Intelligence 53 (2015) 194–201

Contents lists available at ScienceDirect

Intelligence

Exploratory bifactor analysis of the Wechsler Intelligence Scale for Children—Fifth Edition with the 16 primary and secondary subtests

Stefan C. Dombrowski [a,*], Gary L. Canivez [b], Marley W. Watkins [c], A. Alexander Beaujean [c]

[a] Department of Graduate Education, Leadership and Counseling, Rider University, United States
[b] Department of Psychology, Eastern Illinois University, United States
[c] Department of Educational Psychology, Baylor University, United States

ARTICLE INFO

Article history:
Received 27 July 2015
Received in revised form 26 October 2015
Accepted 26 October 2015
Available online 08 November 2015

Keywords:
WISC-V
Bifactor rotation
Exploratory factor analysis
Intelligence
Structural validity
Model-based reliability

ABSTRACT

Despite substantial revisions involved in creating the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V; Wechsler, 2014a), the test publisher relied exclusively upon confirmatory factor analytic procedures to determine the instrument's structure and failed to apportion the variance among factors and subtests. To fill this lacuna, the factor structure of the 16 primary and secondary subtests of the WISC-V standardization sample was examined with exploratory bifactor analysis (EBFA). EBFA results provided strong support for a general intelligence (*g*) factor, but nominal evidence for three group factors (i.e., Processing Speed, Working Memory, and Perceptual Reasoning). There was no evidence for distinct verbal, fluid reasoning or visual–spatial factors. The *g* factor accounted for large portions of total and common subtest variance while the group factors accounted for negligible portions of total and common variance. These results suggest that clinical interpretation of the WISC-V should reside primarily at the global level (i.e., Full Scale IQ).

© 2015 Published by Elsevier Inc.

1. Introduction

The developers of the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V; Wechsler, 2014a) stated that they not only used current cognitive, intellectual, and neuropsychological theories (Carroll, 1993, 2003; Cattell & Horn, 1978; Horn, 1991; Horn & Blankson, 2012; Horn & Cattell, 1966; McCloskey, Whitaker, Murphy, & Rogers, 2012; Miller & Maricle, 2012) to guide its creation, but also retained its longstanding linkage to Spearman's (1904) notion of general intelligence. Evidence of structural validity was established via confirmatory factor analyses (CFA) and reported in the WISC-V *Technical and Interpretive Manual* (Wechsler, 2014b), which included the specification of higher-order factor models with a single second-order general intelligence (*g*) factor *indirectly* influencing subtests through five first-order factors. However, scholars have raised a number of concerns regarding that structure (Canivez & Watkins, in press; Canivez, Watkins, & Dombrowski, 2015). Canivez and Watkins (in press) and Canivez, Watkins, James, James, and Good (2014) noted that there was insufficient detail in describing how the factors were defined and why weighted least squares estimation was used. For example, WLS estimation is typically used with categorical or non-normal data, requires much larger sample sizes, and can lead to model misspecification more readily than maximum likelihood estimation (Hu, Bentler, & Kano, 1992; Olsson, Foss, Troye, & Howell, 2000; Yuan & Chan, 2005). Canivez and colleagues also indicated that the preferred CFA model abandoned the parsimony of simple structure by allowing cross-loadings of the Arithmetic subtest. Further, there was a standardized path coefficient of 1.00 between the higher-order general intelligence factor and the first-order Fluid Reasoning (FR) factor, suggesting that *g* and FR were empirically redundant (Le, Schmidt, Harter, & Lauver, 2010). Canivez et al. also expressed concern about the use of chi-square difference tests of nested models to identify the five-factor model because this approach has been shown to be misleading when the base model is misspecified (Yuan & Bentler, 2004) and is overly powerful with large sample sizes (Millsap, 2007).

There are five additional issues with the test publisher's approach to documenting the WISC-V structure. First, the test publisher did not examine rival models, such as a bifactor model. Bifactor models are sometimes preferred over higher-order models (Canivez, in press; Reise, 2012) and have been recommended for tests of cognitive ability because they allow for partitioning of general and group factor variance (Beaujean, Parkin, & Parker, 2014; Canivez, 2014b; Canivez et al., 2015, 2014; Carroll, 1997; Gignac, 2005, 2006; Gignac & Watkins, 2013; Nelson, Canivez, & Watkins, 2013; Watkins, 2010; Watkins & Beaujean, 2014; Watkins, Canivez, James, James, & Good, 2013; Brunner, Nagy, & Wilhelm, 2012) and are more in line with Carroll's three-stratum theory cognitive ability (Beaujean, 2015). This inclusion would aid clinicians and researchers in determining the interpretability

* Corresponding author at: School Psychology Program, 2083 Lawrenceville Road, Lawrenceville, NJ 08648, United States.
E-mail address: sdombrowski@rider.edu (S.C. Dombrowski).

http://dx.doi.org/10.1016/j.intell.2015.10.009
0160-2896/© 2015 Published by Elsevier Inc.

of group factors (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; Gustafsson & Aberg-Bengtsson, 2010).

Second, model-based reliability estimates including omega-hierarchical (ω_h) and omega-subscale (ω_s) (Gignac & Watkins, 2013; Reise, 2012; Reise, Bonifay, & Haviland, 2013; Shrout & Lane, 2012; Zinbarg, Revelle, Yovel, & Li, 2005, 2009) were not included in the *Technical and Interpretive Manual*. Several researchers (e.g. Canivez, 2010; Canivez, 2014a; Canivez & Kush, 2013) as well as the *Standards for educational and psychological testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014), have emphasized the need for these statistics in IQ test manuals that recommend the interpretation of subscores. Along with the measurement of total and common variance for general- and group/specific, ω estimates can aid in determining how much interpretive emphasis should be placed upon scores designed to measure primary and secondary factors.

Third, the WISC-V authors did not furnish EFA results; instead they relied exclusively upon CFA procedures when providing structural validity evidence. Gorsuch (1983) and others (e.g. Brown, 2015; Carroll, 1993; Reise, 2012) indicated that EFA and CFA are complementary, and test users can have greater confidence in an instrument's structure when both procedures are in agreement, particularly when an instrument has been revised and reformulated. For instance, elimination of the Word Reasoning and Picture Completion subtests and addition of Visual Puzzles, Figure Weights, and Picture Span subtests could have caused unexpected changes to the WISC-V factor structure that would benefit from EFA prior to the use of CFA (Strauss, Spreen, & Hunter, 2000).

Fourth, previous independent investigations of intelligence test factor structures using EFA methods have produced divergent results from those offered by CFA-based models of extant IQ subtests (e.g. Canivez, 2008; Canivez & Watkins, 2010a, 2010b; DiStefano & Dombrowski, 2006; Dombrowski, 2013; Dombrowski, 2014a, 2014b; Dombrowski & Watkins, 2013; Dombrowski, Watkins, & Brogan, 2009; Watkins, 2006). In fact, some researchers contend that present day IQ tests are overfactored (Frazier & Youngstrom, 2007).

