Received: 18 September 2017 | Revised: 31 January 2018 | Accepted: 6 April 2018
DOI: 10.1002/pits.22138

WILEY

RESEARCH ARTICLE

Factor structure of the WISC-V in four standardization age groups: Exploratory and hierarchical factor analyses with the 16 primary and secondary subtests

Gary L. Canivez[1] | Stefan C. Dombrowski[2] | Marley W. Watkins[3]

[1]Eastern Illinois University
[2]Rider University
[3]Baylor University

Correspondence
Gary L. Canivez, Ph.D., Professor of Psychology, Eastern Illinois University, Department of Psychology, 600 Lincoln Avenue, Charleston, IL 61920–3099.
Email: glcanivez@eiu.edu

Preliminary results were presented at the 2015 Annual Convention of the American Psychological Association, Toronto, Ontario, Canada.

Abstract

This study examined the factor structure of the Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) with four standardization sample age groups (6–8, 9–11, 12–14, 15–16 years) using exploratory factor analysis (EFA), multiple factor extraction criteria, and hierarchical EFA not included in the WISC-V *Technical and Interpretation Manual*. Factor extraction criteria suggested that one to four factors might be sufficient despite the publisher-promoted, five-factor solution. Forced extraction of five factors resulted in only one WISC-V subtest obtaining a salient pattern coefficient on the fifth factor in all four groups, rendering it inadequate. Evidence did not support the publisher's desire to split Perceptual Reasoning into separate Visual Spatial and Fluid Reasoning dimensions. Results indicated that most WISC-V subtests were properly associated with the four theoretically oriented first-order factors resembling the WISC-IV, the *g* factor accounted for large portions of total and common variance, and the four first-order group factors accounted for small portions of total and common variance. Results were consistent with EFA of the WISC-V total standardization sample.

KEYWORDS
exploratory factor analysis, factor extraction criteria, Schmid–Leiman higher-order analysis, structural validity, WISC-V

The Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V; Wechsler, 2014a) includes 16 intelligence-related subtests, five first-order factor index scores (Verbal Comprehension [VC], Visual Spatial [VS], Fluid Reasoning [FR], Working Memory [WM], and Processing Speed [PS]), and the hierarchically ordered Full Scale score (FSIQ). The Word Reasoning and Picture Completion subtests of the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003) were removed and three new subtests were added. New subtests include Picture Span (adapted from the Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition [Wechsler, 2012]) to measure visual working memory and Visual Puzzles and Figure Weights (adapted from the Wechsler Adult Intelligence Scale-Fourth Edition [Wechsler, 2008]) to measure visual spatial and fluid reasoning, respectively. Separating the former Perceptual Reasoning factor into separate and distinct Visual Spatial and Fluid Reasoning factors was a major goal in developing and marketing the WISC-V.

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 are used in producing the FSIQ; and three additional "Primary" subtests (Visual Puzzles [VP], Picture Span [PS], and Symbol Search [SS]) that are used in producing the five-factor index scores (two subtests each). There are six "Secondary" subtests (Information [IN], Comprehension [CO], Picture Concepts [PC], Arithmetic [AR], Letter–Number Sequencing [LN], and Cancellation [CN]) that are used for substitution in FSIQ estimation or in estimating newly created (Quantitative Reasoning, Auditory Working Memory, Nonverbal) and previously existing (General Ability, Cognitive Proficiency) Ancillary Index Scores. Like other recent editions of intelligence tests (e.g., WISC-IV, Stanford–Binet Intelligence Scales-Fifth Edition [SB5; Roid, 2003a], Kaufman Assessment Battery for Children-Second Edition [KABC-II; Kaufman & Kaufman, 2004], Reynolds Intellectual Assessment Scales [RIAS; Reynolds & Kamphaus, 2003a], Wide Range Intelligence Test [WRIT; Glutting, Adams, & Sheslow, 2000]), the WISC-V attempted to reflect conceptualizations of intellectual measurement articulated by Spearman (1927), Carroll, Cattell, and Horn (Carroll, 1993, 2003; Cattell & Horn, 1978; Horn, 1991; Horn & Blankson, 2012; Horn & Cattell, 1966), as well as other neuropsychological constructs.

