

Structural Validity of the Wechsler Intelligence Scale for Children–Fifth Edition: Confirmatory Factor Analyses With the 16 Primary and Secondary Subtests

Gary L. Canivez
Eastern Illinois University

Marley W. Watkins
Baylor University

Stefan C. Dombrowski
Rider University

The factor structure of the Wechsler Intelligence Scale for Children–Fifth Edition (WISC-V; Wechsler, 2014a) standardization sample ($N = 2,200$) was examined using confirmatory factor analyses (CFA) with maximum likelihood estimation for all reported models from the WISC-V *Technical and Interpretation Manual* (Wechsler, 2014b). Additionally, alternative bifactor models were examined and variance estimates and model-based reliability estimates (ω coefficients) were provided. Results from analyses of the 16 primary and secondary WISC-V subtests found that all higher-order CFA models with 5 group factors (VC, VS, FR, WM, and PS) produced model specification errors where the Fluid Reasoning factor produced negative variance and were thus judged inadequate. Of the 16 models tested, the bifactor model containing 4 group factors (VC, PR, WM, and PS) produced the best fit. Results from analyses of the 10 primary WISC-V subtests also found the bifactor model with 4 group factors (VC, PR, WM, and PS) produced the best fit. Variance estimates from both 16 and 10 subtest based bifactor models found dominance of general intelligence (g) in accounting for subtest variance (except for PS subtests) and large ω -hierarchical coefficients supporting general intelligence interpretation. The small portions of variance uniquely captured by the 4 group factors and low ω -hierarchical subscale coefficients likely render the group factors of questionable interpretive value independent of g (except perhaps for PS). Present CFA results confirm the EFA results reported by Canivez, Watkins, and Dombrowski (2015); Dombrowski, Canivez, Watkins, and Beaujean (2015); and Canivez, Dombrowski, and Watkins (2015).

Keywords: WISC-V, confirmatory factor analysis, bifactor model, hierarchical CFA, intelligence

The Wechsler Intelligence Scale for Children–Fifth Edition (WISC-V; Wechsler, 2014a) is the newest version of one of the most frequently used tests of cognitive abilities among school and clinical psychologists (Alfonso, Oakland, LaRocca, & Spanakos, 2000; Alfonso & Pratt, 1997; Belter, & Piotrowski, 2001; Goh, Teslow, & Fuller, 1981; Hutton, Dubes, & Muir, 1992; Kaufman & Lichtenberger, 2000; Pfeiffer, Reddy, Kletzel, Schmelzer, & Boyer, 2000; Stinnett, Havey, & Oehler-Stinnett, 1994; Watkins, Campbell, Nieberding, & Hallmark, 1995). Worldwide use of Wechsler scales (Flanagan & Kaufman, 2009) has been facilitated by numerous translations and adaptations of the Wechsler Intelli-

gence Scale for Children–Third Edition (WISC–III; cf., Georgas, van de Vijver, Weiss, and Saklofske (2003), Wechsler Intelligence Scale for Children–Fourth Edition (WISC–IV; cf., Grégoire et al., 2008), Wechsler Preschool and Primary Scale of Intelligence (WPPSI; cf., Liu & Lynn, 2011), and the Wechsler Adult Intelligence Scale–Third Edition (WAIS–III; cf., Golay & Lecerf, 2011).

Development and construction of the WISC-V was reported to reflect conceptualizations of intellectual measurement influenced by Carroll, Cattell, and Horn (Carroll, 1993, 2003; Cattell & Horn, 1978; Horn, 1991; Horn & Blankson, 2005; Horn & Cattell, 1966) as well as other neuropsychological constructs (Wechsler, 2014b). The Word Reasoning and Picture Completion subtests from the WISC-IV were eliminated and three new subtests were added. Picture Span (adapted from the Wechsler Preschool and Primary Scale of Intelligence–Fourth Edition [Wechsler, 2012]) was added to measure visual working memory and Visual Puzzles and Figure Weights (adapted from the Wechsler Adult Intelligence Scale–Fourth Edition [Wechsler, 2008]) were added to measure visual spatial and fluid reasoning, respectively. One major revision goal in WISC-V development was to divide the former Perceptual Reasoning factor into distinct Visual Spatial and Fluid Reasoning factors. Similar attempts were made with the WAIS-IV (Weiss, Keith, Zhu, & Chen, 2013a) and WISC-IV (Weiss, Keith, Zhu, & Chen, 2013a), but Canivez and Kush (2013) pointed out numerous

This article was published Online First July 21, 2016.

Gary L. Canivez, Department of Psychology, Eastern Illinois University; Marley W. Watkins, Department of Educational Psychology, Baylor University; Stefan C. Dombrowski, School Psychology Program, Rider University.

This research was partially supported by a 2015 Summer Research Grant from the Council on Faculty Research, Eastern Illinois University to Gary L. Canivez.

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

psychometric problems with these proposed five-factor higher-order models in both the WAIS-IV and WISC-IV.

The WISC-V includes seven “Primary” subtests (Similarities [SI], Vocabulary [VC], Block Design [BD], Matrix Reasoning [MR], Figure Weights [FW], Digit Span [DS], and Coding [CD]) that produce the FSIQ and three additional “Primary” subtests (Visual Puzzles [VP], Picture Span [PS], and Symbol Search [SS]) to produce the five factor index scores (two subtests each for Verbal Comprehension [VC], Visual Spatial [VS], Fluid Reasoning [FR], Working Memory [WM], and Processing Speed [PS]). There are six “Secondary” subtests (Information [IN], Comprehension [CO], Picture Concepts [PC], Arithmetic [AR], Letter-Number Sequencing [LN], and Cancellation [CN]) that are used either for substitution in FSIQ estimation (when one primary subtest is spoiled) or in estimating newly created (Quantitative Reasoning, Auditory Working Memory, Nonverbal) and previously existing (General Ability and Cognitive Proficiency) Ancillary Index Scores.

WISC-V structural validity evidence reported in the WISC-V *Technical and Interpretive Manual* (Wechsler, 2014b) was based exclusively on confirmatory factor analyses (CFA) that evaluated numerous models beginning with a one-factor (g) model. All other models with two through five first-order factors were *higher-order* models with the general intelligence factor (g) *indirectly* influencing subtests via full mediation through the first-order factors (Yung, Thissen, & McLeod, 1999). All CFA models tested are illustrated with subtest assignments to latent factors in the WISC-V *Technical and Interpretive Manual* (Table 5.3). The standardized measurement model for the final five-factor higher-order model of the WISC-V primary and secondary subtests with the total standardization sample is presented in the WISC-V *Technical and Interpretive Manual* Figure 5.10 (reproduced in modified form here as Figure 1). This model includes a higher-order general intelligence dimension with five first-order factors (VC, VS, FR, WM, and PS) and the 16 subtest indicators are uniquely assigned to one latent first-order factor except for Arithmetic, which cross-loads on VC, FR, and WM. This final measurement model includes a standardized path coefficient of 1.00 between the higher-order general intelligence factor and the FR factor; suggesting they may be redundant. This final model was also reported to fit five different WISC-V age groups (6–7, 8–9, 10–11, 12–13, and 14–16) equally well (Wechsler, 2014b). Subsequently, H. Chen, Zhang, Raiford, Zhu, and Weiss (2015) reported factorial invariance of this final higher-order model across gender, although they did not examine invariance for alternative models.

Although information presented in the WISC-V *Technical and Interpretive Manual* appears to be quite favorable, a number of concerns were pointed out (Canivez, Dombrowski, & Watkins, 2015; Canivez & Watkins, 2016; Canivez, Watkins, & Dombrowski, 2015) including use of Weighted Least Squares (WLS) estimation (without explicit justification) rather than maximum likelihood (ML) estimation, retention of a complex CFA measurement model (cross-loading Arithmetic on three group factors) thereby abandoning parsimony of simple structure, possible redundancy of FR and g because of the standardized path coefficient of 1.0 between general intelligence and the Fluid Reasoning factor, no consideration or testing of bifactor models, omission of decomposed sources of variance between the higher-order general intelligence factor and lower-order group factors, and absence of latent

factor reliabilities estimated for general intelligence and the lower-order group factors. Many of these concerns were previously identified and discussed with other Wechsler scale versions (Canivez, 2010, 2014a; Canivez & Kush, 2013; Gignac & Watkins, 2013) but were not addressed in the WISC-V *Technical and Interpretive Manual*.

