been the focus of several factor analytic investigations (e.g., Axelrod, 2002; Canivez, Konold, Collins, & Wilson, 2009; Ryan et al., 2003). While many of these studies have supported the publisher suggested two-factor Verbal/Performance model, the lone EFA investigation (Canivez et al., 2009) that utilized higher-order modeling procedures is a notable exception. In a conjoint EFA/CFA using measures from both the WASI-II and the Wide-Range Intelligence Test (WRIT; Glutting, Adams, & Sheslow, 2000), the researchers found no support for the extraction of two factors in EFA, and when two factors were forced followed by the SL procedure, the residualized first-order factors accounted for trivial proportions of total and common variance in WASI/WRIT subtests when compared to the variance accounted for by the second-order *g* factor. To date, the higher-order structure of the WASI has yet to be explored on its own suggesting our understanding of the relationship of WASI/WASI-II variables is presently unknown.

Purpose and Goals of the Current Study

The goal of this study is to disclose the higher-order structure of the WASI-II across the age span (ages 9-90) using two EFA methods. As recommended by Carroll (1993, 1995), the first method used the SL orthognalization procedure (Schmid & Leiman, 1957) which posits an indirect hierarchical measurement model similar to Carroll's three-stratum model wherein the general factor is conceptualized as a superordinate factor produced from the correlations between the first-order factors. The SL procedure has a history of use in cognitive ability literature. As previously noted, it helps to uncover simple structure in the presence of correlated traits by partialing out the influence of higher-order factors.

The second method used the bifactor rotation in the R statistical programming language (R Development Core Team, 2016). In a bifactor or *direct hierarchical* model, all factors (i.e., general, group-specific) are conceptualized as orthogonal first-order dimensions with direct effects on the measured variables (Beaujean, 2015). Although it has been suggested that the SL procedure produces an approximate bifactor structure, Dombrowksi, Canivez, Watkins, and Beaujean (2015) noted that the SL procedure "is simply a re-parameterization of the higher-order model to show how the measured variables relate to the second-order factor and residualized versions of the first-order factors" (p. 195). Additionally, researchers have noted that the SL procedure may suffer from a proportionality constraint which can produce biased parameter estimates. As a consequence, Jennrich and Bentler (2011) developed an alternative to the SL

procedure for EFA to capitalize on the advantages afforded by the bifactor model relative to the indirect hierarchical model that they termed exploratory bifactor analysis (EBFA). Recently, EBFA has been employed to examine the latent structure of several contemporary cognitive measures.

Whereas Dombrowski (2014) found consistent results between EBFA and SL solutions for the Woodcock-Johnson Tests of Cognitive Abilities-Third Edition (WJ-III Cog; Woodcock, McGrew, & Mather, 2001), Dombrowksi and colleagues (2015) reported EBFA structures for the WISC-V that conflicted with those reported by Canivez, Watkins, and Dombrowski (2016b) using the SL procedure. As a consequence, it is believed that an analysis of the WASI-II using both methods will be instructive for furthering our understanding of the higher-order structuring of WASI-II variables as well as the degree to which SL and EBFA solutions converge for the measurement instrument.

Method

Participants. Participants were children and adults ages 6 to 90 ($n = 2,300$) drawn from the WASI-II standardization sample. Demographic characteristics are provided in detail in the WASI-II Manual (Wechsler, 2011b). The standardization sample was obtained using stratified proportional sampling across demographic variables of age, sex, race/ethnicity, parent educational level, and geographic region. Examination of the tables provided in the manual revealed a close correspondence to the 2008 U. S. census estimates across the stratification variables.

Measurement Instrument. The WASI-II is a brief multidimensional test of intelligence for children and adults ages 6 to 90 years. The measure is comprised of four core subtests which contribute to the measurement of two first-order index scores: Verbal Comprehension and Perceptual Reasoning. The subtests also linearly combine to form a second-order full scale FSIQ composite. Subtest scores are expressed as *T* scores with a mean of 50 and a standard deviation of 10. All index and composite variables on the WASI-II are expressed as standard scores with a mean of 100 and a standard deviation of 15. The total norming sample ($N = 2,300$) is nationally representative based upon 2008 U.S. census estimates for gender, race/ethnicity, educational attainment, and geographic region. Extensive normative and psychometric data can be found in the WASI-II Manual (Wechsler, 2011b). Internal consistency estimates ranged from .87 to .92 for the subtests and from .94 to .95 for the index scores across the child and adult samples. The average internal consistency estimate for the FSIQ ranged from .96 to .97. Validity evidence is provided in several forms in the manual and independent reviews are available (e.g., Johnson & Sandoval, 2014; McCrimmon & Smith, 2013).

Procedures and Data Analyses. Following best practice guidelines (e.g., Fabrigar et al., 1999; Preacher & MacCallum, 2003; Thompson, 2004), EFA was implemented using the *psych* package (Revelle, 2016) in R using data extracted from the correlation matrices reported in the WASI-II Manual (Wechsler, 2011b, p. 123). In order to maintain consistency with the structural analyses reported in the manual, three correlation matrices were examined: core subtests for the total sample (ages 6-90, $n = 2,300$), core subtests for the total child sample (ages 6-16, $n = 1,100$), and core subtests for the total adult sample (ages 17-90, $n = 1,200$).