Evidence of WISC-V structural validity reported in the WISC-V *Technical and Interpretive Manual* was based exclusively on confirmatory factor analyses (CFA). A one-factor model served as the baseline and all other models were higher-order models with a general intelligence factor indirectly influencing subtests via full mediation through two through five first-order factors. Table 5.3 in the WISC-V *Technical and Interpretive Manual* illustrates all CFA models tested and Figure 5.10 (reproduced in modified form here as Figure 1) presents the standardized measurement model for the final publisher-preferred, five-factor, higher-order model for WISC-V primary and secondary subtests for the total standardization sample. This model included a higher-order general intelligence dimension with five first-order factors (VC, VS, FR, WM, PS) and the 16 subtest indicators were uniquely associated with one latent first-order factor except for Arithmetic, which was cross-loaded on VC, FR, and WM. This preferred measurement model, however, included a standardized path coefficient of 1.00 between the higher-order general intelligence factor and the FR factor, which indicates that FR may be empirically redundant. This final model was also reported to fit five different age groupings (6–7, 8–9, 10–11, 12–13, 14–16) equally well (Wechsler, 2014b) and a subsequent study by Chen, Zhang, Raiford, Zhu, and Weiss (2015) showed factorial invariance of this final model across gender.

CFA reported in the WISC-V *Technical and Interpretive Manual* contained numerous notable psychometric concerns (Beaujean, 2016; Canivez & Watkins, 2016; Canivez, Watkins, & Dombrowski, 2016, 2017a). Details regarding CFA methods are lacking, such as the absence of explanation for selecting weighted least squares (WLS) estimation rather than maximum likelihood (ML) estimation. Latent constructs (i.e., factors) have no natural scale of measurement, so specification by the analyst is necessary to achieve model identification. The choice of metric can affect unstandardized parameters and may "yield different conclusions regarding the statistical significance of freely estimated parameters" (Brown, 2015, p. 133). Kline (2011) noted that "use of an estimation method other than ML requires explicit justification" (p. 154). WLS is typically used with data that are categorical or nonnormally distributed and may not produce chi-values nor approximate fit indices equivalent to those produced by ML estimation (Yuan & Chan, 2005); neither of which pertains to WISC-V subtest scores (Chen et al., 2015). Thus, the use of WLS is perplexing, and a significant departure from the typical use of ML estimation in CFA of intelligence tests. Further, Beaujean (2016) replicated the WISC-V CFA results reported in Wechsler (2014b), deducing that an effects-coding method (Little, Slegers, & Card, 2006) was probably used. Additionally, Beaujean demonstrated that the effects-coding method was modified and caused degrees of freedom to be understated, which has consequences for fit statistics that rely on degrees of freedom for their computation.

The complex CFA model adopted by the publisher (as a result of including Arithmetic subtest cross-loadings) is also problematic because it abandons the parsimony of simple structure (Thurstone, 1947). Further, the publisher's