Exploratory factor analyses (EFA) were not reported in the WISC-V *Technical and Interpretive Manual*. Results of independent EFA of the WISC-V did not support the existence of five-factors in the total WISC-V standardization sample (Canivez, Watkins, & Dombrowski, 2015) or in four age groups (6–8, 9–11, 12–14, and 15–16) within the WISC-V standardization sample (Canivez, Dombrowski, & Watkins, 2015) as the fifth extracted factor included only one salient subtest loading. Schmid and Leiman (1957) orthogonalization of the second-order EFA for the total WISC-V standardization sample and the four age groups found substantial portions of variance apportioned to the hierarchical general factor and substantially smaller portions of variance apportioned to the group factors. Omega-hierarchical coefficients (Reise, 2012; Rodriguez, Reise, & Haviland, 2015) for the general factor ranged from .817 (Canivez, Watkins, & Dombrowski, 2015) to .847 (Canivez, Dombrowski, & Watkins, 2015) and exceeded the preferred level (.75) for clinical interpretation (Reise, 2012; Reise, Bonifay, & Haviland, 2013). Omega-subscale (ω_s) coefficients (Reise, 2012) for the four WISC-V group factors ranged from .131 to .530 and no ω_s coefficients for VC, PR, or WM approached or exceeded the minimum criterion (.50) for clinical interpretation (Reise, 2012; Reise et al., 2013). However, ω_s coefficients for PS approached or exceeded the .50 criterion for possible clinical interpretation. Dombrowski, Canivez, Watkins, and Beaujean (2015) also failed to find support for five-factors in the total WISC-V standardization sample using exploratory bifactor analysis through the bifactor rotation criterion (Jennrich & Bentler, 2011).

Independent CFA results produced with other Wechsler scales have found bifactor models to fit data as well or better than higher-order models and variance estimates for the general intelligence factor has far exceeded variance estimates of the group factors (Canivez, 2014b; Gignac, 2005, 2006; Gignac & Watkins, 2013; Golay, Reverte, Rossier, Favez, & Lecerf, 2013; Nelson, Canivez, & Watkins, 2013; Watkins, 2010; Watkins & Beaujean, 2014; Watkins, Canivez, James, James, & Good, 2013). Given such results in the literature and the advantages of bifactor modeling for understanding test structure (Canivez, 2016; Reise, 2012; Gignac, 2008), comparisons of bifactor models to the higher-order models for the WISC-V are needed.

Carroll (1995) insisted on the use of the Schmid and Leiman (1957) transformation of EFA loadings to apportion subtest variance to the first-order and higher-order dimensions because intelligence test subtests are influenced by both first-order factors and the higher-order g factor. Interpretation of higher-order models requires this partitioning of variance in CFA as well as EFA to determine the relative influence of the first-order factors in comparison to the higher-order factor(s). Within CFA models, a higher-order representation of intelligence test structure is an indirect hierarchical model (Gignac, 2005, 2006, 2008) where the g factor influences subtests *indirectly* through full mediation through the first-order factors (Yung et al., 1999). Thus, g is conceptualized as a *superordinate* factor and is an abstraction from abstrac-

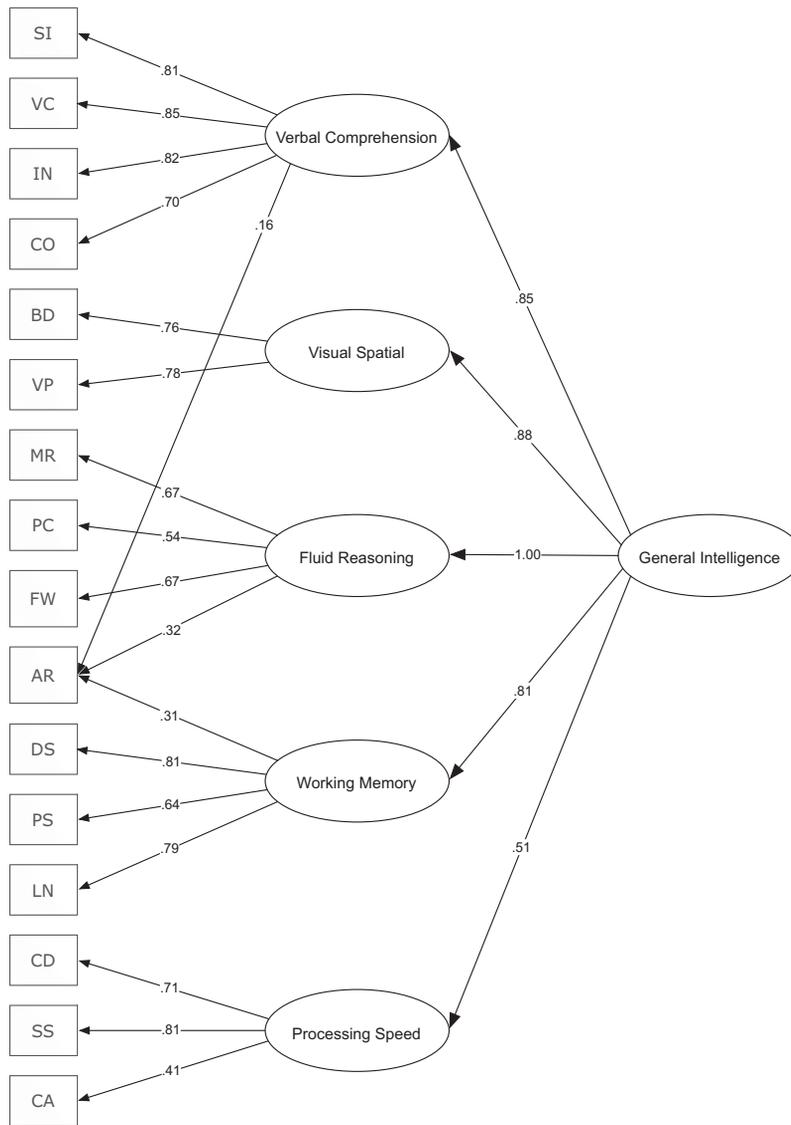


Figure 1. Higher-order measurement model with standardized coefficients (adapted from Figure 5.1; Wechsler, 2014b), for WISC-V standardization sample ($N = 2,200$) 16 Subtests. SI = Similarities; VC = Vocabulary; IN = Information; CO = Comprehension; BD = Block Design; VP = Visual Puzzles; MR = Matrix 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 (WISC-V). Copyright, 2014 NCS Pearson, Inc. Reproduced with permission. All rights reserved. "Wechsler Intelligence Scale for Children" and "WISC" are trademarks, in the United States and/or other countries, of Pearson Education, Inc. or its affiliates(s).

tions (Thompson, 2004). While higher-order models have been commonly applied to assess the "construct-relevant psychometric multidimensionality" (Morin, Arens, & Marsh, 2016, p. 117) of intelligence tests, an alternative conceptualization was originally specified by Holzinger and Swineford (1937) as the bifactor model (alternatively referred to as a direct hierarchical [Gignac, 2005, 2006, 2008] or nested factors model [Gustafsson, & Balke, 1993]). In bifactor models, both the general (g) and the group factors directly influence the subtests and g is conceptualized as a *breadth* factor (Gignac, 2008). This means that both g and first-order group

factors are simultaneous abstractions derived from the observed subtest indicators and therefore a less complicated conceptual model (Gignac, 2008).

Canivez (2016) and Reise (2012) noted several advantages of the bifactor (direct hierarchical/nested factors) model including the direct influences of the general factor are easy to interpret, both general and specific influences on indicators (subtests) can be examined simultaneously, and the psychometric properties necessary for determining scoring and interpretation of subscales can be directly examined. Gignac (2006) also noted that the direct hier-

archical model (bifactor) can also be considered to be more parsimonious because it specifies a unidimensional general factor.

Understanding the structural validity of tests is essential for evaluating interpretability of provided scores (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). Given the results from EFA suggesting the WISC-V was overfactored as presented in the WISC-V *Technical and Interpretive Manual* and that reported WISC-V CFA did not: (a) use ML estimation, (b) compare bifactor models to higher-order models, (c) decompose variance estimates, nor (d) report latent factor reliabilities; the present study examines these methods by independently conducting CFA with the total WISC-V standardization sample. CFA with ML estimation was used to examine the WISC-V structure with the 16 primary and secondary subtests *and* also with the 10 primary subtests, while also examining bifactor models, presenting decomposed variance estimates, and estimating bifactor model-based reliabilities.

Method

Participants

NCS Pearson, Inc. denied without rationale our request of WISC-V standardization sample raw data to conduct these independent analyses. Absent raw data, the summary statistics (correlations and descriptive statistics) in the WISC-V *Technical and Interpretive Manual* (Table 5.1; Wechsler, 2014b) used in our analyses were produced by participants who were members of the full WISC-V standardization sample ($N = 2,200$) and ranged in age from 6–16 years. Detailed demographic characteristics provided in the WISC-V *Technical and Interpretive Manual* illustrated the demographically representative standardization sample obtained using stratified proportional sampling across variables of age, sex, race/ethnicity, parental education level, and geographic region. Demographic information revealed a close match to the U.S. census across stratification variables.

Instrument

The WISC-V, an individual general intelligence test for children aged 6–16 years, like the WISC-IV (Wechsler, 2003), overlaps in age with the WPPSI-IV (Wechsler, 2012) at age 6 years and the WAIS-IV at age 16 years. This allows clinicians the option of selecting the most appropriate instrument depending on the referral question and child characteristics. Consistent with Wechsler's definition of intelligence (i.e., "global capacity;" Wechsler, 1939, p. 229) the WISC-V includes numerous subtests that provide estimates of general intelligence but also are combined to measure various group factors.

WISC-V organization and subtest administration order reflect a new four-level organization. The Full Scale IQ (FSIQ) is composed of seven primary subtests across the five factors (VC, VS, FR, WM, and PS), but if one of the FSIQ subtests is invalid or missing, that subtest may be substituted by a secondary subtest from within the same factor; however, only one substitution is allowed. The Primary Index Scale level is composed of 10 WISC-V subtests (primary subtests) and are used to estimate the five WISC-V factor index scores (VCI, VSI, FRI, WMI, and PSI).