Bartlett's test of sphericity was used to ensure that the correlation matrix was not random, and the Keiser-Meyer-Olkin statistic was required to be above a minimum standard of .60 to ensure that the matrices were suitable for factor analysis (Kaiser, 1974). The PAF method with promax rotation was used because of its ability to disclose latent structure with measurement models that are just identified (de Winter & Dodou, 2012). Next, Horn's parallel analysis (HPA; Horn, 1965), and minimum average partials (MAP; Velicer, 1976) were used to compliment a visual scree test (Cattell, 1966) to determine the number of factors to retain for rotation (Frazier & Youngstrom, 2007; Velicer, Eaton, & Fava, 2000). As per Keith, Cammerer, and Reynolds (2016), parallel analysis based on PAF was used to simulate random eigenvalues for comparison to the obtained values produced from our EFA. The oblique first-order factors were then orthogonalized with the SL procedure, as recommended by Carroll (1995). Additionally, the WASI-II correlation matrices were subjected to EBFA using the bifactor rotation (Jennrich & Bentler, 2011). Criteria for determining factor adequacy were established *a priori*. In accord with Dombrowksi (2013), salient coefficients were defined as those≥ .30 for the oblique solution and ≥ .20 for the orthogonalized solution.

Finally, the bifactor model hypothesizes that each WASI-II subtest is influenced simultaneously by two orthogonal latent constructs: a general intelligence factor (g) and a first-order domain-specific group factor (e.g., VC, PR)[1]. As a consequence, Omega (ω) and omega-hierarchical/hierarchical subscale ($\omega_h \backslash \omega_{hs}$) were estimated as model-based reliability estimates of the latent factors (Gignac & Watkins, 2013). Whereas ω estimates the variance accounted for by both of the constructs in a given domain, ω_h estimates the variance accounted for by a single target construct. Chen, Hayes, Carver, Laurenceau, and Zhang (2012) stressed that "for multidimensional constructs, the alpha coefficient is complexly determined, and McDonald's omega-hierarchical (ω_h; 1999) provides a better estimate for the composite score and thus should be used" (p. 228). Omega estimates were produced using the *Omega* program (Watkins, 2013). Albeit subjective, omega coefficients should at a minimum exceed .50, but .75 would be preferred (Reise, 2012; Reise, Bonifay, & Haviland, 2013).

Results

Factor Extraction Criteria Comparisons. Across the three WASI-II configurations, parallel analysis (Horn, 1965) suggested the retention of two factors whereas the MAP criterion (Velicer, 1976) recommended retention of one factor. Additionally, visual inspection of the scree plots produced by the three WASI-II configurations suggested the presence of two factors. Given that it is better to over factor than under factor (Wood, Tatryn, & Gorsuch, 1996), two factors were extracted to accord with the theoretical structure delineated in the WASI-II Manual (Wechsler, 2011b).

Exploratory (First-Order) Analysis. Results from Bartlett's test of sphericity indicated that the correlation matrix for the total sample was not random, $\chi^2 = 3,759.72$, $df = 6$, $p < .001$. The KMO statistic was .75, well above the minimum standard for conducting factor analysis (Kaiser, 1974). Communality estimates ranged from .54 to .74. On the basis of these values, it was determined that the correlation matrix was appropriate for the EFA procedures that were

---

[1] In a "restricted" Bifactor model, cross loading on multiple group factors is not permitted.

employed. Table 1 presents the results from extracting two WASI-II factors for the total sample. All subtests were saliently and properly associated with their theoretical factor demonstrating desirable simple structure. The correlation between the factors was .71, implying the presence of a higher-order dimension requiring explication (Gignac, 2007; Gorsuch, 1983; Reise, 2012; Thompson, 2004).

Table 1

Exploratory Factor Analysis with Oblique and Orthogonalized Pattern Coefficients of the Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II) Total Normative Sample, Ages 6-90 (N = 2,300)

| | Oblique Solution | | Orthogonolized Schmid-Leiman Solution | | | | |
Subtest	I	II	g	I	II	h^2	u^2
Vocabulary	**.82**	.05	.73	**.45**	.03	.74	.26
Similarities	**.80**	.08	.74	**.43**	.04	.74	.26
Block Design	-.01	**.75**	.62	-.01	**.41**	.55	.45
Matrix Reasoning	.05	**.72**	.65	.03	**.39**	.56	.44
Total Variance (%)			47.2	9.7	8.1	65.0	35.0
Common Variance (%)			72.6	14.9	12.4		

Note. As per Dombrowski (2013), salient loadings ≥ .30 for the oblique solution and ≥ .20 for the orthogonalized solution are denoted in bold. h^2 = communality; u^2= uniqueness.