Finally, Canivez et al. (2015, 2014) recently subjected the WISC-V total sample correlation matrix to EFA using the Schmid–Leiman (SL) procedure. The SL procedure mathematically transforms a second-order factor solution into an orthogonal first-order structure where general and group factors both directly influence indicator variables. Schmid and Leiman (1957) argued that this process "preserves the desired characteristics of the oblique solution" and "discloses the hierarchical structure of the variables" (p. 53). Carroll (1995) also emphasized that orthogonal factors are appropriate only when produced in the context of a Schmid–Leiman solution. Canivez et al.'s SL analysis resulted in a four-factor solution where the fluid reasoning and visual spatial subtests combined to form the WISC-IV's previously identified perceptual reasoning factor. Additionally, their results revealed the preeminence of the higher-order *g* factor and prompted them to recommend that primary interpretative emphasis should be placed on the FSIQ with possible secondary interpretive emphasis on the processing speed index score.

Although useful, 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. As with higher-order models in general, loading values from the SL transformation may be biased if there are cross-loadings (Reise, 2012). Likewise, the loadings of all measured variables on a group factor are constrained to be proportional (Schmiedek & Li, 2004). Given these issues, Jennrich and Bentler (2011) developed an alternative to the SL procedure for EFA: exploratory bifactor analysis (EBFA). They described EBFA as "simply exploratory factor analysis using a bi-factor rotation criterion" (p. 2). EBFA is designed to estimate loadings from bifactor models directly, which Jennrich and Bentler contend can be better than the SL transformation in some cases. The only independently published article comparing the two procedures on cognitive ability data, however, found consistent results between EBFA and the SL (Dombrowski, 2014b).

2. Method

2.1. Participants

Participants included the entire WISC-V standardization sample ($N = 2200$), ranging in age from 6 to 16 years. Detailed demographic characteristics are available in the WISC-V *Technical and Interpretive Manual* (Wechsler, 2014b). The standardization sample was obtained using stratified proportional sampling across variables of age, sex, race/ethnicity, parental education level, and geographic region. Education level was used as a proxy for socioeconomic status. Examination of tables in the *Technical and Interpretive Manual* (Wechsler, 2014b) revealed a close match to the U.S. census across stratification variables.

2.2. Instrument

The WISC-V is an individually administered test of cognitive ability for children aged 6–16 years. The Full Scale IQ (FSIQ) is composed of seven primary subtests across five domains [Verbal Comprehension (VC), Visual Spatial (VS), Fluid Reasoning (FR), Working Memory (WM), and Processing Speed (PS)]. The Primary Index Scale level is composed of 10 WISC-V subtests (primary subtests) that are used to estimate the five WISC-V factor index scores (VCI, VSI, FRI, WMI, PSI). Fig. 1 illustrates the publisher's proposed latent factor structure.

2.3. Procedure and analyses

The WISC-V subtest correlation matrix for the total standardization sample was obtained from the *Technical and Interpretive Manual* (Table 5.1; Wechsler, 2014b). In addition to purposefully extracting 2 to 5 factors to map onto the CFA models posited in the WISC-V *Technical and Interpretive Manual*, multiple empirical factor extraction criteria were also examined (Gorsuch, 1983) as well as factor interpretability and compliance with simple structure (Thurstone, 1947). Specifically, eigenvalues > 1 (Kaiser, 1960), the scree test (Cattell, 1966), standard error of scree (SE_{Scree}; Zoski & Jurs, 1996), Horn's parallel analysis (HPA; Horn, 1965), and minimum average partials (MAP; Velicer, 1976) were examined. After determining the number of factors to extract, factors were then rotated using the bifactor rotation with orthogonal group factors (Jennrich & Bentler, 2011). All EBFA analyses were conducted using the R statistical programming language (R Development Core Team, 2015) using the BaylorEdPsych and Psych packages (cf. Beaujean, 2013, 2014; Revelle, 2012). Omega estimates (ω_h and ω_s) were produced using the *Omega* program developed by Watkins (2013).

3. Results

3.1. Factor extraction criteria comparisons

MAP suggested one factor; eigenvalue > 1, scree, and HPA suggested 2–3 factors; whereas the test publisher recommended five factors. Given that it is better to overextract than underextract (Wood, Tataryn, & Gorsuch, 1996), and to attempt a replication of the publisher's five-factor model, we began by extracting five factors. Models with four, three, and two factors were also successively examined for adequacy.