No substitutions are allowed for the Primary Index Scales. The Ancillary Index level is composed of five scales that are not factorially derived: Quantitative Reasoning (QR), Auditory Working Memory (AWM), Nonverbal (NV), General Ability (GA), and Cognitive Proficiency (CP) and reflect various combinations of primary and secondary subtests. The Complementary Index level is composed of three scales: Naming Speed, Symbol Translation, and Storage and Retrieval derived from the newly created complementary subtests (Naming Speed Literacy, Naming Speed Quality, Immediate Symbol Translation, Delayed Symbol Translation, and Recognition Symbol Translation). Complementary subtests are *not* intelligence subtests and may not be substituted for primary or secondary subtests. As such these are not included in present analyses nor were they included in WISC-V CFAs reported in the WISC-V *Technical and Interpretive Manual* (Wechsler, 2014b).

Analyses

EQS 6.2 (Bentler & Wu, 2012) was used to conduct CFA using maximum likelihood estimation. Covariance matrices were produced for CFA using the correlation matrix, means, and SDs from the total WISC-V standardization sample presented in the WISC-V *Technical and Interpretive Manual* (Table 5.1). Some first-order factors were underidentified because they were measured by only two subtests. In those CFA, the two subtests were constrained to equality before estimating bifactor models to ensure identification (Little, Lindenberger, & Nesselroade, 1999).

The structural models specified in Table 5.3 of the WISC-V *Technical and Interpretive Manual* are reproduced in Figures 2–4 with the addition of alternative bifactor models that were not included in analyses reported in the WISC-V *Technical and Interpretive Manual*. Although there are no universally accepted cutoff values for approximate fit indices (McDonald, 2010), overall model fit was evaluated using the comparative fit index (CFI), standardized root mean squared residual (SRMR), Tucker-Lewis index (TLI), and the root mean square error of approximation (RMSEA). Higher values indicate better fit for the CFI and TLI whereas lower values indicate better fit for the SRMR and RMSEA. Applying the Hu and Bentler (1999) combinatorial heuristics, criteria for adequate model fit were CFI and TLI $\geq .90$ along with SRMR $\leq .09$ and RMSEA $\leq .08$. Good model fit required CFI and TLI ≥ 0.95 with SRMR and RMSEA ≤ 0.06 (Hu & Bentler, 1999). For a model to be considered superior, it had to exhibit adequate to good overall fit *and* display meaningfully better fit ($\Delta\text{CFI} > .01$ and $\Delta\text{RMSEA} > .015$) than alternative models (Cheung & Rensvold, 2002; F. F. Chen, 2007). Additionally, the Akaike Information Criterion (AIC) was considered. AIC does not have a meaningful scale but the model with the smallest AIC values is most likely to replicate (Kline, 2016) and would be preferred.

Model-based reliabilities were estimated with coefficients ω -hierarchical (ω_H) and ω -hierarchical subscale (ω_{HS}), which estimate reliability of unit-weighted scores produced by the indicators (Reise, 2012; Rodriguez et al., 2015). ω_H is the model based reliability estimate for the general intelligence factor with variability of group factors removed. ω_{HS} is the model based reliability estimate of a group factor with all other group *and* general factors removed (Brunner, Nagy, & Wilhelm, 2012; Reise, 2012). Omega

| | Model 2 | Model 3 | Model 4a | Model 4a Bi-factor | Model 4b | Model 4c | Model 4d | Model 4e | Model 4e Bifactor |
|---------|---------|---------|----------|--------------------|----------|----------|----------|----------|-------------------|
| Subtest | F1 F2 | F1F2F3 | F1F2F3F4 | F1F2F3F4 | F1F2F3F4 | F1F2F3F4 | F1F2F3F4 | F1F2F3F4 | F1F2F3F4 |
| SI | • | • | • | •• | • | • | • | • | •• |
| VC | • | • | • | •• | • | • | • | • | •• |
| IN | • | • | • | •• | • | • | • | • | •• |
| CO | • | • | • | •• | • | • | • | • | •• |
| BD | • | • | • | •• | • | • | • | • | •• |
| VP | • | • | • | •• | • | • | • | • | •• |
| MR | • | • | • | •• | • | • | • | • | •• |
| FW | • | • | • | •• | • | • | • | • | •• |
| PC | • | • | • | •• | • | • | • | • | •• |
| AR | • | • | • | •• | • | •• | ••• | • | •• |
| DS | • | • | • | •• | • | • | • | • | •• |
| PS | • | • | • | •• | • | • | • | • | •• |
| LN | • | • | • | •• | • | • | • | • | •• |
| CD | • | • | • | •• | • | • | • | • | •• |
| SS | • | • | • | •• | • | • | • | • | •• |
| CA | • | • | • | •• | • | • | • | • | •• |

Figure 2. WISC-V Primary and Secondary Subtest configuration for CFA models with 2–4 factors. VC = Vocabulary; IN = Information; CO = Comprehension; BD = Block Design; VP = Visual Puzzles; MR = Matrix Reasoning; FW = Figure Weights; PC = Picture Concepts; AR = Arithmetic; DS = Digit Span; PS = Picture Span; LN = Letter-Number Sequencing; CD = Coding; SS = Symbol Search; CA = Cancellation. All models include a higher-order general factor except for the bifactor models.

estimates (ω_H and ω_{HS}) may be obtained from CFA bifactor solutions or decomposed variance estimates from higher-order models and were produced using the *Omega* program (Watkins, 2013), which is based on the tutorial by Brunner et al. (2012) and the work of Zinbarg, Revelle, Yovel, and Li (2005) and Zinbarg, Yovel, Revelle, and McDonald (2006). Omega coefficients should at a minimum exceed .50, but .75 is preferred (Reise, 2012; Reise, Bonifay, & Haviland, 2013).

Results

16 WISC-V Primary and Secondary Subtests

Results from CFAs for the 16 WISC-V primary and secondary subtests are presented in Table 1. All five of the higher-order models that included five first-order factors (including the final WISC-V model presented in the WISC-V *Technical and Interpretative Manual*) resulted in inadmissible solutions (i.e., negative

variance estimates for the FR factor) potentially caused by misspecification of the models. Imposing an equality constraint of zero on the FR variance estimate allowed the models to converge properly, but this “only masks the underlying problem” (Hair, Anderson, Tatham, & Black, 1998, p. 610) indicating that these models “should not be trusted” (Kline, 2016, p. 237). Accordingly, neither fit indices nor loadings for these models are reported in Table 1.

A bifactor model that included five first-order factors produced an admissible solution and fit the standardization data well. However, this bifactor model produced results where the Matrix Reasoning, Figure Weights, and Picture Concepts subtests did not have statistically significant loadings on the FR group factor. All five models that included four first-order factors provided good fit to these data. No single four-factor model was superior in terms of $\Delta CFI > .01$ and $\Delta RMSEA > .015$, but the AIC value was lowest for the bifactor version where the FR and VS dimensions ($r = .91$) collapsed into a single (PR) factor (see Figure 5).

| Subtest | Model 5a | | | | | Model 5a Bifactor | | | | | Model 5b | | | | | Model 5c | | | | | Model 5d | | | | | Model 5e | | | | |
|---------|----------|----|----|----|----|-------------------|----|----|----|----|----------|----|----|----|----|----------|----|----|----|----|----------|----|----|----|----|----------|----|----|----|----|
| | F1 | F2 | F3 | F4 | F5 | g | F1 | F2 | F3 | F4 | F5 | F1 | F2 | F3 | F4 | F5 | F1 | F2 | F3 | F4 | F5 | F1 | F2 | F3 | F4 | F5 | F1 | F2 | F3 | F4 |
| SI | • | | | | | • | • | | | | • | | | | | • | | | | | • | | | | | • | | | | |
| VC | • | | | | | • | • | | | | • | | | | | • | | | | | • | | | | | • | | | | |
| IN | • | | | | | • | • | | | | • | | | | | • | | | | | • | | | | | • | | | | |
| CO | • | | | | | • | • | | | | • | | | | | • | | | | | • | | | | | • | | | | |
| BD | | • | | | | • | | • | | | • | | | | | • | | | | | • | | | | | • | | | | |
| VP | | • | | | | • | | • | | | • | | | | | • | | | | | • | | | | | • | | | | |
| MR | | | • | | | • | | | • | | • | | | | | • | | | | | • | | | | | • | | | | |
| FW | | | • | | | • | | | • | | • | | | | | • | | | | | • | | | | | • | | | | |
| PC | | | • | | | • | | | • | | • | | | | | • | | | | | • | | | | | • | | | | |
| AR | | | | • | | • | | | | • | | | | • | • | • | | | | • | | • | | | • | | • | • | • | |
| DS | | | | • | | • | | | | • | | | | • | | • | | | | • | | • | | | • | | | | | |
| PS | | | | • | | • | | | | • | | | | • | | • | | | | • | | • | | | • | | | | | |
| LN | | | | • | | • | | | | • | | | | • | | • | | | | • | | • | | | • | | | | | |
| CD | | | | | • | • | | | | • | | | | • | | • | | | | • | | • | | | • | | | | | |
| SS | | | | | • | • | | | | • | | | | • | | • | | | | • | | • | | | • | | | | | |
| CA | | | | | • | • | | | | • | | | | • | | • | | | | • | | • | | | • | | | | | |

Figure 3. WISC-V Primary and Secondary Subtest configuration for CFA models with 5 factors. VC = Vocabulary; IN = Information; CO = Comprehension; BD = Block Design; VP = Visual Puzzles; MR = Matrix Reasoning; FW = Figure Weights; PC = Picture Concepts; AR = Arithmetic; DS = Digit Span; PS = Picture Span; LN = Letter-Number Sequencing; CD = Coding; SS = Symbol Search; CA = Cancellation. All models include a higher-order general factor except for the bifactor model.