Total Sample (*N* = 2,300)

       Hierarchical Factor Analyses (SL Orthogonalization and EBFA). Results from the Schmid-Leiman (1957) procedure and Jennrich and Bentler's (2011) EBFA procedure on the two factor solution for the total normative sample are found in Tables 1 and 2. With EBFA, the general factor accounted for 59.3% of the total variance and 74.3% of the common variance. The SL analysis accounted for 47.2% of the total variance and 72.6% of the common variance. The general factor accounted for between 44% to 72% of individual subtest variance in the EBFA. The *g* factor accounted for between 38% to 54% of individual subtest variability in the SL analysis. For EBFA, the two group factors accounted for a small to moderate proportions of the total variance (6.9%-13.7%) and common variance (8.6%-17.1%). The general and group factors combined to measure 79.9% of the variance in the WASI-II, reflecting 20.1% unique variance. For the SL analysis, the two group factors accounted for a small to moderate proportions of the total variance (8.1%-9.7%) and common variance (12.4%-14.9%). The general and group factors combined to measure 65% of the variance in the WASI-II, reflecting 20% unique variance. The results of both analyses demonstrate the robust manifestation of general intelligence, which exceeded the contributions made by both of the group specific factors combined.

_____

Table 2

Sourced of Variance in the WASI-II for the Total Sample According to an Exploratory Bifactor Model

	Factor						
Test	*g*	I	II[a]	*h²*	*u²*	ERROR	*s²*
Vocabulary	.77	.38		.74	.26	.08	.18
Similarities	.78	.36		.74	.26	.10	.16
Block Design	.85		.00	.72	.28	.10	.18
Matrix Reasoning	.67		.74	.99	.01	.11	.00
Total Variance (%)	59.3	6.9	13.7	79.9	20.1	9.7	13.0
Common Variance (%)	74.3	8.6	17.1				
ω	.92	.84	.91				
$\omega_{h/hs}$	.83	.15	.17				

Note. g = general intelligence, b = standardized loading of subtest on factor; S^2 = variance explained; h^2 = communality; u^2 = uniqueness; Error = 1-reliability from Wechsler (2011b); s^2 = u^2-Error; ω = omega; ω_h = Omega hierarchical; ω_{hs} = Omega hierarchical subscale.
[a] Empirically under-identified.

Omega-hierarchical and omega-subscale coefficients were estimated based on the EBFA results in Table 2. The ω_h coefficient for general intelligence (.83) was high and sufficient for scale interpretation; however, the ω_s coefficients for the two specific WASI-II group factors (Verbal and Perceptual) were considerably lower (.15 and .17, respectively). Thus, the two specific WASI-II group factors likely possess too little true score variance for confidant clinical interpretation (Reise, 2012; Reise et al., 2013).

Child Sample (*N* = 1,100)

Exploratory (First-Order) Analysis. Results from Bartlett's test of sphericity indicated that the correlation matrix for the total sample was not random, χ^2 = 1,680.43, *df* = 6, *p* < .001. The KMO statistic was .74, well above the minimum standard for conducting factor analysis (Kaiser, 1974). Communality estimates ranged from .54 to .72. On the basis of these values, it was determined that the correlation matrix was appropriate for the EFA procedures that were employed. Table 3 presents the results from extracting two WASI-II factors for the child sample. All subtests were saliently and properly associated with their theoretical factor demonstrating desirable simple structure. The correlation between the factors was .71, implying the presence of a higher-order dimension requiring explication (Gignac, 2007; Gorsuch, 1983; Reise, 2012; Thompson, 2004).

Table 3

Exploratory Factor Analysis with Oblique and Orthogonalized Pattern Coefficients of the Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II) Child Sample, Ages 6-16 (N = 1,100)

Subtest	Oblique Solution		Orthogonolized Schmid-Leiman Solution				
	I	II	g	I	II	h^2	u^2
Vocabulary	**.82**	.03	.71	**.45**	.02	.71	.29
Similarities	**.81**	.06	.73	**.44**	.03	.73	.27
Block Design	.00	**.74**	.62	.00	**.40**	.55	.45
Matrix Reasoning	.06	**.72**	.66	.03	**.39**	.58	.42
Total Variance (%)			46.5	9.8	7.9	64.2	35.8
Common Variance (%)			72.4	15.3	12.3		

Note. As per Dombrowski (2013), salient loadings ≥ .30 for the oblique solution and ≥ .20 for the orthogonalized solution are denoted in bold. h^2 = communality; u^2 = uniqueness.

Hierarchical Factor Analyses (SL Orthogonalization and EBFA). Results from the Schmid-Leiman (1957) procedure and Jennrich and Bentler's (2011) EBFA procedure on the two factor solution for the total child sample are found in Tables 3 and 4. With EBFA, the general factor accounted for 58.2% of the total variance and 74.3% of the common variance. The SL analysis accounted for 46.5% of the total variance and 72.4% of the common variance. The general factor accounted for between 44% to 72% of individual subtest variance in the EBFA. The *g* factor accounted for between 39% to 53% of individual subtest variability in the SL analysis. For EBFA, the two group factors accounted for small to moderate proportions of the total variance (6.9%-13.3%) and common variance (8.7%-17%).