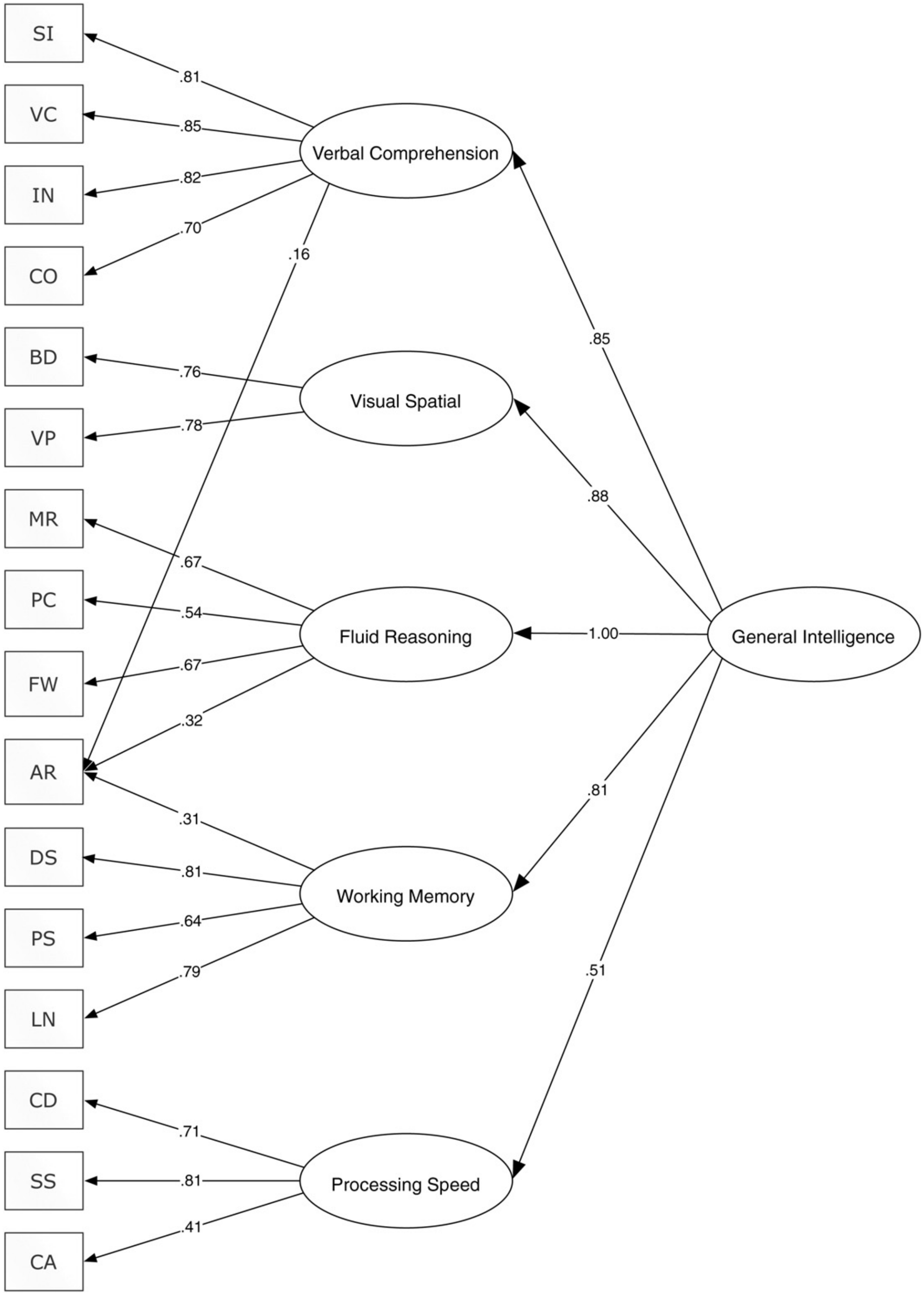

Fig 1. Higher-order measurement model with standardized coefficients (adapted from Figure 5.1 [Wechsler, 2014b]), for WISC-V standardization sample (N = 2200) 16 subtests. SI = Similarities, VC = Vocabulary, IN = Information, CO = Comprehension, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, PC = Picture Concepts, FW = Figure Weights, AR = Arithmetic, DS = Digit Span, PS = Picture Span, LN = Letter–Number Sequencing, CD = Coding, SS = Symbol Search, CA = Cancellation. Wechsler Intelligence Scale for Children, Fifth Edition (WAIS-V). Copyright 2014 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "Wechsler Intelligence Scale for Children" and "WAIS" are trademarks, in the United States and/or other countries, of Pearson Education, Inc. or its affiliates(s).

3.2. Exploratory bifactor analyses

Table 1 presents summary results of extracting two through five factors and then rotating the factors using a bifactor rotation (Jennrich & Bentler, 2011). Across all factor extractions, the general factor consistently accounted for over 38% of the subtests' total variance (38.4% to 38.8%) and anywhere from 66.1% to 77.2% of the subtests' common variance. The general factor accounted for 4.8% to 67.2% (*Mdn* range: 38.4% to 39.7%) of the subtest variance. Across all models, the group factors accounted for a small proportion of the subtests' total variance (1.5% to 6.4%) and common variance (2.6% to 12.1%). The general and group factors across two- to five-factor extractions combined to measure 49.7% to 58.2% of the subtests' variance in the WISC-V, indicating that between 41.8% and 50.3% of the subtests' variance was unexplained (i.e., unique variance).

Table 1
Summary results of exploratory bifactor analysis models including two to five group factors.

	5 Factor	4 Factor	3 Factor	2 Factor
g				
Total S^2 (%)	38.6	38.5	38.8	38.4
Common S^2 (%)	68.6	66.1	73.6	77.2
Subtest S^2 (%)	7.3 to 64.0 Mdn = 38.4	7.8 to 64.0 Mdn = 39.1	4.8 to 67.2 Mdn = 39.7	6.3 to 62.4 Mdn = 38.4
Group				
Total S^2 (%)	1.5 to 5.6	2.8 to 5.9	3.3 to 6.4	5.0 & 6.3
Common S^2 (%)	2.6 to 10.0	4.8 to 10.1	6.3 to 12.1	12.8 & 10.1
h^2	56.3	58.2	52.7	49.7
u^2	47.7	41.8	47.3	50.3
g loadings	.27 to .80	.28 to .80	.22 to .82	.25 to .79
Subtests with poor g loadings	PC & CO	CO & CN	CO, SS & CN	CO & CN
ω_h	.85	.849	.802	.828
ω_s	.166 to .247 (only 3 factors)	.166 to .244 (only 3 factors)	.159 to .281	.25 & .185

Note: All factors were rotated using the analytic bifactor rotation (Jennrich & Bentler, 2011). S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_h = omega hierarchical, ω_s = omega subscale. PC = Picture Concepts; CO = Coding; CN = Cancellation and SS = Symbol Search.

The WISC-V subtests' g-loadings ranged from .22 to .82 across all models and most were within the fair to good range based on Kaufman's (1994) criteria (≥.70 = good, .50–.69 = fair, <.50 = poor). The exceptions included Coding across all models, Picture Concepts with five-factor extraction, Cancellation with two- through four-factors, and Symbol Search with three-factors.

The latent factor reliabilities of the WISC-V were estimated with ω_h and ω_s. Reise (2012) and Reise, Bonifay, and Haviland (2013) tentatively suggested that ω values should be greater than .50, and preferred values greater than .75, but these cutoff values have not yet been thoroughly investigated. Low ω_s values suggest that little interpretive weight should be placed on scores representing these factors (e.g., WISC-V index scores) since little true score variance exists at the group level that is independent of the general factor. The ω_h coefficient for the general factor was high across all factor models (.802 to .850) and sufficient for interpretation. Omega subscale (ω_s) coefficients were only produced for three of the five group factors due to a single subtest loading on two of the group factors in the five and four factor models. Omega subscale coefficients ranged from .159 to .281, which is much lower than the minimum suggested for interpretation.

When extracting five WISC-V factors (Table 2), the results produced three factors that cohered to varying degrees with the structure presented in the *Technical and Interpretive Manual* (Wechsler, 2014b). Block Design and Visual Puzzles had loadings of approximately .40 on a Perceptual Reasoning factor. However, Matrix Reasoning and Figure Weights *did not* form a distinct FR factor. Instead, they loaded on the PR factor, but both of their loadings were relatively low (<.30). Digit Span, Picture Span, and Letter–Number Sequencing loaded a Working Memory (WM) factor, although their loadings were not very strong as they only ranged from .35 to .52. The Arithmetic subtest had its highest group loading on the WM factor, but it also was relatively low (<.30). Coding and Symbol Search had relatively strong loadings on what appears to be a PS factor. However, Cancellation had a low loading on this factor and Vocabulary and Information loaded negatively, though weakly, on PS. The fourth and fifth factors from this model were not interpretable as each had only one subtest with a salient loading (i.e., Cancellation on Factor 4 and Picture Concepts on Factor 5; see Table 2). This is an indication of over factoring the data. Even then, the Similarities and Comprehension subtests did not saliently load any group factor.