Table 2 presents sources of variance from the 16 WISC-V primary and secondary subtests according to the bifactor model with four group factors. Most subtest variance is associated with the general intelligence dimension and substantially smaller por-

tions of variance are uniquely associated with the four WISC-V group factors. Omega-hierarchical and ω_{HS} coefficients were estimated based on the bifactor results from Table 2 and the ω_H coefficient for general intelligence (.849) was high and sufficient

| Subtest | Model 2 | | Model 3 | | | Model 4 | | | Model 4 Bifactor | | | | Model 5 | | | | Model 5 Bifactor | | | | | | | | |
|---------|---------|----|---------|----|----|---------|----|----|------------------|---|----|----|---------|----|----|----|------------------|----|----|---|----|----|----|----|----|
| | F1 | F2 | F1 | F2 | F3 | F1 | F2 | F3 | F4 | g | F1 | F2 | F3 | F4 | F1 | F2 | F3 | F4 | F5 | g | F1 | F2 | F3 | F4 | F5 |
| SI | • | | • | | | • | | | | • | • | | | | • | | | | | • | • | | | | |
| VC | • | | • | | | • | | | | • | • | | | | • | | | | | • | • | | | | |
| BD | | • | | • | | | • | | | • | | • | | | • | | | | | • | | • | | | |
| VP | | • | | • | | | • | | | • | | • | | | • | | | | | • | | • | | | |
| MR | | • | | • | | | • | | | • | | • | | | • | | | | | • | | • | | | |
| FW | | • | | • | | | • | | | • | | • | | | • | | | | | • | | • | | | |
| DS | | • | | • | | | • | | | • | | • | | | • | | | | | • | | • | | | |
| PS | | • | | • | | | • | | | • | | • | | | • | | | | | • | | • | | | |
| CD | | • | | • | | | • | | | • | | • | | | • | | | | | • | | • | | | |
| SS | | • | | • | | | • | | | • | | • | | | • | | | | | • | | • | | | |

Figure 4. WISC-V Primary Subtest alignment for CFA models. SI = Similarities; VC = Vocabulary; BD = Block Design; VP = Visual Puzzles; MR = Matrix Reasoning; FW = Figure Weights; DS = Digit Span; PS = Picture Span; CD = Coding; SS = Symbol Search. All models include a higher-order general factor except for the bifactor models.

Table 1
CFA Fit Statistics for WISC-V 16 Subtests for the Total Standardization Sample (N = 2,200)

| Model ^a | χ^2 | df | CFI | TLI | SRMR | RMSEA | RMSEA 90% CI | AIC |
|--------------------------------------|---|-----------|-------------|-------------|-------------|-------------|---------------------|--------------|
| 1 (g) | 2538.5 | 104 | .843 | .819 | .065 | .103 | [.100, .107] | 2330.5 |
| 2 (V, P) ^b | 2178.0 | 102 | .866 | .843 | .062 | .096 | [.093, .100] | 1974.0 |
| 3 (V, P, and PS) | 1350.5 | 101 | .920 | .904 | .044 | .075 | [.071, .079] | 1148.5 |
| 4a (VC, PR, WM, and PS) | 577.6 | 100 | .969 | .963 | .032 | .047 | [.043, .050] | 377.6 |
| 4a Bifactor (VC, PR, WM, and PS) | 312.8 | 88 | .986 | .980 | .023 | .034 | [.030, .038] | 136.8 |
| 4b (VC, PR, WM, and PS) | 756.3 | 100 | .958 | .949 | .034 | .055 | [.051, .058] | 556.3 |
| 4c (VC, PR, WM, and PS) | 467.1 | 99 | .976 | .971 | .028 | .041 | [.037, .045] | 269.1 |
| 4d (VC, PR, WM, and PS) | 433.4 | 98 | .978 | .974 | .027 | .039 | [.036, .043] | 237.4 |
| 4e (VC, PR, WM, and PS) ^c | 573.2 | 100 | .970 | .963 | .030 | .046 | [.043, .050] | 373.2 |
| 4e Bifactor ^d | 327.3 | 88 | .985 | .979 | .023 | .035 | [.031, .039] | 151.3 |
| 5a (VC, VS, FR, WM, and PS) | Model specification error (FR produced negative variance), improper model fit statistics not reported | | | | | | | |
| 5b (VC, VS, FR, WM, and PS) | Model specification error (FR produced negative variance), improper model fit statistics not reported | | | | | | | |
| 5c (VC, VS, FR, WM, and PS) | Model specification error (FR produced negative variance), improper model fit statistics not reported | | | | | | | |
| 5d (VC, VS, FR, WM, and PS) | Model specification error (FR produced negative variance), improper model fit statistics not reported | | | | | | | |
| 5e (VC, VS, FR, WM, and PS) | Model specification error (FR produced negative variance), improper model fit statistics not reported | | | | | | | |
| 5a Bifactor (VC, VS, FR, WM, and PS) | 350.9 | 89 | .983 | .977 | .024 | .037 | [.033, .041] | 172.9 |

Note. CFI = comparative fit index; TLI = Tucker-Lewis Index; SRMR = standardized root mean square; RMSEA = root mean square error of approximation; AIC = Akaike’s Information Criterion; g = general intelligence; V = verbal; P = performance; PS = Processing Speed; VC = Verbal Comprehension; PR = Perceptual Reasoning; WM = Working Memory; VS = Visual Spatial; FR = Fluid Reasoning. Bold text illustrates best fitting model.

^a Model numbers and letters correspond to those reported in the WISC-V *Technical and Interpretive Manual* (except 4e and its bifactor variant, which were added for examination) and are higher-order models (unless otherwise specified) when more than one first-order factor was specified (Models 2–5). ^b Disturbance for Factor 2 (Performance) was linearly dependent on other parameters so variance estimate set to zero for model estimation and loss of 1 df. ^c Model 4e placed AR on PR factor alone with no cross-loadings. ^d Negative loading of Arithmetic subtest on PR factor. In Model 5a Bifactor the Matrix Reasoning, Figure Weights, and Picture Concepts subtests *did not* have statistically significant loadings on the FR group factor.

for confident scale interpretation. The ω_{HS} coefficients for the four WISC-V factors (PR, VC, PS, and WM), however, were considerably lower ranging from .109 (PR) to .516 (PS). Thus, the four WISC-V first-order factors, with the possible exception of PS, likely possess too little true score variance to support clinical interpretation (Reise, 2012; Reise et al., 2013).

10 WISC-V Primary Subtests

CFA results for the 10 WISC-V primary subtests are presented in Table 3 and both four- and five-factor models provided good fit

to these data. No single four- or five-factor model was superior in terms of $\Delta CFI > .01$ and $\Delta RMSEA > .015$, but the AIC value was lowest for the bifactor version of the four-factor model where the VS and FR dimensions ($r = .90$) collapsed into a single (PR) factor (see Figure 6). Because of constraining each factor’s loadings to equality because of underidentified latent factors (VC, VS, FR, WM, and PS), the bifactor version of Model 5 (see Figure 4) is mathematically equivalent to higher-order Model 5.

Table 4 presents sources of variance from the 10 WISC-V primary subtests according to the bifactor model with four group

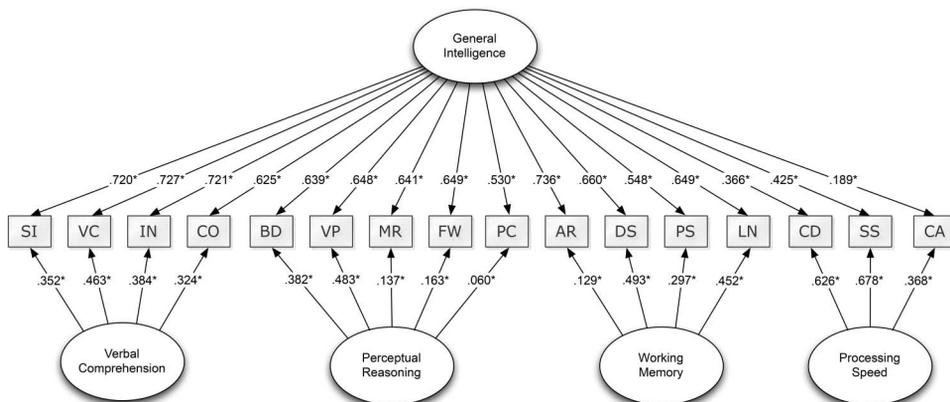


Figure 5. Bifactor measurement model (4a Bifactor), with standardized coefficients, for WISC-V standardization sample (N = 2,200) 16 Subtests. SI = Similarities; VC = Vocabulary; IN = Information; CO = Comprehension; BD = Block Design; VP = Visual Puzzles; MR = Matrix Reasoning; FW = Figure Weights; PC = Picture Concepts; AR = Arithmetic; DS = Digit Span; PS = Picture Span; LN = Letter-Number Sequencing; CD = Coding; SS = Symbol Search; CA = Cancellation. * $p < .05$.