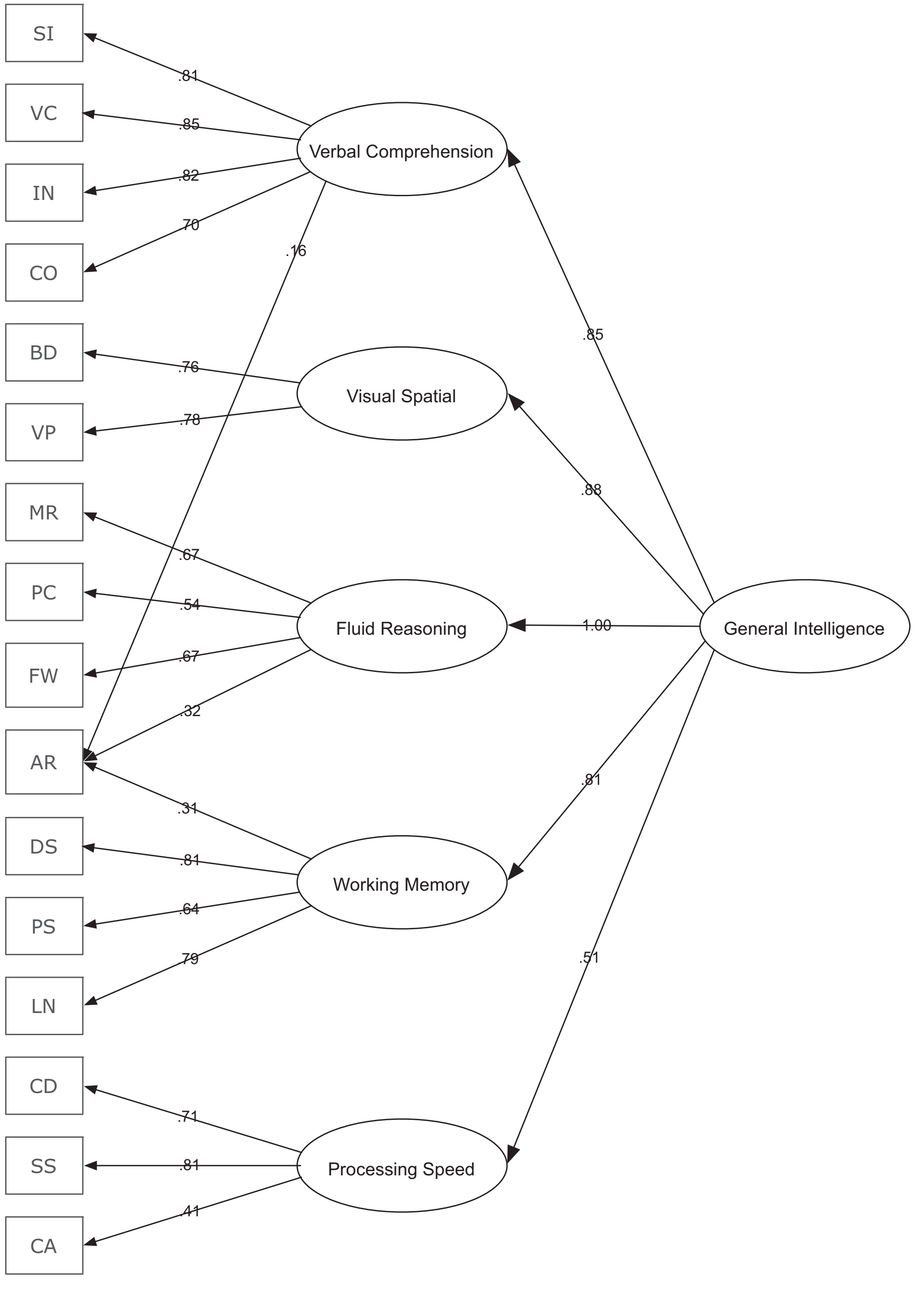

FIGURE 1 Higher-order measurement model with standardized coefficients (adapted from Figure 5.1 [Wechsler, 2014b]), for WISC-V standardization sample (N = 2,200) 16 Subtests. SI = Similarities, VC = Vocabulary, IN = Information, CO = Comprehension, BD = Block Design, VP = Visual Puzzles, MR = Matrix 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 US and/or other countries, of Pearson Education, Inc. or its affiliates(s).

preferred model produced a standardized path coefficient of 1.00 between the latent general intelligence (mislabeled in Figure 5.1 [Wechsler, 2014b] as “Full Scale”) factor and the Fluid Reasoning factor; indicating *g* and FR were empirically redundant (Le, Schmidt, Harter, & Lauver, 2010). This constitutes a major threat to discriminant validity and indicates that the WISC-V has likely been overfactored (Frazier & Youngstrom, 2007).

Another issue concerning the WISC-V *Technical and Interpretive Manual* was the acknowledgment of the sensitivity of the chi-square test to trivial differences with large samples, but the subsequent use of chi-square difference tests of nested models to identify the preferred five-factor model (Wechsler, 2014b). The same sensitivity to large samples is true for chi-square difference tests (Millsap, 2007), suggesting that the model differences reported in the WISC-V *Technical and Interpretive Manual* might be statistically significant yet trivial. For example, Table 5.4 in Wechsler (2014b, p. 82) reveals that the difference between models 4a and 5a was statistically significant but those two models exhibited identical comparative fit index (CFI) and root mean square error of approximation (RMSEA) values. Likewise, the preferred five-factor higher-order model was significantly different from other five-factor models but all exhibited identical CFI and RMSEA values (e.g., .98 and .04, respectively). Cheung and Rensvold (2002) demonstrated, in the context of factorial invariance, that practical differences independent of sample size and model complexity could be identified by ΔCFI > .01; this condition was not met when moving from a four- to a five-factor solution.