When extracting four WISC-V factors (Table 3), the results produced three factors that cohered to varying degrees with the theoretical structure presented in the *Technical and Interpretive Manual*. Coding and Symbol Search saliently loaded on what appears to be a Processing Speed (PS) factor. As with the five-factor model, Cancellation had a low loading while Vocabulary and Information had weak negative loadings. Block Design and Visual Puzzles loaded on a common Perceptual Reasoning (PR) factor. Matrix Reasoning and Figure Weights *did not* form a distinct FR factor; instead, they loaded – albeit relatively weakly – on the PR factor. Digit Span, Picture Span, and Letter–Number Sequencing loaded a Working Memory (WM) factor, although their loadings were not very strong. The Arithmetic subtest had its highest group loading on this WM factor, but it was relatively low (<.30). The fourth factor was not interpretable as it only had one subtest (Cancellation) with a salient loading. Similarities, Comprehension, and Picture Concepts did not align with any group factor.

Table 2
Sources of variance in the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V) for the total standardization sample (N = 2200) using exploratory bifactor analysis with five first-order factors.

	General		F1: Processing Speed		F2: Working Memory		F3: Perceptual Reasoning		F4: ?		F5: ?			
WISC-V subtest	b	S^2	b	S^2	b	S^2	b	S^2	b	S^2	b	S^2	h^2	u^2
Similarities	.78	.608	−.170	.029	−.020	.000	−.090	.008	−.110	.012	−.05	.003	.660	.340
Vocabulary	.80	.640	*−.240*	*.058*	−.080	.006	−.090	.008	−.120	.014	−.02	.000	.727	.273
Information	.78	.608	*−.210*	*.044*	−.060	.004	−.070	.005	−.070	.005	−.04	.002	.668	.333
Comprehension	.69	.476	−.120	.014	−.010	.000	−.140	.020	−.040	.002	−.02	.000	.512	.488
Block Design	.64	.410	.040	.002	−.010	.000	**.410**	**.168**	.010	.000	−.02	.000	.580	.420
Visual Puzzles	.64	.410	−.090	.008	.000	.000	**.420**	**.176**	−.030	.001	.07	.005	.600	.400
Matrix Reasoning	.60	.360	−.030	.001	.130	.017	*.220*	*.048*	−.030	.001	.05	.003	.430	.570
Figure Weights	.60	.360	−.120	.014	.100	.010	*.260*	*.068*	−.030	.001	−.01	.000	.453	.547
Picture Concepts	.52	.270	−.050	.003	.020	.000	.040	.002	−.020	.000	**.47**	**.221**	.496	.504
Arithmetic	.69	.476	−.020	.000	*.220*	*.048*	.060	.004	−.030	.001	−.01	.000	.530	.471
Digit Span	.62	.384	.010	.000	**.520**	**.270**	.040	.002	−.050	.003	.01	.000	.659	.341
Picture Span	.52	.270	.040	.002	**.350**	**.123**	.050	.003	−.060	.004	.05	.003	.403	.597
Letter–Number Sequencing	.62	.384	−.010	.000	**.490**	**.240**	−.050	.003	−.050	.003	−.01	.000	.630	.370
Coding	.45	.203	**.630**	**.397**	.020	.000	−.040	.002	.080	.006	−.02	.000	.608	.392
Symbol Search	.50	.250	**.540**	**.292**	−.020	.000	.010	.000	.130	.017	.00	.000	.559	.441
Cancellation	.27	.073	*.200*	*.040*	−.060	.004	−.010	.000	**.620**	**.384**	−.01	.000	.501	.499
Total variance		.386		.056		.045		.032		.028		.015	.563	.437
Common variance		.686		.100		.080		.057		.050		.026		
	ω_h = .850		ω_s = .166		ω_s = .242		ω_s = .174							

Note. b = loading of subtest on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_h = omega hierarchical, ω_s = omega subscale. Bold type indicates salient loading ($b \geq .30$). Italic type indicates coefficient and variance estimate alignment ($.20 \leq b < .30$) with theoretically posited factor.

Table 3
Sources of variance in the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V) for the total standardization sample ($N = 2200$) using exploratory bifactor analysis with four first-order factors.

	General		F1: Processing Speed		F2: Working Memory		F3: Perceptual Reasoning		F4: ?			
WISC-V subtest	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	h^2	u^2
Similarities	0.77	.593	−.180	.032	−.020	.000	−.070	.005	−.110	.012	.800	.200
Vocabulary	0.80	.640	*−.240*	*.058*	−.080	.006	−.070	.005	−.130	.017	.740	.260
Information	0.78	.608	*−.220*	*.048*	−.060	.004	−.050	.003	−.070	.005	.660	.340
Comprehension	0.69	.476	−.130	.017	−.010	.000	−.120	.014	−.050	.003	.510	.490
Block Design	0.63	.397	.030	.001	.000	.000	**.410**	**.168**	.020	.000	.710	.290
Visual Puzzles	0.63	.397	−.100	.010	.010	.000	**.450**	**.203**	−.030	.001	.600	.400
Matrix Reasoning	0.59	.348	−.030	.010	.140	.020	*.240*	*.058*	−.020	.000	.440	.560
Figure Weights	0.60	.360	−.120	.014	.100	.010	*.270*	*.073*	−.040	.002	.510	.490
Picture Concepts	0.51	.260	−.060	.004	.050	.003	.100	.010	−.020	.000	.390	.610
Arithmetic	0.69	.476	−.030	.001	*.220*	*.048*	.070	.005	−.030	.001	.660	.340
Digit Span	0.62	.384	.010	.000	**.530**	**.281**	.040	.002	−.040	.002	.700	.300
Picture Span	0.52	.270	.030	.001	**.350**	**.123**	.070	.005	−.050	.003	.480	.520
Letter–Number Sequencing	0.63	.397	−.010	.000	**.490**	**.240**	−.040	.002	−.050	.003	.680	.320
Coding	0.47	.221	**.690**	**.476**	.000	.000	−.040	.002	.060	.004	.570	.430
Symbol Search	0.50	.250	**.490**	**.240**	.000	.000	.030	.001	.160	.026	.630	.370
Cancellation	0.28	.078	.190	.036	−.070	.005	−.010	.000	**.610**	**.372**	.230	.770
Total variance		.385		.059		.046		.035		.028	.582	.418
Common variance		.661		.101		.079		.059		.048		
	$\omega_h = .849$		$\omega_s = .162$		$\omega_s = .190$		$\omega_s = .244$					

Note. All factors were rotated using the analytic bifactor rotation b = loading of subtest on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_h = omega hierarchical, ω_s = omega subscale. Bold type indicates salient loading ($b \geq .30$). Italic type indicates coefficient and variance estimate alignment ($.20 \leq b < .30$) with theoretically posited factor.