Table 2
Sources of Variance in the WISC-V 16 Subtests for the Total Standardization Sample (N = 2,200) According to a CFA Bifactor Model

| WISC-V subtest | General | | Verbal Comprehension | | Perceptual Reasoning | | Working Memory | | Processing Speed | | h ² | u ² |
|---------------------------------|---------|----------------|----------------------|----------------|----------------------|----------------|----------------|----------------|------------------|----------------|----------------|----------------|
| | b | S ² | b | S ² | b | S ² | b | S ² | b | S ² | | |
| Similarities | .720 | .518 | .352 | .124 | | | | | | | .642 | .358 |
| Vocabulary | .727 | .529 | .463 | .214 | | | | | | | .743 | .257 |
| Information | .721 | .520 | .384 | .147 | | | | | | | .667 | .333 |
| Comprehension | .625 | .391 | .324 | .105 | | | | | | | .496 | .504 |
| Block Design | .639 | .408 | | | .382 | .146 | | | | | .554 | .446 |
| Visual Puzzles | .648 | .420 | | | .483 | .233 | | | | | .653 | .347 |
| Matrix Reasoning | .641 | .411 | | | .137 | .019 | | | | | .430 | .570 |
| Figure Weights | .649 | .421 | | | .163 | .027 | | | | | .448 | .552 |
| Picture Concepts | .530 | .281 | | | .060 | .004 | | | | | .285 | .716 |
| Arithmetic | .736 | .542 | | | | | .129 | .017 | | | .558 | .442 |
| Digit Span | .660 | .436 | | | | | .493 | .243 | | | .679 | .321 |
| Picture Span | .548 | .300 | | | | | .297 | .088 | | | .389 | .611 |
| Letter-Number Sequencing | .649 | .421 | | | | | .452 | .204 | | | .626 | .374 |
| Coding | .366 | .134 | | | | | | | .626 | .392 | .526 | .474 |
| Symbol Search | .425 | .181 | | | | | | | .678 | .460 | .640 | .360 |
| Cancellation | .189 | .036 | | | | | | | .368 | .135 | .171 | .829 |
| Total Variance | | .372 | | .037 | | .027 | | .035 | | .062 | .532 | .468 |
| Common Variance | | .699 | | .069 | | .050 | | .065 | | .116 | | |
| ω _H /ω _{HS} | | .849 | | .200 | | .109 | | .182 | | .516 | | |

Note. b = standardized loading of subtest on factor; S² = variance explained in the subtest; h² = communality; u² = uniqueness; ω_H = omega hierarchical (general factor); ω_{HS} = omega hierarchical subscale (group factors).

factors and, as with the 16 WISC-V primary and secondary subtests, most subtest variance is associated with the general intelligence dimension and smaller portions of variance are uniquely associated with the group factors. Omega-hierarchical and ω_{HS} coefficients were estimated based on the bifactor results from Table 4 and the ω_H coefficient for general intelligence (.817) was high and sufficient for scale interpretation. The ω_{HS} coefficients for the four WISC-V group factors (PR, VC, PS, and WM), however, were considerably lower and ranged from .087 (PR) to .543 (PS). Thus, the four WISC-V group factors, with the possible exception of PS, likely possess too little true score variance to support clinical interpretation (Reise, 2012; Reise et al., 2013).

Discussion

Results from the present independent CFA challenge the WISC-V structure promoted in the WISC-V *Technical and Interpretive Manual* for which standard scores and interpretive guidelines are provided. The present results confirm the outcome of three WISC-V EFA studies, two with the full WISC-V standardization sample (Canivez, Watkins, & Dombrowski, 2015; Dombrowski et al., 2015) and one examining four age groups (Canivez, Dombrowski, & Watkins, 2015), that found a lack of empirical support for five first-order WISC-V factors. Present results found that when modeling five first-order factors and one higher-order

Table 3
CFA Fit Statistics for WISC-V 10 Primary Subtests for the Total Standardization Sample (N = 2,200)

| Model ^a | χ ² | df | CFI | TLI | SRMR | RMSEA | RMSEA 90% CI | AIC |
|---------------------------------|----------------|-----------|-------------|-------------|-------------|-------------|---------------------|-------------|
| 1 (g) | 1296.5 | 35 | .848 | .804 | .071 | .128 | [.122, .134] | 1226.5 |
| 2 (V, P) ^b | 1125.9 | 33 | .868 | .820 | .072 | .123 | [.117, .129] | 1059.9 |
| 3 (V, P, and PS) | 871.1 | 32 | .899 | .858 | .062 | .109 | [.103, .115] | 807.1 |
| 4 (VC, PR, WM, and PS) | 184.9 | 31 | .981 | .973 | .027 | .048 | [.041, .054] | 122.9 |
| 4 Bifactor (VC, PR, WM, and PS) | 126.0 | 28 | .988 | .981 | .024 | .040 | [.033, .047] | 70.0 |
| 5 (VC, VS, FR, WM, and PS) | 134.0 | 30 | .987 | .981 | .025 | .040 | [.033, .047] | 74.0 |

Note. CFI = comparative fit index; TLI = Tucker-Lewis Index; SRMR = standardized root mean square; RMSEA = root mean square error of approximation; AIC = Akaike's Information Criterion; g = general intelligence; V = verbal; P = performance; PS = Processing Speed; VC = Verbal Comprehension; PR = Perceptual Reasoning; WM = Working Memory; VS = Visual Spatial; FR = Fluid Reasoning. VC, WM, and PS subtest loadings were constrained to equality to identify the bifactor version of Model 4 because of under-identified latent factors (VC, WM, and PS). Because of constraining each factor's loadings to equality because of under-identified latent factors (VC, VS, FR, WM, and PS), bifactor version of Model 5 is equivalent to higher-order Model 5. Bold text illustrates best fitting model.

^a Model numbers correspond to those reported in the WISC-V *Technical and Interpretive Manual* and are higher-order models (unless otherwise specified) when more than one first-order factor was specified. ^b Factor 1 (Verbal) and the higher-order factor (g) were linearly dependent on other parameters so variance estimate set to zero for model estimation and loss of 1 df.

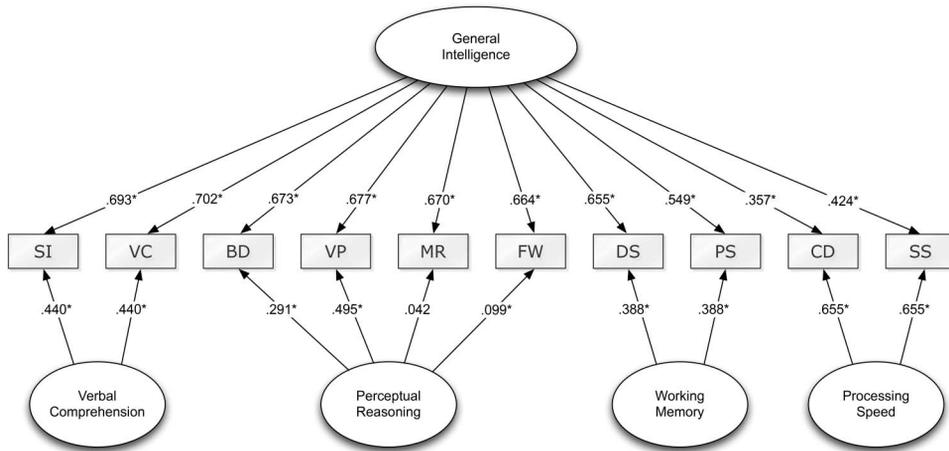


Figure 6. Bifactor measurement model (4 Bifactor), with standardized coefficients, for WISC-V standardization sample ($N = 2,200$) 10 Primary Subtests. SI = Similarities; VC = Vocabulary; BD = Block Design; VP = Visual Puzzles; MR = Matrix Reasoning; FW = Figure Weights; DS = Digit Span; PS = Picture Span; CD = Coding; SS = Symbol Search. * $p < .05$.

factor with all 16 primary and secondary subtests as promoted by the publisher, inadmissible results were produced because of negative variance estimates for the FR factor. A bifactor representation of the WISC-V with general intelligence (g) and five group factors produced admissible results but Matrix Reasoning, Figure Weights, and Picture Concepts subtests did not have statistically significant loadings on the FR group factor, thereby questioning its viability. All models that included four group factors (higher-order and bifactor) were statistically similar but the bifactor model produced the smallest AIC estimate and was selected as the best representation of the WISC-V structure with the 16 primary and secondary subtests. Similar results were also observed in the CFA of only the 10 primary WISC-V subtests.