Table 4

Sourced of Variance in the WASI-II for the Child Sample According to an Exploratory Bifactor Model

Test	Factor			h^2	u^2	ERROR	s^2
	g	I	II[a]				
Vocabulary	.76	.38		.72	.28	.09	.19
Similarities	.77	.36		.72	.28	.11	.17
Block Design	.85		.00	.72	.28	.11	.17
Matrix Reasoning	.66		.73	.97	.03	.13	.00
Total Variance (%)	58.2	6.9	13.3	78.4	21.6	11.0	13.3
Common Variance (%)	74.3	8.7	17.0				
ω	.92	.84	.90				
$\omega_{h/hs}$	.83	.16	.17				

Note. g = general intelligence, b = standardized loading of subtest on factor; S^2 = variance explained; h^2 = communality; u^2 = uniqueness; Error = 1-reliability from Wechsler (2011b); s^2 = u^2-Error; ω = omega; ω_h = Omega hierarchical; ω_{hs} = Omega hierarchical subscale.
[a] Empirically under-identified.

The general and group factors combined to measure 78.4% of the variance in the WASI-II, reflecting 21.6% unique variance. For the SL analysis, the two group factors accounted for small to moderate proportions of the total variance (7.9%-9.8%) and common variance (12.3%-15.3%). The general and group factors combined to measure 64% of the variance in the WASI-II, reflecting 36% unique variance. The results of both analyses demonstrated the robust manifestation of general intelligence, which exceeded the contributions made by both of the group specific factors combined.

Omega-hierarchical and omega-subscale coefficients were estimated based on the EBFA results in Table 2. The ω_h coefficient for general intelligence (.83) was high and sufficient for scale interpretation; however, the ω_s coefficients for the two specific WASI-II group factors (Verbal and Perceptual) were considerably lower (.16 and .17, respectively). Thus, the two specific WASI-II group factors likely possess too little true score variance for confidant clinical interpretation (Reise, 2012; Reise et al., 2013).

Table 5

Exploratory Factor Analysis with Oblique and Orthogonalized Pattern Coefficients of the Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II) Adult Sample, Ages 17-90 (N = 1,200)

	Oblique Solution		Orthogonolized Schmid-Leiman Solution				
Subtest	I	II	g	I	II	h^2	u^2
Vocabulary	**.84**	.04	.75	**.44**	.02	.76	.24
Similarities	**.82**	.06	.75	**.43**	.03	.75	.25
Block Design	-.01	**.76**	.64	-.01	**.40**	.57	.43
Matrix Reasoning	.07	**.74**	.69	.04	**.39**	.63	.37
Total Variance (%)			50.0	9.6	7.8	67.4	32.6
Common Variance (%)			74.2	14.2	11.6		

Note. As per Dombrowski (2013), salient loadings ≥ .30 for the oblique solution and ≥ .20 for the orthogonalized solution are denoted in bold. h^2 = communality; u^2 = uniqueness.

Adult Sample ($N = 1,200$)

Exploratory (First-Order) Analysis. Results from Bartlett's test of sphericity indicated that the correlation matrix for the total sample was not random, $\chi^2 = 2,065.69$, $df = 6$, $p < .001$. The KMO statistic was .75, well above the minimum standard for conducting factor analysis (Kaiser, 1974). Communality estimates ranged from .56 to .75. On the basis of these values, it was determined that the correlation matrix was appropriate for the EFA procedures that were employed. Table 5 presents the results from extracting two WASI-II factors for the child sample.

All subtests were saliently and properly associated with their theoretical factor demonstrating desirable simple structure. The correlation between the factors was .72, implying the presence of a higher-order dimension requiring explication (Gignac, 2007; Gorsuch, 1983; Reise, 2012; Thompson, 2004).

      Hierarchical Factor Analyses (SL Orthogonalization and EBFA). Results from the Schmid-Leiman (1957) procedure and Jennrich and Bentler's (2011) EBFA procedure on the two factor solution for the total adult sample are found in Tables 5 and 6. With EBFA, the general factor accounted for 61.2% of the total variance and 75.7% of the common variance. The SL analysis accounted for 50% of the total variance and 74.2% of the common variance. The general factor accounted for between 48% to 74% of individual subtest variance in the EBFA. The g factor accounted for between 41% to 56% of individual subtest variability in the SL analysis. For EBFA, the two group factors accounted for a small to moderate proportions of the total variance (7%-12.6%) and common variance (8.7%-15.6%). The general and group factors combined to measure 80.8% of the variance in the WASI-II, reflecting 19.2% unique variance. For the SL analysis, the two group factors accounted for small to moderate proportions of the total variance (7.8%-9.6%) and common variance (11.6%-14.2%). The general and group factors combined to measure 67% of the variance in the WASI-II, reflecting 33% unique variance. The results of both analyses demonstrated the robust manifestation of general intelligence, which exceeded the contributions made by both of the group specific factors combined.