When extracting three WISC-V factors (Table 4), the results produced three factors that cohered with the theoretical structure presented in the *Technical and Interpretive Manual*. The results produced a Processing Speed (PS) factor that was consistent with the theoretically posited structure (Coding, Symbol Search, and Cancellation), although Vocabulary had a negative loading on this factor. The second factor was comprised of the Working Memory subtests (Digit Span, Picture Span, and Letter–Number Sequencing). The third factor measured Perceptual Reasoning, which combined Block Design and Visual Puzzles, with relatively strong loadings, and Matrix Reasoning and Figure Weights, with relatively weak loadings. Similarities, Information, Comprehension, Picture Concepts, and Arithmetic *did not* saliently load any group factor.

When extracting two WISC-V factors (Table 5), the results suggested that two factors cohered to varying degrees with the theoretical structure presented in the *Technical and Interpretive Manual*. The first factor was a Processing Speed factor that is consistent with the structure presented in the *Technical and Interpretive Manual* (Coding, Symbol Search, and Cancellation), but Similarities, Vocabulary, and Information had negative loadings on this factor. The second factor was comprised of the Working Memory subtests (Digit Span, Picture Span, and Letter–Number Sequencing) as well as Arithmetic, although its loading was relatively small. Comprehension, Block Design, Visual Puzzles, Matrix Reasoning, Figure Weights, and Picture Concepts *did not* saliently load any group factor.

Table 4
Sources of variance in the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V) for the total standardization sample ($N = 2200$) according to exploratory bifactor analysis with three first-order factors.

	General		F1: Processing Speed		F2: Working Memory		F3: Perceptual Reasoning			
WISC-V subtest	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	h^2	u^2
Similarities	0.79	.624	−.140	.020	−.060	.004	−.080	.006	.654	.346
Vocabulary	0.82	.672	*−.200*	.040	−.120	.014	−.090	.008	.735	.265
Information	0.79	.624	−.150	.023	−.100	.010	−.060	.004	.660	.340
Comprehension	0.69	.476	−.070	.005	−.040	.002	−.130	.017	.500	.501
Block Design	0.63	.397	.090	.008	−.020	.000	**.400**	**.160**	.565	.435
Visual Puzzles	0.64	.410	−.060	.004	−.010	.000	**.440**	**.194**	.607	.393
Matrix Reasoning	0.60	.360	.010	.000	.120	.014	*.230*	*.053*	.427	.573
Figure Weights	0.61	.372	−.080	.006	.070	.005	*.260*	*.068*	.451	.549
Picture Concepts	0.51	.260	−.020	.000	.020	.000	.090	.008	.269	.731
Arithmetic	0.70	.490	.010	.000	.190	.036	.060	.004	.530	.470
Digit Span	0.64	.410	.030	.001	**.500**	**.250**	.030	.001	.661	.339
Picture Span	0.53	.281	.040	.002	**.340**	**.116**	.050	.003	.401	.599
Letter–Number Sequencing	0.64	.410	.000	.000	**.460**	**.212**	−.060	.004	.625	.375
Coding	0.40	.160	**.630**	**.397**	.040	.002	−.030	.001	.559	.441
Symbol Search	0.46	.212	**.620**	**.384**	−.010	.000	.020	.000	.597	.404
Cancellation	0.22	.048	**.360**	**.130**	−.080	.006	.030	.001	.185	.815
Total variance		.388		.064		.042		.033	.527	.473
Common variance		.736		.121		.080		.063		
	$\omega_h = .802$		$\omega_s = .159$		$\omega_s = .240$		$\omega_s = .281$			

Note. b = loading of subtest on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_h = omega hierarchical, ω_s = omega subscale.
Bold type indicates salient loading ($b \geq .30$). Italic type indicates coefficient and variance estimate alignment ($.20 \leq b < .30$) with theoretically posited factor.

Table 5
Sources of variance in the Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V) for the total standardization sample ($N = 2200$) according to exploratory bifactor analysis with two first-order factors.

	General		F1: Processing Speed		F2: Working Memory			
WISC-V subtest	b	S^2	b	S^2	b	S^2	h^2	u^2
Similarities	0.76	.578	*−.200*	*.040*	.020	.000	.618	.382
Vocabulary	0.79	.624	*−.270*	*.073*	−.030	.001	.698	.302
Information	0.77	.593	*−.220*	*.048*	−.020	.000	.642	.358
Comprehension	0.66	.436	−.130	.017	.030	.001	.453	.557
Block Design	0.67	.449	.040	.002	.030	.001	.451	.549
Visual Puzzles	0.67	.449	−.080	.006	.040	.002	.457	.543
Matrix Reasoning	0.62	.384	−.030	.001	.150	.023	.408	.592
Figure Weights	0.62	.384	−.011	.012	.110	.012	.409	.591
Picture Concepts	0.52	.270	−.050	.003	.060	.004	.277	.724
Arithmetic	0.69	.476	−.030	.001	*.240*	*.058*	.535	.465
Digit Span	0.60	.360	.010	.000	**.550**	**.303**	.663	.337
Picture Span	0.51	.260	.020	.000	**.370**	**.137**	.397	.603
Letter–Number Sequencing	0.60	.360	−.020	.000	**.500**	**.250**	.610	.390
Coding	0.44	.194	**.590**	**.348**	.040	.002	.543	.457
Symbol Search	0.51	.260	**.590**	**.348**	−020	.000	.609	.391
Cancellation	0.25	.063	**.340**	**.116**	−.090	.008	.186	.815
Total variance		.384		.063		.050	.497	.503
Common variance		.772		.128		.101		
	$\omega_h = .828$		$\omega_s = .250$		$\omega_s = .185$			

Note. b = loading of subtest on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_h = omega hierarchical, ω_s = omega subscale. Bold type indicates salient loading ($b \geq .30$). Italic type indicates coefficient and variance estimate alignment ($.20 \leq b < .30$) with theoretically posited factor.

4. Discussion

The WISC-V factor structure was examined using exploratory bifactor factor analysis (EBFA; i.e., exploratory factor analysis with a bi-factor rotation). This posits that the general factor *directly* influences performance on the subtests instead of *indirectly* influencing subtest performance through full mediation by the first-order factors (Gignac, 2008). Neither the five- nor four-factor models were appropriate for the WISC-V subtests using EBFA because they produced ill-defined factors. Instead, results indicated that a three factor solution was the most plausible, consisting of Processing Speed (PS), Working Memory (WM), and Perceptual Reasoning (PR) group factors. Block Design, Visual Puzzles, Matrix Reasoning, and Figure Weights all converged to form a single Perceptual Reasoning factor rather than separate Visual Spatial and Fluid Reasoning factors. The Processing Speed and Working Memory factors contained all subtests suggested by the *Technical and Interpretive Manual*, although many of the Verbal Comprehension (VC) subtests also loaded negatively on the PS factor.