Difficulty in disentangling g and Gf factors is not unique to the WISC-V. An early example was provided by Gustafsson (1984)

who found identical fluid and general factors among a sample of Swedish students. Other studies using different samples of indicators and participants produced similar results (Chang, Paulson, Finch, McIntosh, & Rothlisberg, 2014; DiStefano & Dombrowski, 2006; Reynolds, Keith, Flanagan, & Alfonso, 2013; Undheim & Gustafsson, 1987). Prior Wechsler scales did not explicitly include a fluid intelligence factor and attempts to identify that factor in recent Wechsler scales often resulted in unity or near unity of Gf - g loadings (Keith, Fine, Taub, Reynolds, & Kranzler, 2006; Watkins & Beaujean, 2014; Weiss et al., 2013a, 2013b).

Carroll's (2003) review of the evidence led him to believe that g and Gf are not identical but "factor Gf is inherently difficult to measure reliably independently of its dependence on g " (p. 14). Vernon (1965) did not incorporate a Gf factor in his model of intelligence because "most of the common variance of

Table 4
Sources of Variance in the WISC-V 10 Primary Subtests for the Total Standardization Sample ($N = 2,200$) According to a CFA Bifactor Model

| Subtest | General | | Verbal Comprehension | | Perceptual Reasoning | | Working Memory | | Processing Speed | | h^2 | u^2 |
|------------------------|---------|-------|----------------------|-------|----------------------|-------|----------------|-------|------------------|-------|-------|-------|
| | b | S^2 | b | S^2 | b | S^2 | b | S^2 | b | S^2 | | |
| Similarities | .693 | .480 | .440 | .194 | | | | | | | .674 | .326 |
| Vocabulary | .702 | .493 | .440 | .194 | | | | | | | .686 | .314 |
| Block Design | .673 | .453 | | | .291 | .085 | | | | | .538 | .462 |
| Visual Puzzles | .677 | .458 | | | .495 | .245 | | | | | .703 | .297 |
| Matrix Reasoning | .670 | .449 | | | .042 | .002 | | | | | .451 | .549 |
| Figure Weights | .664 | .441 | | | .099 | .010 | | | | | .451 | .549 |
| Digit Span | .655 | .429 | | | | | .388 | .151 | | | .580 | .420 |
| Picture Span | .549 | .301 | | | | | .388 | .151 | | | .452 | .548 |
| Coding | .357 | .127 | | | | | | | .655 | .429 | .556 | .444 |
| Symbol Search | .424 | .180 | | | | | | | .655 | .429 | .609 | .391 |
| Total Variance | | .381 | | .039 | | .034 | | .030 | | .086 | .570 | .430 |
| Common Variance | | .669 | | .068 | | .060 | | .053 | | .151 | | |
| ω_H/ω_{HS} | | .817 | | .230 | | .087 | | .199 | | .543 | | |

Note. b = standardized loading of subtest on factor; S^2 = variance explained in the subtest; h^2 = communality; u^2 = uniqueness; ω_h = omega hierarchical (general factor); ω_s = omega subscale (group factors).

reasoning tests is apt to be absorbed into *g*" (p. 725). Likewise, Johnson and Bouchard's (2005) VPR extension of Vernon's theory did not include a fluid reasoning factor. Other researchers believe that a *Gf* factor may not emerge unless its indicators represent diverse content areas (Gustafsson & Wolff, 2015) or unless participants have had differential learning opportunities (Kvist & Gustafsson, 2008). Regardless of the theoretical implications, this study found that there was insufficient covariation among the purported FR subtests to form an independent FR group factor.

Bifactor model-based reliability estimates indicated that while the broad *g* factor would allow individual interpretation (16 subtest $\omega_H = .849$, 10 subtest $\omega_H = .817$) the ω_{HS} estimates for the four WISC-V group factors were generally low (see Tables 2 and 4), and extremely limited for measuring unique constructs (Brunner et al., 2012; Reise, 2012). The ω_{HS} estimates for the four WISC-V group factors were not high enough for individual interpretation (except perhaps for PS). For comparison purposes, standardized path coefficients from the WISC-V higher-order model (Figure 5.2, Wechsler, 2014b) of the 10 primary subtests that are used to provide factor index scores were used to decompose subtest variance among the higher- and lower-order factors (absent from the WISC-V *Technical and Interpretive Manual*) and are presented along with ω_H and ω_{HS} estimates in Table 5. Results were nearly identical to those from the present bifactor model and illustrated strong measurement of the latent general intelligence factor but poor unique measurement of VC, VS, and WM and extremely poor unique measurement of FR.

These results are not unique to the WISC-V and have been observed in both EFA and CFA studies of the WISC-IV (Bodin et al., 2009; Canivez, 2014; Keith, 2005; Watkins, 2006, 2010; Watkins, Wilson, Kotz, Carbone, & Babula, 2006) and with other versions of Wechsler scales (Canivez & Watkins, 2010a, 2010b; Golay & Lecerf, 2011; Golay et al., 2013; Gignac, 2005, 2006; McGill & Canivez, 2016; Watkins & Beaujean, 2014; Watkins et al., 2013). Further, these results are also not unique among Wechsler scales as similar results were also observed with the

DAS-II (Canivez & McGill, 2016), SB5 (Canivez, 2008), WASI and WRIT (Canivez et al., 2009), RIAS (Dombrowski, Watkins, & Brogan, 2009; Nelson & Canivez, 2012; Nelson et al., 2007), CAS (Canivez, 2011), WJ III (Dombrowski, 2013, 2014a, 2014b; Dombrowski & Watkins, 2013; Strickland, Watkins, & Caterino, 2015), and the WJ IV Cognitive (Dombrowski, McGill, & Canivez, in press).

Some researchers have questioned the theoretical appropriateness of bifactor models of intelligence, stating that "we believe that higher-order models are theoretically more defensible, more consistent with relevant intelligence theory (e.g., Jensen, 1998), than are less constrained hierarchical [bifactor] models (Reynolds & Keith, 2013, p. 66). This conclusion has been contested by other researchers. For example, Gignac (2006, 2008) contended that the most substantial factor of a battery of tests (i.e., *g*) should be directly modeled whereas its full mediation in the higher-order model demands an explicit theoretical justification. That is, a rationale for why general intelligence should directly influence group factors but not subtests. Subtest scores reflect variation on both a general and more specific group factor. The effect is that the subtest scores may appear reliable, but the reliability is primarily a function of the general factor, not the specific group factor. Other researchers maintained that a bifactor model better represents the conceptualizations of intelligence expressed by Spearman and Carroll than the higher-order model (Beaujean, 2015; Brunner et al., 2012; Frisby & Beaujean, 2015; Gignac, 2006, 2008; Gignac & Watkins, 2013; Gustafsson & Balke, 1993). In an extended discussion, Beaujean (2015) reported that Spearman's conception of general intelligence was of a factor "that was directly involved in all cognitive performances, not indirectly involved through, or mediated by, other factors" (Spearman, 1927, p. 130). Beaujean (2015) also noted that "Carroll was explicit in noting that a bi-factor model best represents his theory" (p. 130).

It has also been suggested that bifactor models might benefit from statistical bias when compared to higher-order models by better accounting for unmodeled complexity (Murray & Johnson, 2013). However, subsequent Monte Carlo simulations found that

Table 5
Decomposed Sources of Variance in the WISC-V 10 Primary Subtests for the Total Standardization Sample (N = 2,200) According to the Higher-Order Model (Figure 5.2, Wechsler, 2014b)

| Subtest | General | | Verbal Comprehension | | Visual Spatial | | Fluid Reasoning | | Working Memory | | Processing Speed | | h^2 | u^2 |
|------------------------|----------|-------|----------------------|-------|----------------|-------|-----------------|-------|----------------|-------|------------------|-------|-------|-------|
| | <i>b</i> | S^2 | <i>b</i> | S^2 | <i>b</i> | S^2 | <i>b</i> | S^2 | <i>b</i> | S^2 | <i>b</i> | S^2 | | |
| Similarities | .689 | .474 | .442 | .196 | | | | | | | | | .670 | .330 |
| Vocabulary | .697 | .486 | .452 | .204 | | | | | | | | | .690 | .310 |
| Block Design | .684 | .468 | | | .335 | .112 | | | | | | | .580 | .420 |
| Visual Puzzles | .702 | .493 | | | .342 | .117 | | | | | | | .610 | .390 |
| Matrix Reasoning | .673 | .453 | | | | | .082 | .007 | | | | | .460 | .540 |
| Figure Weights | .673 | .453 | | | | | .130 | .017 | | | | | .470 | .530 |
| Digit Span | .647 | .419 | | | | | | | .437 | .191 | | | .610 | .390 |
| Picture Span | .540 | .291 | | | | | | | .359 | .129 | | | .420 | .580 |
| Coding | .357 | .127 | | | | | | | | | .602 | .363 | .490 | .510 |
| Symbol Search | .423 | .179 | | | | | | | | | .715 | .511 | .690 | .310 |
| Total Variance | | .348 | | .040 | | .023 | | .002 | | .032 | | .087 | | |
| Common Variance | | .676 | | .070 | | .040 | | .004 | | .056 | | .154 | | |
| ω_H/ω_{HS} | | .823 | | .238 | | .144 | | .015 | | .210 | | .548 | | |