Another criticism of WISC-V CFA reported in the WISC-V *Technical and Interpretive Manual* is that there was a failure to test rival bifactor measurement models against the higher-order measurement models. Bifactor models have several benefits over higher-order models (Canivez, 2016; Reise, 2012), have been found to fit data from other Wechsler scales (viz., Canivez, 2014a; Canivez, Watkins, Good, James, & James, 2017b; Gignac & Watkins, 2013; Lecerf & Canivez, 2017; Nelson, Canivez, & Watkins, 2013; Watkins, 2010; Watkins & Beaujean, 2014; Watkins, Canivez, James, James, & Good, 2013), and have been recommended for cognitive tests (Brunner, Nagy, & Wilhelm, 2012; Canivez, 2016; Cucina & Byle, 2017; Cucina & Howardson, 2017; Gignac, 2005, 2006; Morin, Arens, Tran, & Caci, 2016). A higher-order structural model posits general intelligence as a superordinate construct that is fully mediated by the lower-order factors and indirectly influences the subtest indicators. In contrast, the bifactor model hypothesizes general intelligence as a breadth factor with direct influence on subtests in addition to direct influence on subtests by group factors (Canivez, 2016; Gignac, 2008). The bifactor model appears to be more consistent with Spearman's (1927) conceptualization of intelligence and a more conceptually parsimonious explanation than the higher-order model (Canivez, 2016; Cucina & Howardson, 2017; Gignac, 2006). Further, the structure of intelligence described by Carroll (1993) is better represented by the bifactor model (Beaujean, 2015a; Cucina & Howardson, 2017).

Another significant problem is that the publisher did not provide decomposed variance estimates to disclose how much subtest variance is a result of the hierarchical *g* factor and how much is a result of the lower-order group factors. This makes it difficult for clinicians and researchers to judge the adequacy of the group factors (VC, VS, FR, WM, PS) based on how much unique variance the group factors capture when purged of the effects of general intelligence (Reise, Moore, & Haviland, 2010), although this could be computed by hand from the model. As noted by DeVellis (2017), relying on statistical fit alone “may obscure the fact that some statistically significant factors may account for uninterestingly small proportions of variance” (p. 199).

Also missing from the WISC-V *Technical and Interpretive Manual* are model-based reliability estimates (omega). It has long been argued that classical estimates of reliability are biased (Raykov, 1997). Model-based estimates, such as omega-hierarchical (ω_H) and omega-hierarchical subscales (ω_{HS}), have been recommended as superior metrics for determining construct-based reliability (Rodriguez, Reise, & Haviland, 2016; Watkins, 2017). These problems were highlighted in several reviews and critiques of Wechsler scales including the WAIS-IV, WPPSI-IV, and WISC-IV (Canivez, 2010, 2014b; Canivez & Kush, 2013); however, omega estimates are notably absent from the WISC-V *Technical and Interpretive Manual*.

Although Chen et al. (2015) used ML estimation in their WISC-V invariance study, their chosen model replicated the standardized path coefficient of 1.0 from the FSIQ to FR and cross-loading of Arithmetic on three first-order factors. Further, there was no consideration of rival bifactor models nor was there decomposition of subtest variance or estimation of latent factor reliabilities to understand the relative contributions of the higher-order versus first-order factors. Reynolds and Keith (2017) examined WISC-V invariance across standardization sample age groups, but the model examined for invariance was an oblique five-factor model rather than the bifactor or higher-order model, which thus ignored general intelligence and its unmodeled variance.

Reynolds and Keith (2017) also explored numerous post hoc modifications for first-order models with five factors and then for both higher-order and bifactor models with five group factors in an attempt to better understand WISC-V measurement. Whereas such explorations are possible, they may capitalize on chance and it could be argued that such exploratory interest might be better served by using EFA (Carroll, 1995) or exploratory structural equation modeling (Asparouhov & Muthén, 2009). Their final best fitting WISC-V higher-order model was different from the publisher-preferred model in that Arithmetic was given a direct loading from general intelligence and a “cross-loading” on Working Memory, but Reynolds and Keith also added correlated disturbance of Visual Spatial and Fluid Reasoning group factors yet the model still produced a standardized path coefficient of .97 from general intelligence to Fluid Reasoning. Further, decomposed variance estimates of their higher-order model showed that the WISC-V subtests primarily reflected general intelligence variance with small portions of variance unique to the group factors (except for the Processing Speed subtests). Their best WISC-V bifactor model also added a covariance estimate between Visual Spatial and Fluid Reasoning (.62), which appears necessary to salvage five group factors. Watkins, Dombrowski, and Canivez (2017) also tested a similar bifactor model with the Canadian WISC-V (WISC–V^{CDN}), but this bifactor model with five group factors and VS-FR covariance estimate was not superior to the bifactor model with four group factors.