Such bipolar factors are common in unrotated factor solutions, and typically suggest that the factors require rotation to eliminate the negative loadings (Comrey & Lee, 1992). As this bipolar factor is already rotated, we suggest that the small negative loadings of the VC subtests reflect the nature of the PS construct. Specifically, the PS subtests rely strongly on speed of information processing because they all have time constraints. However, the Verbal Comprehension subtests have no time constraints so are minimally constrained by information processing speed. Thus, the small negative loadings of the Verbal Comprehension subtests bolster the interpretation of this factor as one representing Processing Speed.

Neither the three-factor model nor any other model showed evidence of distinct Fluid Reasoning or Verbal Comprehension factors. The lack of evidence for a Fluid Reasoning factor is consistent with Canivez et al.'s (2015, 2014) analysis of the WISC-V, as well as the analysis of other Wechsler scales (e.g., Beaujean et al., 2014). The failure to find a Verbal Comprehension factor is inconsistent with the theoretical structure presented in the WISC-V *Technical and Interpretive Manual* and with the extant literature regarding the Wechsler scales (Boake, 2002; Kamphuas, 2005), and is a more unusual and perplexing finding. Wechsler scales have traditionally had an excess of verbal subtests and factor models have traditionally found evidence for the Verbal Comprehension factor (e.g. Wechsler, 1939; Wechsler, 2003; Wechsler, 2008; Wechsler, 2012). Perhaps the results from EBFA of the WISC-V data merely reflect the arbitrary nature of group factors, which Spearman (1933) noted many years ago:

> Moreover, this secondary subdivision [into group factors], unlike the primary bisection [into general and specific factors], is unstable …. Speaking generally, all these sub-divisions of an ability depend on what other abilities we choose to put into one and the same set; they therefore come and go at our will. Whereas the primary bisection into universal and non-universal factors remains inviolate; it is not dependent on any chance composition of a particular set of abilities, but instead marks the most fundamental feature in ability as a whole (p. 600).

There may be questions as to why the present analysis produced a different outcome (i.e., did not locate a verbal factor) from the Schmid–Leiman procedure that Canivez et al. (2015, 2014) used. While bifactor models are often thought to be functionally equivalent to higher-order models with a SL transformation (e.g., Weiss, Keith, Zhu, & Chen, 2013), they are very different, both conceptually and methodologically (Beaujean, 2015; Frisby & Beaujean, 2015; Jennrich & Bentler, 2011; Reise, 2012). In a bifactor model, all factors – including *g* – are first-order, so are extracted directly from the indicator variables. This is different from higher-order models, which posit *g* as a second- (or higher) order factor, extracted from first- (or higher) order factors. In bifactor models the factors are all competing with each other to explain the subtests' covariance. Typically, *g* is formed first and then group factors are formed from any of the remaining covariance unexplained by *g*. In higher-order models, the first-order factors are formed first without any reference to *g*; then, *g* is formed from already identified first-order oblique factors. The SL rotation of such higher-order models simultaneously calculates all the subtests' indirect relationship to *g* and the group factors' residuals/error; it does not produce a bifactor model as *g*'s relationship to the subtests is still indirect. For a more in-depth explanation, see Beaujean (2015) and Beaujean et al. (2014).

In the current exploratory bifactor analysis of the WISC-V subtest data, it appears that *g* explains all that is common among the verbal subtests, so the residual covariance from these tests is best explained by specific factors and error variance. This explanation is supported by the consistency of not finding a VC factor for any of the models examined. If a VC factor was not found because the three-factor model extracted too few factors, then it should have been present when extracting more factors (i.e., four- or five-factor model), but this did not occur. Of note, it is not atypical to find some group factors diminish when using a bifactor model/rotation if the general factor is well defined (e.g., Beaujean et al., 2014).

By contrast, higher-order models require group factors be extracted first, followed by the extraction of *g*. For *g* to have strong indirect relationships to the subtests (which is what the SL transformation is calculating), two things must occur: subtests must have strong loadings on group factors and group factors must have strong loadings on *g*. If either of these conditions is not met, *g*'s relationship to a given subtest will be small. Consequently, in higher-order models of the WISC-V data, if a Verbal Comprehension factor was not strongly defined (or the verbal subtests had strong loadings on a different factor), then *g* would not have a sizable relationship with these subtests, which is contrary to the corpus of literature examining the relationship with verbal abilities and *g* (e.g. Carroll, 1993; Jensen, 2001).

In summary, the results of this study revealed that the WISC-V is primarily a measure of *g*, as it accounts for a majority of the subtests' total and common variance. The preeminence of *g* found in this study is similar to the findings of other studies of Wechsler scales using both EFA and CFA methods (Bodin, Pardini, Burns, & Stevens, 2009; Canivez, 2014b; Canivez & Watkins, 2010a, 2010b; Gignac & Watkins, 2013; Nelson, Canivez, & Watkins, 2013; Watkins, 2006, 2010; Watkins & Beaujean, 2014; Watkins et al., 2013; Watkins, Wilson, Kotz, Carbone, & Babula, 2006) and other intelligence tests (Canivez, 2008; Canivez et al., 2009; Canivez, 2011; DiStefano & Dombrowski, 2006; Dombrowski, 2013, 2014a, 2014b; Dombrowski & Watkins, 2013; Dombrowski et al., 2009; Nelson & Canivez, 2012; Nelson, Canivez, Lindstrom, & Hatt, 2007). Likewise, these results are consistent with the broader professional literature on the importance of general intelligence (Deary, 2013; Jensen, 1998; Lubinski, 2000; Ree, Carretta, & Green, 2003). Unfortunately, a limitation of this analysis was the use of a correlation matrix for factor analysis, but at the time of this research there was no independent access to the WISC-V raw data.

Given that most of the WISC-V variance was contributed by g and that ω_s coefficients were low, primary interpretive emphasis should be placed upon this general factor as manifested in the FSIQ score. These results provide little justification for the clinical interpretation of group factors or their manifestations in the surplus of index scores and comparisons that can be calculated from the WISC-V subtests (Reise, 2012; Reise et al., 2013). The *Standards* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) state that interpretation of subscores requires demonstration of the scores' "distinctiveness and reliability" (Standard 1.14), which do not appear to be present for the WISC-V index scores. Thus, while focusing on WISC-V index scores may be well-intentioned, it may be promulgating inaccurate interpretation paradigms that lead clinicians down a blind alley (Dombrowski, 2015).

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014c). *Standards for educational and psychological testing* (4th ed.). Washington, DC: Authors.