Table 6

Sourced of Variance in the WASI-II for the Adult Sample According to an Exploratory Bifactor Model

Test	Factor			h^2	u^2	ERROR	s^2
	g	I	II[a]				
Vocabulary	.78	.39		.76	.24	.08	.16
Similarities	.79	.36		.75	.25	.09	.17
Block Design	.86		.00	.74	.26	.09	.17
Matrix Reasoning	.69		.71	.98	.02	.10	.00
Total Variance (%)	61.2	7.0	12.6	80.8	19.2	9.0	12.5
Common Variance (%)	75.7	8.7	15.6				
ω	.93	.86	.91				
$\omega_{h/hs}$	.84	.16	.16				

Note. g = general intelligence, b = standardized loading of subtest on factor; S^2 = variance explained; h^2 = communality; u^2 = uniqueness; Error = 1-reliability from Wechsler (2011b); s^2 = u^2-Error; ω = omega; ω_h = Omega hierarchical; ω_{hs} = Omega hierarchical subscale.
[a] Empirically under-identified.

      Omega-hierarchical and omega-subscale coefficients were estimated based on the EBFA results in Table 6. The ω_h coefficient for general intelligence (.84) was high and sufficient for scale interpretation; however, the ω_s coefficients for the two specific WASI-II group factors (Verbal and Perceptual) were considerably lower (.16). Thus, the two specific WASI-II group

factors likely possess too little true score variance for confidant clinical interpretation (Reise, 2012; Reise et al., 2013).

Discussion

The recently revised joint standards for educational and psychological testing (American Educational Research Association [AERA], American Psychological Association [APA], & The National Council on Measurement in Education [NCME], 2014) stipulate that the dimensions purported to be measured by a psychological test must be supported with appropriate psychometric evidence (e.g. reliability and validity studies). Despite the converging empirical and theoretical evidence presented in the manual that the WASI-II is a two-factor test, its hypothesized multilevel structure was not examined. Whereas users are encouraged to focus most of their interpretive weight on the second-order FSIQ composite, the EFA and CFA results reported in the WASI-II Manual tend to provide evidence only for the interpretability of the VC and PR scores (Carreta & Ree, 2001). As per Gustaffson and Åberg-Bengtsen (2010), it is impossible to make diagnostic inferences about a more general factor's (i.e., g) influence on observed variables unless such a dimension is estimated in a structural model.

Accordingly, the purpose of the present study was to examine the structural validity of the WASI-II using two exploratory factor analytic procedures that take into account the higher-order structure of the measurement instrument using more rigorous factor extraction criteria and theory-based rotation methods. Although use of the SL procedure (Schmid & Leiman, 1957) is well established within the cognitive assessment literature, EBFA (Jennrich & Bentler, 2011) is a promising technique that has recently emerged in applied assessment research. The present study sought to clarify the structural validity of the WASI-II scores and to apportion subtest variance correctly according to general and group-specific dimensions (Carroll, 1995). It is believed that the results from these analyses will better inform interpretive practice for the measurement instrument.

When examining the variance accounted for by the general factor in the publisher suggested two-factor solution, the EBFA procedure accounted for higher proportions of total variance (47.2% for SL and 59.3% for EBFA) and common variance (72.6% for SL and 74.3% for EBFA) in the WASI-II normative sample. Inspection of the subtest specific g loadings revealed that loadings were relatively consistent across both methods with the exception of Block Design. As an example, in the total normative sample analyses, the g loading for Block Design was significantly higher in the EBFA solution (.62 in SL and .85 in EBFA).

Interestingly, first run PAF analyses produced first-order factor loadings consistent with those reported in the manual from the WISC-IV/WAIS-IV analyses further calling into question the concerns noted in the manual about potential construct under-identification in the WASI-II normative sample. Application of the SL procedure resulted in all WASI-II subtests being saliently and properly associated with their theoretically consistent group-factors with none of the subtests demonstrating cross-loading on multiple factors, resulting in desirable simple structure. Similar results were observed in EBFA with the exception of Block Design, which failed to load saliently on a latent dimension apart from g, resulting in a mathematically impermissible PR factor. It should be noted that these results were also replicated in the child and adult samples.

In sum, the SL procedure provided stronger support for the presence of the VC and PR factors, even after variance attributable to the general factor was residualized from the group-specific factors. It is interesting to note that despite their more parsimonious configurations on

the WASI-II, the variance accounted for by the VC and PR factors in the SL solutions from the present study are consistent with similar estimates produced from SL solutions for the WISC-IV (Watkins, 2006) and WAIS-IV (Canivez & Watkins, 2010), suggesting that the sacrifice in quality of measurement provided by brief measures such as the WASI-II may be overstated (Floyd, Clark, & Shadish, 2008). Nevertheless, the hierarchical *g*–factor accounted for substantially greater total and common variance in WASI-II subtests for all age groups.

Conversely, the EBFA solution consistently produced an impermissible PR factor with only a single loading from Matrix Reasoning, suggesting that the WASI-II may be overfactored (Frazier & Youngstrom, 2007). While this discrepancy may seem unusual at first glance, it is important to reiterate that the bifactor model and the indirect hierarchical model estimated by SL are not functionally equivalent (Beaujean, 2015; Reise, 2012). In the bifactor model, all factors are extracted directly from the measured variables. Typically, *g* is formed first with the group-specific factors being formed from any of the remaining variance not explained by *g*. As a result, "the factors are all competing with each other to explain the subtests' variance" (Dombrowski et al., 2015, p. 199). In the present analyses, it appears that *g* explains most of the common variance in Block Design resulting in minimal covariance left for the PR factor. Similar effects on VC measures were observed by Dombrowksi and colleagues (2015) in their EBFA of the WISC-V. According to Beaujean, Parkin, & Parker (2014), it is not uncommon for cognitive group-specific factors such as PR to be degraded in the presence of a strong general factor. Given the small portions of total and common variance uniquely attributed to the WASI-II VC and PR factors and the low portions of true score variance in these factors (as estimated by ω_s coefficients) there appears to be little variance apart from *g* in these factor scores to warrant clinical interpretation (Reise, 2012; Reise et al., 2013), or if interpreted, done with extreme caution.