A final criticism is that the WISC-V *Technical and Interpretive Manual* includes explicit preference for CFA over EFA methods rather than taking advantage of each method's unique strengths. EFA and CFA are complementary procedures, so greater confidence in the latent factor structure is achieved when EFA and CFA are in agreement (Gorsuch, 1983). Carroll (1995) and Reise (2012) both noted that EFA procedures are especially useful in suggesting possible models to be tested in CFA, and Carroll (1998) suggested that “CFA should derive its initial hypotheses from EFA results, rather than starting from scratch or from a priori hypotheses…[and] CFA analyses should be done to check my EFA analyses” (p. 8). The deletion of Word Reasoning and Picture Completion subtests; the addition of Visual Puzzles, Figure Weights, and Picture Span subtests; and the inclusion of new or revised items across all WISC-V subtests suggests that relationships among retained and new subtests might result in associations and latent structure unanticipated by *a priori* conceptualizations (Beaujean, 2015b; Strauss, Spreen, & Hunter, 2000).

Intelligence test factor structure research using EFA procedures have consistently produced serious and substantial challenges to the optimistic conclusions from CFA-based latent structures reported in test technical manuals. DiStefano and Dombrowski (2006) and Canivez (2008), using data from the SB5 (Roid, 2003a) standardization sample, obtained markedly different results for the SB5 than CFA results presented in the technical manual (Roid, 2003b) and concluded that the SB5 essentially measured one dimension (*g*). Three studies of the WISC-IV (Wechsler, 2003) and two studies of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV; Wechsler, 2008) using EFA (Bodin, Pardini, Burns, & Stevens, 2009; Canivez & Watkins, 2010a, 2010b; Watkins, 2006; Watkins, Wilson, Kotz, Carbone, & Babula, 2006) indicated that most variance was associated with general intelligence (substantially lesser amounts at the factor level) and suggested that interpretation of both the WISC-IV and WAIS-IV should focus on the global FSIQ score because it accounts for most of the common variance and additional research showing FSIQ superiority in predictive validity with little to no meaningful incremental prediction by the factor index scores (Canivez, 2014a; Canivez, Watkins, James, James, & Good, 2014; Glutting, Watkins, Konold, & McDermott, 2006; Glutting, Youngstrom, Ward, Ward, & Hale, 1997; Nelson et al., 2013). The limited unique variance captured by the first-order factors is likely responsible for the poor incremental predictive validity of the WISC-IV and WAIS-IV factor index scores. EFA studies of other intelligence tests such as RIAS (Reynolds & Kamphaus, 2003a) have also indicated that fundamental measurement is primarily that of general intelligence (Dombrowski, Watkins, & Brogan, 2009; Nelson & Canivez, 2012; Nelson, Canivez, Lindstrom, & Hatt, 2007), which was by design its primary goal (Reynolds & Kamphaus, 2003b). Similar findings were obtained with a joint examination of the Wide Range Intelligence Test (WRIT; Glutting et al., 2000) and Wechsler Abbreviated Scale of Intelligence (WASI; Psychological Corporation, 1999) where most subtest variability was associated with a hierarchical general intelligence dimension and smaller portions of variance were apportioned to the first-order factors; supporting primary interpretations of the FSIQ and general intelligence test (Canivez, Konold, Collins, & Wilson, 2009).