Beaujean, A. A. (2013). Factor analysis using R. *Practical assessment, research, and evaluation*, *18*(4), 1–11 Retrieved from http://pareonline.net/pdf/v18n4.pdf

Beaujean, A. A. (2014). *R syntax to accompany best practices in exploratory factor analysis (2014) by Jason Osborne*. Waco, TX: Baylorhttps://dl.dropboxusercontent.com/u/18489687/EFAbook/BestPracticesEFA_Rcode.pdf

Beaujean, A. A. (2015). *John's views on intelligence: Bi-factor vs. higher-order models.* Manuscript submitted for publication.

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

Boake, C. (2002). From the Binet–Simon to the Wechsler–Bellvue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology*, *24*, 383–405. http://dx.doi.org/10.1076/jcen.24.3.383.981.

Bodin, D., Pardini, D. A., Burns, T. G., & Stevens, A. B. (2009). Higher order factor structure of the WISC-IV in a clinical neuropsychological sample. *Child Neuropsychology*, *15*, 417–424. http://dx.doi.org/10.1080/09297040802603661.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, *80*, 796–846. http://dx.doi.org/10.1111/j.1467–6494.2011.00749.x.

Canivez, G. L. (2008). Orthogonal higher-order factor structure of the Stanford–Binet Intelligence Scales for children and adolescents. *School Psychology Quarterly*, *23*, 533–541. http://dx.doi.org/10.1037/a0012884.

Canivez, G. L. (2010). Review of the Wechsler Adult Intelligence Test—Fourth Edition. In R. A. Spies, J. F. Carlson, & K. F. Geisinger (Eds.), *The eighteenth mental measurements yearbook* (pp. 684–688). Lincoln, NE: Buros Institute of Mental Measurements.

Canivez, G. L. (2011). Hierarchical factor structure of the cognitive assessment system: Variance partitions from the Schmid–Leiman (1957) procedure. *School Psychology Quarterly*, *26*, 305–317. http://dx.doi.org/10.1037/a0025973.

Canivez, G. L. (2014a). Review of the Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition. In J. F. Carlson, K. F. Geisinger, & J. L. Jonson (Eds.), *The nineteenth mental measurements yearbook* (pp. 732–737). Lincoln, NE: Buros Institute of Mental Measurements.

Canivez, G. L. (2014b). Construct validity of the WISC-IV with a referred sample: Direct versus indirect hierarchical structures. *School Psychology Quarterly*, *29*, 38–51. http://dx.doi.org/10.1037/spq0000032.

Canivez, G. L. (2015). 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*. Gottingen, Germany: Hogrefe (in press).

Canivez, G. L., & Kush, J. C. (2013). WISC-IV and WAIS-IV structural validity: Alternate methods, alternate results. Commentary on Weiss et al. (2013a) and Weiss et al. (2013b). *Journal of Psychoeducational Assessment*, *31*, 157–169. http://dx.doi.org/10.1177/0734282913478036.

Canivez, G. L., & Watkins, M. W. (2010a). 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. http://dx.doi.org/10.1037/a0020429.

Canivez, G. L., & Watkins, M. W. (2010b). Exploratory and higher-order factor analyses of the Wechsler Adult Intelligence Scale—Fourth Edition (WAIS-IV) adolescent subsample. *School Psychology Quarterly*, *25*, 223–235. http://dx.doi.org/10.1037/a0022046.

Canivez, G. L., & Watkins, M. W. (2015). Review of the Wechsler Intelligence Scale for Children—Fifth Edition: Critique, commentary, and independent analyses. (Authors) In A. S. Kaufman, S. E. Raiford, & D. L. Coalson (Eds.), *Intelligent testing with the WISC-V*. Hoboken, NJ: Wiley (in press).

Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2015). Factor structure of the Wechsler Intelligence Scale for Children—Fifth Edition: Exploratory factor analyses with the 16 primary and secondary subtests. *Manuscript submitted for publication.*

Canivez, G. L., Watkins, M. W., James, T., James, K., & Good, R. (2014). Incremental validity of WISC-IV^{UK} factor index scores with a referred Irish sample: Predicting performance on the WIAT-II^{UK}. *British Journal of Educational Psychology*, *84*, 667–684. http://dx.doi.org/10.1111/bjep.12056.

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. http://dx.doi.org/10.1207/s15327906mbr3003_6.

Carroll, J. B. (1997). The three-stratum theory of cognitive abilities. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 183–208). New York: Guilford.

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, NY: Pergamon Press.

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

Cattell, R. B., & Horn, J. L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, *15*, 139–164. http://dx.doi.org/10.1111/j.1745–3984.1978.tb00065.x.

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

Deary, I. J. (2013). Intelligence. *Current Biology*, *23*, 673–676. http://dx.doi.org/10.1016/j.cub.2013.07.021.

Development Core Team, R. (2015). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

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

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

Dombrowski, S. C. (2014a). Exploratory bifactor analysis of the WJ-III cognitive in adulthood via the Schmid–Leiman procedure. *Journal of Psychoeducational Assessment*, *32*, 330–341. http://dx.doi.org/10.1177/0734282913508243.

Dombrowski, S. C. (2014b). 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. http://dx.doi.org/10.1177/0734282914530838.

Dombrowski, S. C. (2015). *Psychoeducational assessment and reporting.* New York, NY: Springer Science.

Dombrowski, S. C., & Watkins, M. W. (2013). Exploratory and higher order factor analysis of the WJ-III full test battery: A school aged analysis. *Psychological Assessment*, *25*, 442–455. http://dx.doi.org/10.1037/a0031335.

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. http://dx.doi.org/10.1177/0734282909333179.

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. http://dx.doi.org/10.1016/j.intell.2006.07.002.

Frisby, C. L., & Beaujean, A. A. (2015). Testing Spearman's hypotheses using a bi-factor model. *Intelligence*, *51*, 79–97. http://dx.doi.org/10.1016/j.intell.2015.04.007.

Gignac, G. E. (2005). Revisiting the factor structure of the WAIS-R: Insights through nested factor modeling. *Assessment*, *12*, 320–329. http://dx.doi.org/10.1177/1073191105278118.

Gignac, G. E. (2006). The WAIS-III as a nested factors model: A useful alternative to the more conventional oblique and higher-order models. *Journal of Individual Differences*, *27*, 73–86. http://dx.doi.org/10.1027/1614–0001.27.2.73.

Gignac, G. E. (2008). Higher-order models versus direct hierarchical models: *g* as superordinate or breadth factor? *Psychology Science Quarterly*, *50*, 21–43.

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. http://dx.doi.org/10.1080/00273171.2013.804398.

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

Gustafsson, J. -E., & Aberg-Bengtsson, L. (2010). Unidimensionality and interpretability of psychological instruments. In S. E. Embretson (Ed.), *Measuring psychological* *constructs: Advances in model-based approaches* (pp. 97–121). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/12074-005.