Note. *b* = standardized loading of subtest on factor; S^2 = variance explained in the subtest; h^2 = communality; u^2 = uniqueness; ω_h = omega hierarchical (general factor); ω_s = omega subscale (group factors).

the bifactor model “did not generally produce a better fit when the true underlying structure was not a bi-factor one” (Morgan, Hodge, Wells, & Watkins, 2015, p. 15). In any case, Murray and Johnson concluded that the “bifactor model factor scores should be preferred” (Murray & Johnson, 2013, p. 420) when there is an attempt to estimate or account for domain-specific abilities; something critical in evaluation of the construct validity of the WISC-V because of publisher claims of what factor index scores measure as well as the numerous comparisons of factor index scores and inferences made from such comparisons. Researchers and clinicians must know how well WISC-V group factors (domain-specific) perform independent of the general intelligence (*g*) factor (F. F. Chen, Hayes, Carver, Laurenceau, & Zhang, 2012; F. F. Chen, West, & Sousa, 2006). Reise, Moore, and Haviland (2010) concluded that a bifactor model, which contains a general factor but permits multidimensionality, is better than the higher-order model for determining the relative contribution of group factors independent of the general factor (i.e., general intelligence). As can be seen in Table 5, however, decomposed variance estimates and model-based reliability estimates based on the publisher presented higher-order model with five group factors produced very similar results to the bifactor models generated in the present study.

Limitations

While the present study examined CFA for the full WISC-V standardization sample it is possible that different age groups within the WISC-V standardization sample might produce different results. Consequently, CFA with different age groups should be conducted to examine structural invariance across age. Further, these data pertain to the standardization sample and may not generalize to other populations such as different clinical groups or independent samples of nonclinical groups, participants of different races/ethnicities, or language minorities. While structural invariance across gender has been reported (H.-L. Chen et al., 2015), bifactor models were not examined so invariance of a bifactor model should also be examined across gender. Finally, the results of the present study only pertain to the latent factor structure and do not fully test the validity of the WISC-V, which also involves relations with external criteria. Other methods such as incremental predictive validity (Canivez, 2013a; Canivez, Watkins, James, Good, & James, 2014; Glutting, Watkins, Konold, & McDermott, 2006) to determine if reliable achievement variance is incrementally accounted for by the factor index scores beyond that accounted for by the FSIQ (or through latent factor scores; see Kranzler, Benson, & Floyd, 2015) and diagnostic utility (see Canivez, 2013b) studies should also be examined. However, given the small portions of true score variance uniquely contributed by the group factors it is difficult to imagine that they will provide substantial value.

Conclusion

Based on the present results as well as WISC-V EFA studies (Canivez, Dombrowski, & Watkins, 2015; Canivez, Watkins, & Dombrowski, 2015; Dombrowski et al., 2015), the WISC-V as presented in the WISC-V *Technical and Interpretive Manual* appears to be overfactored. The FR factor is implausible given

the negligible amount of unique true score variance in either the bifactor or higher-order models. Further, EFA revealed that the fifth factor was defined by only one subtest in the total WISC-V standardization sample (Canivez, Watkins, & Dombrowski, 2015; Dombrowski et al., 2015) and with four age subgroups (Canivez, Dombrowski, & Watkins, 2015). Thus, it appears the attempt to divide the Perceptual Reasoning factor into separate Visual Spatial and Fluid Reasoning factors was unsuccessful and generating standard scores and comparisons for FR potentially misleading. If FR does not have unique contribution then the publisher should provide normative scores for four (VC, PR, WM, and PS) rather than five first-order factors.

Based on the present results and replication of previous findings, primary interpretive emphasis should be placed on the FSIQ. If going beyond the FSIQ and interpreting factor index scores then clinicians must exercise caution to guard against misinterpretation or overinterpretation of scores because of the limitations of the group factors. Further, confidence intervals for the observed factor index scores are narrow because of the strong influence of general intelligence (except for PS). Confidence intervals based on unique factor true score variance would be considerably larger because of substantially less unique true score variance. Those using the WISC-V must be mindful of how much unique variance is conveyed by the different scores provided, which is unavailable in the WISC-V *Technical and Interpretive Manual*, so this and other studies (Canivez, Dombrowski, & Watkins, 2015; Canivez, Watkins, & Dombrowski, 2015; Dombrowski et al., 2015) are critical for proper interpretation of the WISC-V.

References

- Alfonso, V. C., Oakland, T. D., LaRocca, R., & Spanakos, A. (2000). The course on individual cognitive assessment. *School Psychology Review*, 29, 52–64.
- Alfonso, V. C., & Pratt, S. I. (1997). Issues and suggestions for training professionals in assessing intelligence. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 326–347). New York, NY: Guilford Press.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher-order models. *Journal of Intelligence*, 3, 121–136. <http://dx.doi.org/10.3390/jintelligence3040121>
- Belter, R. W., & Piotrowski, C. (2001). Current status of doctoral-level training in psychological testing. *Journal of Clinical Psychology*, 57, 717–726. <http://dx.doi.org/10.1002/jclp.1044>
- Bentler, P. M., & Wu, E. J. C. (2012). *EQS for Windows*. Encino, CA: Multivariate Software, Inc.
- 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>
- 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-Fifth Edition 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. In

- R. A. Spies, J. F. Carlson, & K. F. Geisinger (Eds.), *The eighteenth mental measurements yearbook* (4th ed., 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. (2013a). Incremental criterion validity of WAIS-IV factor index scores: Relationships with WIAT-II and WIAT-III subtest and composite scores. *Psychological Assessment*, *25*, 484–495. <http://dx.doi.org/10.1037/a0032092>
- Canivez, G. L. (2013b). Psychometric versus actuarial interpretation of intelligence and related aptitude batteries. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwane (Eds.), *The Oxford handbook of child psychological assessments* (pp. 84–112). New York, NY: Oxford University Press.
- Canivez, G. L. (2014a). Review of the Wechsler Preschool and Primary Scale of Intelligence. In J. F. Carlson, K. F. Geisinger, & J. L. Jonson (Eds.), *The nineteenth mental measurements yearbook* (4th ed., pp. 732–737). Lincoln, NE: Buros Center for Testing.
- 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. (2016). Bifactor modeling in construct validation of multifaceted 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., Dombrowski, S. C., & Watkins, M. W. (2015, August). *Hierarchical exploratory factor analyses of the WISC-V in four standardization sample age groups*. Paper presented at the 2015 Annual Convention of the American Psychological Association, Toronto, Ontario, Canada.
- Canivez, G. L., Konold, T. R., Collins, J. M., & Wilson, G. (2009). Construct validity of the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test: Convergent and structural validity. *School Psychology Quarterly*, *24*, 252–265. <http://dx.doi.org/10.1037/a0018030>
- Canivez, G. L., & Kush, J. C. (2013). WAIS-IV and WISC-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., & McGill, R. J. (2016, January 25). Factor structure of the Differential Ability Scales–2nd ed.: Exploratory and hierarchical factor analyses with the core subtests. *Psychological Assessment*. Advance online publication. <http://dx.doi.org/10.1037/pas0000279>
- 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. (2016). Review of the Wechsler Intelligence Scale for Children–Fifth Edition: Critique, commentary, and independent analyses. In A. S. Kaufman, S. E. Raiford, & D. L. Coalson (Eds.), *Intelligent testing with the WISC-V* (pp. 683–702). Hoboken, NJ: Wiley.
- Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2015, November 16). Factor structure of the Wechsler Intelligence Scale for Children-Fifth Edition: Exploratory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment*. Advance online publication. <http://dx.doi.org/10.1037/pas0000238>
- Canivez, G. L., Watkins, M. W., James, T., Good, R., & James, K. (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, United Kingdom: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511571312>
- 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. (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. <http://dx.doi.org/10.1016/B978-008043793-4/50036-2>
- 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>
- Chang, M., Paulson, S. E., Finch, W. H., McIntosh, D. E., & Rothlisberg, B. A. (2014). Joint confirmatory factor analysis of the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition, and the Stanford-Binet Intelligence Scales, Fifth Edition, with a preschool population. *Psychology in the Schools*, *51*, 32–57. <http://dx.doi.org/10.1002/pits>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464–504. <http://dx.doi.org/10.1080/10705510701301834>
- 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. <http://dx.doi.org/10.1111/j.1467-6494.2011.00739.x>
- Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*, 189–225. http://dx.doi.org/10.1207/s15327906mbr4102_5
- Chen, H., Zhang, O., Raiford, S. E., Zhu, J., & Weiss, L. G. (2015). Factor invariance between genders on the Wechsler Intelligence Scale for Children-Fifth Edition. *Personality and Individual Differences*, *86*, 1–5. <http://dx.doi.org/10.1016/j.paid.2015.05.020>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5
- 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., 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. <http://dx.doi.org/10.1016/j.intell.2015.10.009>