Limitations

The present results were derived from covariance matrices representing a large proportion of the WASI-II age span. Although these results provide a relevant comparison to the psychometric analyses presented in the manual, additional independent examinations of the construct validity of the measurement instrument at different points of the age span would be beneficial.

Also, it is not possible to determine which manifestation of the higher-order model (i.e., direct versus indirect) is most appropriate for the data via EFA. As noted by Beujean (2015), "While the higher-order model [indirect hierarchical] is technically nested within the bi-factor model, they provide very different conceptualizations of *g* and other factors in the model" (p. 122). Given its advantages, it has been argued that the bifactor model, in which the effects of the general and specific factors on measured variables are estimated directly, is preferred as the effects of the latent variables are easier to interpret (Canivez, 2016; Gignac, 2007, 2016). Nevertheless, some researchers have questioned whether the bifactor model is a tenable structure for human cognitive abilities (Murray & Johnson, 2013; Reynolds & Keith, 2013). While adjudication of this issue is beyond the scope of the present discussion, the present results indicate that, regardless of one's preference in terms of a higher-order model, the effects of the general factor are superior to the group factors on the WASI-II.

Finally, while it may be argued that the present results do nothing but confirm the long-standing guidance that brief measures such as the WASI-II should be used as screening tests for general intelligence (Gorth-Marnat, 2009; Sattler, 2008), test publishers have increased the

factorial complexity of recent instruments providing an array of scores for practitioners to potentially interpret. Given the lack of target construct specificity in the VC and PR index scores, it is our position that their provision on the WASI-II is an invitation for misuse. Within the professional literature the distinction between comprehensive and brief intelligence tests is increasingly becoming obscured (see Reynolds, Kamphaus, & Raines, 2012 for an example). In fact, it is now common practice for factors to be identified by only two indicators on comprehensive ability measures (Frazier & Youngstrom, 2007). Unfortunately, examiners all too often operate under the assumption that if a score is provided by a test publisher that it accurately reflects a latent psychological trait (Kane, 2013). As a result, it may be tempting for assessment professionals to consider adopting brief measures for purposes that they were not intended for (e.g., specific learning disability identification) due to their perceived advantages in terms of cost and administration time when compared to more comprehensive measures (Glutting, Watkins, & Youngstrom, 2003). However, when selecting and/or adopting tests for specific purposes we encourage practitioners to "(a) know what their tests can do and (b) act accordingly" (Weiner, 1989, p. 829).

Conclusion

The present study provides clinicians with important information for interpreting the WASI-II (Wechsler, 2011a). As "the ultimate responsibility for appropriate test use and interpretation lies predominantly with the test user" (AERA, APA, & NCME, 2014, p. 141), clinicians using the WASI-II in clinical evaluations must seriously consider the present information to make informed decisions about which scores have satisfactory reliability, validity, and utility. Whereas our results suggest that users can be reasonably confident in their interpretations of the FSIQ composite, the contribution of the group factors was less consistent. As a consequence, it is recommended that users of the WASI-II, should focus most, if not all, of their interpretive weight at that level of measurement.

About the Authors

Brittany McGeehan, B.A. is a doctoral student in the school psychology program at Texas Woman's University.

Nadine Ndip, B.A. is a doctoral student in the school psychology program at Texas Woman's University.

Ryan J. McGill, Ph.D., BCBA-D, NCSP is assistant professor and director of the school psychology program at the College of William and Mary.

Correspondence concerning this article should be addressed to Brittany McGeehan, Department of Psychology and Philosophy, Texas Woman's University, P. O. Box 425470 Denton, TX. 76204. E-Mail: blerma@twu.edu

References

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

Axelrod, B. N. (2002). Validity of the Wechsler Abbreviated Scale of Intelligence and other very short forms of estimating intellectual functioning. *Assessment, 9,* 17-23. doi: 10.1177/1073191102009001003

Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher-order models. *Journal of Intelligence, 3,* 121-136. doi: 10.3390/jintelligence3040121

Beaujean, A. A., McGlaughlin, & 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

Beaujean, A. A., Parkin, J., & Parker, S. (2014). Comparing Cattell-Horn-Carroll factor models: Differences between bifactor and higher order factor models in predicting language achievement. *Psychological Assessment, 26,* 789-805. doi: 10.1037/a0036745

Beal, A. L., Willis, J. O., & Dumont, R. (2013). Psychological testing by models of cognitive ability. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.), *The oxford handbook of child psychological assessment* (pp. 445-473). New York: Oxford University Press.