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

Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. In K. S. McGrew, J. K. Werder, & R. W. Woodcock (Eds.), *Woodcock–Johnson technical manual* (pp. 197–232) (Rev. ed.). Itasca, IL: Riverside.

Horn, J. L., & Blankson, A. N. (2012). Foundations for better understanding of cognitive abilities. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 73–98) (3rd ed.). New York, NY: Guilford.

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. http://dx.doi.org/10.1037/h0023816.

Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, *112*(2), 351.

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

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

Jensen, A. R. (2001). Vocabulary and general intelligence. *Behavioral and Brain Science*, *24*, 1109–1110.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151. http://dx.doi.org/10.1177/001316446002000116.

Kamphaus, R. W. (2005). *Clinical assessment of child and adolescent intelligence*. New York, NY: Springer Science.

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

Le, H., Schmidt, F. L., Harter, J. K., & Lauver, K. J. (2010). The problem of empirical redundancy of constructs in organizational research: An empirical investigation. *Organizational Behavior and Human Decision Processes*, *112*, 112–125. http://dx.doi.org/10.1016/j.obhdp.2010.02.003.

Lubinski, D. (2000). Scientific and social significance of assessing individual differences: "Sinking shafts at a few critical points.". *Annual Review of Psychology*, *51*, 405–444. http://dx.doi.org/10.1146/annurev.psych.51.1.405.

McCloskey, G., Whitaker, J., Murphy, R., & Rogers, J. (2012). Intellectual, cognitive, and neuropsychological assessment in three-tier service delivery systems in schools. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (pp. 852–881) (3rd ed.). New York, NY: Guilford Press.

Miller, D. C., & Maricle, D. E. (2012). The emergence of neuropsychological constructs into tests of intelligence and cognitive abilities. In D. P. Flanagan, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 800–819) (3rd ed.). New York, NY: Guilford Press.

Millsap, R. E. (2007). Structural equation modeling made difficult. *Personality and Individual Differences*, *42*, 875–881. http://dx.doi.org/10.1016/j.paid.2006.09.021.

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. http://dx.doi.org/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. http://dx.doi.org/10.1016/j.jsp.2007.03.003.

Nelson, J. M., Canivez, G. L., & Watkins, M. W. (2013). Structural and incremental validity of the Wechsler Adult Intelligence Scale—Fourth Edition (WAIS-IV) with a clinical sample. *Psychological Assessment*, *25*, 618–630. http://dx.doi.org/10.1037/a0032086.

Olsson, U. H., Foss, T., Troye, S. V., & Howell, R. D. (2000). The performance of ML, GLS and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. *Structural Equation Modeling*, *7*(4), 557–595.

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, NY: Pergamon Press.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696. http://dx.doi.org/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. http://dx.doi.org/10.1080/00223891.2012.725437.

Revelle, W. (2012). *Psych: Procedures for psychological, psychometric, and personality research (version 1.2.4) [computer software]*. Evanston, IL: Northwestern University.

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

Schmiedek, F., & Li, S. -C. (2004). Toward an alternative representation for disentangling age associated differences in general and specific cognitive abilities. *Psychology and Aging*, *19*, 40–56. http://dx.doi.org/10.1037/0882-7974.19.1.40.

Shrout, P. E., & Lane, S. P. (2012). Reliability. In H. Cooper (Ed.), *APA handbook of research methods in psychology: Foundations, planning, measures, and psychometrics*, *Vol. 1*. (pp. 643–660). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/13619-034.

Spearman, C. (1904). "General intelligence:" Objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Spearman, C. E. (1933). The factor theory and its troubles. III. Misrepresentation of the theory. *Journal of Educational Psychology*, *24*, 591–601. http://dx.doi.org/10.1037/h0073929.

Strauss, E., Spreen, O., & Hunter, M. (2000). Implications of test revisions for research. *Psychological Assessment*, *12*, 237–244. http://dx.doi.org/10.\037/11040–3590.12.3.237.

Thurstone, L. L. (1947). *Multiple-factor analysis*. Chicago, IL: University of Chicago Press.

Velicer, W. F. (1976). Determining the number of components form the matrix of partial correlations. *Psychometrika*, *31*, 321–327. http://dx.doi.org/10.1007/BF02293557.

Watkins, M. W. (2006). Orthogonal higher-order structure of the Wechsler Intelligence Scale for Children—Fourth Edition. *Psychological Assessment*, *18*, 123–125. http://dx.doi.org/10.1037/1040–3590.18.1.123.

Watkins, M. W. (2010). Structure of the Wechsler Intelligence Scale For Children — Fourth Edition among a national sample of referred students. *Psychological Assessment*, *22*, 782–787. http://dx.doi.org/10.1037/a0020043.

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

Watkins, M. W., & Beaujean, A. A. (2014). Bifactor structure of the Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition. *School Psychology Quarterly*, *29*, 52–63. http://dx.doi.org/10.1037/spq0000038.

Watkins, M. W., Canivez, G. L., James, T., James, K., & Good, R. (2013). Construct validity of the WISC-IV^{UK} with a large referred Irish sample. *International Journal of School & Educational Psychology*, *1*, 102–111. http://dx.doi.org/10.1080/21683603.2013.794439.

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. http://dx.doi.org/10.1177/0013164406288168.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore, MD: Williams & Wilkins.

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

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition*. San Antonio, TX: NCS Pearson.

Wechsler, D. (2012). *Wechsler Preschool And Primary Scale of Intelligence—Fourth Edition*. San Antonio, TX: NCS Pearson.

Wechsler, D. (2014a). *Wechsler Intelligence Scale for Children—Fifth Edition*. San Antonio, TX: NCS Pearson.

Wechsler, D. (2014b). *Wechsler Intelligence Scale for Children—Fifth Edition technical and interpretive manual*. San Antonio, TX: NCS Pearson.

Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013). Technical and practical issues in the structure and clinical invariance of the Wechsler scales: A rejoinder to commentaries. *Journal of Psychoeducational Assessment*, *31*(2), 235–243. http://dx.doi.org/10.1177/0734282913478050.

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. http://dx.doi.org/10.1037/1082–989X.1.4.354.

Yuan, K. -H., & Bentler, P. M. (2004). On chi-square difference and z tests in mean and covariance structure analysis when the based model is misspecified. *Educational and Psychological Measurement*, *64*, 737–757. http://dx.doi.org/10.1177/0013164404264853.

Yuan, K. -H., & Chan, W. (2005). On nonequivalence of several procedures of structural equation modeling. *Psychometrika*, *70*(4), 791–798.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's alpha, Revelle's beta, and McDonald's omega h: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*, 123–133. http://dx.doi.org/10.1007/s11336–003–0974–7.

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. http://dx.doi.org/10.1177/0013164496056003006.