- Dombrowski, S. C., McGill, R. J., & Canivez, G. L. (in press). Exploratory and hierarchical factor analysis of the WJ-IV Cognitive at school age. *Psychological Assessment*.
- 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>
- Flanagan, D. P., & Kaufman, A. S. (2009). *Essentials of WISC-IV assessment* (2nd ed.). Hoboken, NJ: Wiley.
- Frisby, C. L., & Beaujean, A. A. (2015). Testing Spearman's hypotheses using a bi-factor model with WAIS-IV/WMS-IV standardization data. *Intelligence*, 51, 79–97. <http://dx.doi.org/10.1016/j.intell.2015.04.007>
- Georgas, J., van de Vijver, F. J. R., Weiss, L. G., & Saklofske, D. H. (2003). A cross-cultural analysis of the WISC-III. In J. Georgas, L. G. Weiss, & F. J. R. van de Vijver (Eds.), *Culture and children's intelligence* (pp. 277–313). San Diego, CA: Academic Press. <http://dx.doi.org/10.1016/B978-012280055-9/50021-7>
- 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. (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>
- Glutting, J. J., Watkins, M. W., Konold, T. R., & McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC-IV in estimating reading and math achievement on the WIAI-II. *The Journal of Special Education*, 40, 103–114. <http://dx.doi.org/10.1177/00224669060400020101>
- Goh, D. S., Teslow, C. J., & Fuller, G. B. (1981). The practice of psychological assessment among school psychologists. *Professional Psychology*, 12, 696–706. <http://dx.doi.org/10.1037/0735-7028.12.6.696>
- Golay, P., & Lecerf, T. (2011). Orthogonal higher order structure and confirmatory factor analysis of the French Wechsler Adult Intelligence Scale (WAIS-III). *Psychological Assessment*, 23, 143–152. <http://dx.doi.org/10.1037/a0021230>
- Golay, P., Reverte, I., Rossier, J., Favez, N., & Lecerf, T. (2013). Further insights on the French WISC-IV factor structure through Bayesian structural equation modeling. *Psychological Assessment*, 25, 496–508. <http://dx.doi.org/10.1037/a0030676>
- Grégoire, J., Georgas, J., Saklofske, D. H., van de Vijver, F., Wierzbicki, C., Weiss, L. G., & Zhu, J. (2008). Cultural issues in clinical use of the WISC-IV. In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical assessment and intervention* (2nd ed., pp. 517–544). San Diego, CA: Academic Press.
- Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 179–203. [http://dx.doi.org/10.1016/0160-2896\(84\)90008-4](http://dx.doi.org/10.1016/0160-2896(84)90008-4)
- Gustafsson, J.-E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28, 407–434. http://dx.doi.org/10.1207/s15327906mbr2804_2
- Gustafsson, J.-E., & Wolff, U. (2015). Measuring fluid intelligence at age four. *Intelligence*, 50, 175–185.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2, 41–54. <http://dx.doi.org/10.1007/BF02287965>
- 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* (Rev. ed., pp. 197–232). Itasca, IL: Riverside.
- Horn, J. L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 41–68). New York, NY: Guilford Press.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57, 253–270. <http://dx.doi.org/10.1037/h0023816>
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <http://dx.doi.org/10.1080/1070519990540118>
- Hutton, J. B., Dubes, R., & Muir, S. (1992). Assessment practices of school psychologists: Ten years later. *School Psychology Review*, 21, 271–284.
- 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.
- Johnson, W., & Bouchard, T. J., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33, 393–416. <http://dx.doi.org/10.1016/j.intell.2004.12.002>
- Kaufman, A. S., & Lichtenberger, E. O. (2000). *Essentials of WISC-III and WPPSI-R assessment*. New York, NY: Wiley.
- Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 581–614). New York, NY: Guilford Press.
- Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler Intelligence Scale for Children-Fourth Edition: What does it measure? *School Psychology Review*, 35, 108–127.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford Press.
- Kranzler, J. H., Benson, N., & Floyd, R. G. (2015). Using estimated factor scores from a bifactor analysis to examine the unique effects of the latent variables measured by the WAIS-IV on academic achievement. *Psychological Assessment*, 27, 1402–1416. <http://dx.doi.org/10.1037/pas0000119>
- Kvist, A. V., & Gustafsson, J.-E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's investment theory. *Intelligence*, 36, 422–436. <http://dx.doi.org/10.1016/j.intell.2007.08.004>
- Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, 4, 192–211. <http://dx.doi.org/10.1037/1082-989X.4.2.192>
- Liu, J., & Lynn, R. (2011). Factor structure and sex differences on the Wechsler Preschool and Primary Scale of Intelligence in China, Japan and United States. *Personality and Individual Differences*, 50, 1222–1226. <http://dx.doi.org/10.1016/j.paid.2011.02.013>

- McDonald, R. P. (2010). Structural models and the art of approximation. *Perspectives on Psychological Science, 5*, 675–686. <http://dx.doi.org/10.1177/1745691610388766>
- McGill, R. J., & Canivez, G. L. (2016, January). Orthogonal higher order structure of the WISC-IV Spanish using hierarchical exploratory factor analytic procedures. *Journal of Psychoeducational Assessment*. Advance online publication. <http://dx.doi.org/10.1177/0734282915624293>
- Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are fit indices biased in favor of bi-factor models in cognitive ability research? A comparison of fit in correlated factors, higher-order, and bi-factor models via Monte Carlo simulations. *Journal of Intelligence, 3*, 2–20. <http://dx.doi.org/10.3390/jintelligence3010002>
- Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling, 23*, 116–139. <http://dx.doi.org/10.1080/10705511.2014.961800>
- 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, 41*, 407–422. <http://dx.doi.org/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. <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 with a clinical sample. *Psychological Assessment, 25*, 618–630. <http://dx.doi.org/10.1037/a0032086>
- Pfeiffer, S. I., Reddy, L. A., Kletzel, J. E., Schmelzer, E. R., & Boyer, L. M. (2000). The practitioner's view of IQ testing and profile analysis. *School Psychology Quarterly, 15*, 376–385. <http://dx.doi.org/10.1037/h0088795>
- 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>
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*, 544–559. <http://dx.doi.org/10.1080/00223891.2010.496477>
- Reynolds, M. R., & Keith, T. Z. (2013). Measurement and statistical issues in child assessment research. In D. H. Saklofske, V. L. Schwann, & C. R. Reynolds (Eds.), *Oxford handbook of child psychological assessment* (pp. 48–83). New York, NY: Oxford University Press.
- Reynolds, M. R., Keith, T. Z., Flanagan, D. P., & Alfonso, V. C. (2013). A cross-battery, reference variable, confirmatory factor analytic investigation of the CHC taxonomy. *Journal of School Psychology, 51*, 535–555. <http://dx.doi.org/10.1016/j.jsp.2013.02.003>
- Rodríguez, A., Reise, S. P., & Haviland, M. G. (2015). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*. Advance online publication. <http://dx.doi.org/10.1037/met0000045>
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53–61. <http://dx.doi.org/10.1007/BF02289209>
- Spearman, C. (1927). *The abilities of man*. New York, NY: Cambridge.
- Stinnett, T. A., Havey, J. M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12*, 331–350. <http://dx.doi.org/10.1177/073428299401200403>
- Strickland, T., Watkins, M. W., & Caterino, L. C. (2015). Structure of the Woodcock-Johnson III cognitive tests in a referral sample of elementary school students. *Psychological Assessment, 27*, 689–697. <http://dx.doi.org/10.1037/pas0000052>
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10694-000>
- Undheim, J. O., & Gustafsson, J.-E. (1987). The hierarchical organization of cognitive abilities: Restoring general intelligence through the use of linear structural relations (LISREL). *Multivariate Behavioral Research, 22*, 149–171. http://dx.doi.org/10.1207/s15327906mbr2202_2
- Vernon, P. E. (1965). Ability factors and environmental influences. *American Psychologist, 20*, 723–733. <http://dx.doi.org/10.1037/h0021472>
- Watkins, C. E., Jr., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26*, 54–60. <http://dx.doi.org/10.1037/0735-7028.26.1.54>
- Watkins, M. W. (2006). Orthogonal higher order structure of the Wechsler Intelligence Scale for Children—Fourth Edition. *Psychological Assessment, 18*, 123–125.
- 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 and 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. <http://dx.doi.org/10.1037/10020-000>
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children* (4th ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (4th ed.). 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* (5th ed.). 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. (2013a). WAIS-IV and clinical validation of the four- and five-factor interpretative approaches. *Journal of Psychoeducational Assessment, 31*, 94–113. <http://dx.doi.org/10.1177/0734282913478030>
- Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013b). WISC-IV and clinical validation of the four- and five-factor interpretative approaches. *Journal of Psychoeducational Assessment, 31*, 114–131. <http://dx.doi.org/10.1177/0734282913478032>

- Yung, Y.-F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*, 113–128. <http://dx.doi.org/10.1007/BF02294531>
- Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α , Revelle's β , and McDonald's ω 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>
- Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's

indicators: A comparison of estimators for ω_h . *Applied Psychological Measurement*, *30*, 121–144. <http://dx.doi.org/10.1177/0146621605278814>

Received September 4, 2015
 Revision received May 28, 2016
 Accepted June 2, 2016 ■

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at Reviewers@apa.org. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

APA now has an online video course that provides guidance in reviewing manuscripts. To learn more about the course and to access the video, visit <http://www.apa.org/pubs/authors/review-manuscript-ce-video.aspx>.