Brown, T. A. (2016). *Confirmatory factor analysis for applied research* (2nd ed.). New York: Guilford Press.

Canivez, G. L. (2016). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements* (pp. 247-271). Gottingen, Germany: Hogrefe.

Canivez, G. L. (2013). Psychometric versus actuarial interpretation of intelligence and related aptitude batteries. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.). *The Oxford handbook of child psychological assessment* (pp. 84-112). New York: Oxford University Press.

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. *School Psychology Quarterly, 24,* 252-265. doi: 10.1037/a0018030

Canivez, G. L., & McGill, R. J. (2016). Factor structure of the Differential Ability Scales-Second Edition: Exploratory and hierarchical factor analyses with the core subtests. *Psychological Assessment*. Advance online publication. doi: 10.1037/pas0000279

Canivez, G. L., & Watkins, M. W. (2010). Investigation of the factor structure of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV): Exploratory and higher-order factor analyses. *Psychological Assessment, 22,* 827-836. doi: 10.1037/a0020429

Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2016a). Structural validity of the Wechsler Intelligence Scale for Children-Fifth Edition: Confirmatory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment*. Advance online publication. doi: 10.1037/pas0000358

Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2016b). Factor structure of the Wechsler Intelligence Scale for Children-Fifth Edition: Exploratory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment, 28,* 975-986. doi: 10.1037/pas0000238

Carretta, T. R., & Ree, J. J. (2001). Pitfalls of ability research. *International Journal of Selection and Assessment, 9,* 325-335. doi: 10.1111/1468-2389.00184

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

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

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.

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

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J. P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality, 80,* 219-251. doi:10.1111/j.1467–6494.2011.00739.x

de Winter, J. C. F., & Dodou, D. (2012). Factor recovery by principle axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics, 39,* 695-710. Doi: 10.1080/02664763.2011.610445

Dombrowski, S. C. (2013). Investigating the structure of the WJ-III Cognitive at school age. *School Psychology Quarterly, 28,* 154-169. doi: 10.1037/spq0000010

Dombrowski, S. C. (2014). Investigating the structure of the WJ-III Cognitive in early school age through two exploratory bifactor analysis procedures. *Journal of Psychoeducational Assessment, 32,* 483-494. doi: 10.1177/0734282914530838

Dombrowski, S. C., Canivez, G. L., Watkins, M. W., & Beaujean, A. (2015). Exploratory bifactor analysis of the Wechsler Intelligence Scale for Children-Fifth Edition with the 16 primary and secondary subtests. *Intelligence, 53,* 194-201. doi: 10.1016/j.intell.2015.10.009

Dombrowski, S. C., McGill, R. J., & Canivez, G. L. (2016). Exploratory and hierarchical factor analysis of the WJ IV Cognitive at school age. *Psychological Assessment*. Advance online publication. doi: 10.1037/pas0000350

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

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

Floyd, R. G., Clark, M. H., & Shadish, W. R. (2008). The exchangeability of IQs: Implications for professional psychology. *Professional Psychology: Research and Practice, 39,* 414-423. doi: 10.1037/0735-7028.39.4.414

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

Gignac, G. E. (2007). Multi-factor modeling in individual differences research: Some recommendations and suggestions. *Personality and Individual Differences, 42,* 37-48. doi: 0.1016/j.paid.2006.06.019

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bi-factor model tends to fit better. *Intelligence, 55,* 57-68. doi: 10.1016/j.intell.2016.01.006

Gignac, G. E., & Watkins, M. W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS–IV. *Multivariate Behavioral Research, 48,* 639-662. doi:10.1080/00273171.2013.804398

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

Glutting, J. J., Watkins, M. W., & Youngstrom, E. A. (2003). Multifactored and cross-battery assessments: Are they worth the effort? In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence*

aptitude, and achievement (2nd ed., pp. 343-374). New York: Guilford.

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

Groth-Marat, G. (2009). *Handbook of psychological assessment* (5th ed.). Hoboken, NJ: John Wiley.

Gustaffson, J. E., & Åberg-Bengtsen, L. (2010). Unidimensionality and interpretability of psychological instruments. In S. E. Embretson (Ed.), *Measuring psychological constructs: Advances in model-based approaches* (pp. 97-122). Washington, DC: American Psychological Association.

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

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika, 76,* 537-549. doi: 10.1007/s11336-011-9218-4

Johnson, K. M., & Sandoval, J. (2014). Wechsler Abbreviated Scale of Intelligence-Second Edition. In J. F. Carlson, K.F. Geisinger, & J. L. Jonson (Eds.), *The nineteenth mental measurements yearbook* (pp. 1-7). Lincoln, NE: Buros Institute of Mental Measurement.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39,* 31-36. doi: 10.1007/BF02291575

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50,* 1-73. doi: 10.1111/jedm.12000

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). 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. doi: 10.1177/073428290101900305

Keith, T. Z., Caemmerer, J. M., & Reynolds, M. R. (2016). Comparison of methods for factor extraction for cognitive test-like data: Which overfactor, which underfactor? *Intelligence, 54,* 37-54. doi: 10.1016/j.intell.2015.11.003

Kranzler, J. H., & Keith, T. Z. (1999). Independent confirmatory factor analysis of the Cognitive Assessment System (CAS): What does the CAS measure? *School Psychology Review, 28,* 117-144. Retrieved from http://www.nasponline.org

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

McCrimmon, A. W., & Smith, A. D. (2013). Review of the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II). *Journal of Psychoeducational Assessment, 31,* 337-341. doi: 10.1177/0734282912467756

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

McGill, R. J. (2016). Investigating the factor structure of the Comprehensive Test of Nonverbal Intelligence-Second Edition (CTONI-2) using exploratory factor analysis. *Journal of Psychoeducational Assessment, 34,* 339-350. doi: 10.1177/0734282915610717

McGill, R. J., & Canivez, G. L. (2016). Orthogonal higher order structure of the WISC–IV Spanish using hierarchical exploratory factor analytic procedures. *Journal of Psychoeducational Assessment, 34,* 600-606. doi: 10.1177/0734282915624293

McGill, R. J., & Spurgin, A. R. (2015). Exploratory higher order analysis of the Luria interpretive model on the Kaufman Assessment Battery for Children-Second Edition (KABC-II) school-age battery. *Assessment*. Advance online publication. doi: 10.1177/1073191115614081

Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence, 37,* 487-494. doi: 10.1016/j.intell.2013.06.004

Nelson, J. M., & Canivez, G. L. (2012). Examination of the structural, convergent, and incremental validity of the Reynolds Intellectual Assessment Scales (RIAS) with a clinical sample. *Psychological Assessment, 24,* 129-140. doi: 10.1037/a0024878

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. doi: 10.1016/j.jsp.2007.03.003

Pierson, E. E., Kilmer, L. M., Rothlisberg, B. A., & McIntosh, D. E. (2012). Use of brief intelligence tests in the identification of giftedness. *Journal of Psychoeducational Assessment, 30,* 10-24. doi: 177/0734282911428193

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics, 2,* 13-43. doi: 10.1207/S15328031US0201_02

R Developmental Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47,* 667-696. doi:10.1080/00273171.2012.715555

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment, 95,*

129-140. doi:10.1080/00223891.2012.725437

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

Reynolds, C. R., & Kamphaus, R. W. (2015). *Reynolds Intellectual Assessment Scales* (2nd ed.). Lutz, FL: Psychological Assessment Resources.

Reynolds, C. R., Kamphaus, R. W., & Raines, T. C. (2012). 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* (3rd ed.; pp. 400-421). New York: Guilford.

Reynolds, M. R., & T. Z. Keith (2013). Measurement and statistical issues in child assessment research. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.). *The Oxford handbook of child psychological assessment* (pp. 48-83). New York: Oxford University Press.

Revelle, W. (2016). *Psych: Procedures for psychological, psychometric, and personality research* (version 1.6.9) [Computer Software]. Evanston, IL: Northwestern University.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98,* 223-237. doi: 10.1080/00223891.2015.1089249

Ryan, J. J., Carruthers, C. A., Miller, L. J., Souheaver, G. T., Gontkovsky, S. T., & Zehr, M. D. (2010). Exploratory factor analysis of the Wechsler Abbreviated Scale of Intelligence (WASI) in adult standardization and clinical samples. *Applied Neuropsychology: Adult, 10,* 252-256. doi: 10.1207/s15324826an1004_8

Sattler, J. M. (2008). *Assessment of children: Cognitive foundations* (5th ed.). La Mesa, CA: Sattler Publishing.

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

Silverstein, A. B. (1990). Short forms of individual intelligence tests. *Psychological Assessment, 2,* 3-11. doi: 10.1037/1040-3590.2.1.3

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. doi: 10.1177/0013164496056002001

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 31,* 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. (2006). Orthogonal higher order structure of the Wechsler Intelligence Scale for Children-Fourth Edition. *Psychological Assessment, 18,* 123-125. doi: 10.1037/1040-3590.18.1.123

Watkins, M. W. (2013). *Omega*. [Computer Software]. Phoenix, AZ: Ed & Psych Associates.

Wechsler, D. (2011a). *Wechsler Abbreviated Scale of Intelligence* (2nd ed.). Bloomington, MN: Pearson.

Wechsler, D. (2011b). *Wechsler Abbreviated Scale of Intelligence—Second Edition manual*. Bloomington, MN: Pearson.

Weiner, I. B. (1989). On competence and ethicality in psychodiagnostic assessment. *Journal of Personality Assessment, 53,* 827-831. doi:10.1207/s15327752jpa5304_18

Wood, J. M., Tatryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction 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. (2011). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside.

Yates, B. T., & Taub, J. (2003). Assessing the costs, benefits, cost-effectiveness, and cost-benefit of psychological assessment: We should, we can, and here's how. *Psychological Assessment, 15,* 478-495. doi: 10.1037/1040-3590.15.4.478

Youngstrom, E. A. (2008). Commentary: Evidence-based assessment is not evidence-based medicine—Commentary on evidence-based assessment of cognitive functioning in pediatric psychology. *Journal of Pediatric Psychology, 33,* 1015-1020. doi: 10.1093/jpepsy/jsn060