Independent assessment of the WISC-V using EFA with the total standardization sample ($n = 2,200$) was reported by Canivez et al. (2016) and no evidence was found for five factors. The intended separation of Visual Spatial and Fluid Reasoning dimensions was not supported as extracting five factors resulted in the fifth factor including only one subtest (Figure Weights) with a salient factor pattern coefficient, and Picture Concepts failed to saliently load on any factor. Extraction of four factors produced a structure very similar to the WISC-IV with Visual Spatial and Fluid Reasoning collapsing into one Perceptual Reasoning factor. Schmid and Leiman (SL, 1957) orthogonalization found the *g* factor accounted for large portions of total and common variance and provided little evidence for interpretation of the lower-ordered factors. The omega-hierarchical coefficient of the *g* factor was large while the omega-hierarchical subscale coefficients for the four lower-order factors were too low for confident interpretation, except perhaps for the Processing Speed factor. Canivez et al. (2017a) replicated WISC-V EFA results with CFA using maximum likelihood estimation, further challenging results in the WISC-V *Technical and Interpretive Manual*. Whereas these results were consistent with other Wechsler scales (WPPSI-IV, WISC-IV, WAIS-IV), and other tests of intelligence, they were obtained with the entire standardization sample and it is possible that different structures might be observed within different age ranges; therefore, Canivez et al. recommended examination of WISC-V structure with different age groups using similar EFA procedures.

Following that recommendation, the present study investigated the factor structure of the WISC-V with four age groups (6–8, 9–11, 12–14, 15–16 years) from the WISC-V standardization sample using EFA followed by a Schmid–Leiman orthogonalization, the same procedures used by Canivez et al. (2016) when investigating the WISC-V total sample to allow for direct comparison of results. The EFA-based SL orthogonalization procedure produces an approximate bifactor solution that is a reparameterization of the higher-order structure and contains proportionality constraints (Yung, Thissen, & McLeod, 1999), but is the dominant exploratory approach to assessing bifactor structure (Reise, 2012). Also, the present study used identical EFA methods to Canivez et al. (2016), which allows for direct comparison of results of the more homogeneous age groups to the full standardization sample results but does not directly test the factorial invariance of the WISC-V across age/development. The primary research questions included (1) how many WISC-V factors should be extracted and retained in each age subgroup; (2) how are subtests associated with the latent factors; (3) was there evidence for the publisher's claim of five first-order factors; and (4) what proportion of variance was a result of general intelligence versus the first-order group ability factors following a Schmid–Leiman orthogonalization?

1 | METHOD

1.1 | Participants

Participants were members of the WISC-V standardization sample and included a total of 2,200 individuals ranging in age from 6 to 16 years. Demographic characteristics are provided in detail in the WISC-V *Technical and Interpretive Manual* (Wechsler, 2014b). Stratified proportional sampling was used across variables of age, sex, race/ethnicity, parental education level, and geographic region in obtaining the standardization sample. Education level was a proxy for socioeconomic status where accurate information about income is often difficult to obtain. Examination of tables in the *Technical and Interpretive Manual* revealed a close match to the U.S. census across stratification variables.

1.2 | Instrument

The WISC-V is an individual test of general intelligence for children ages 6–16 years and originated with the first WISC (Wechsler, 1949). 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 group factors. WISC-V measurement of intelligence continues to include narrow ability subtests (16), broad group factors (5), and general intelligence.

Organization and subtest administration order of the WISC-V reflect a new four-level organization. The FSIQ is composed of seven primary subtests across the five domains (VC, VS, FR, WM, 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 domain. 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, PSI). No substitutions are allowed for the Primary Index Scales. Complementary subtests are not intelligence subtests and so were not included in the present analyses.

1.3 | Procedure

NCS Pearson denied without rationale the request for WISC-V standardization sample raw data to conduct these (and other) independent analyses. Absent raw data, WISC-V subtest scaled score correlation matrices for each age group ($n = 200$) in the standardization sample were obtained from the WISC-V *Technical and Interpretive Manual Supplement* (Wechsler, 2014c) and combined by averaging correlations through Fisher transformations. Four correlation matrices (16 primary and secondary intelligence subtests) were created to represent four broad age subgroups (ages 6–8 [$n = 600$], 9–11 [$n = 600$], 12–14 [$n = 600$], and 15–16 [$n = 400$] years). The sample size of single age groups ($n = 200$) would be too small for stable results (Goldberg & Velicer, 2006; Mundfrom & Shaw, 2005). In contrast, these four age groups should allow developmental differences to emerge while still providing robust factor recovery.

1.4 | Analyses

Principal axis exploratory factor analyses (Fabrigar, Wegener, MacCallum, & Strahan, 1999) were used to analyze the combined WISC-V standardization sample correlation matrices from the four age groups using SPSS 21 for Macintosh OSX. Principal axis EFA was selected for comparison to Canivez et al. (2016) and because it “frequently outperformed ML in the recovery of relatively weak common factors” (Briggs & MacCallum, 2003, p. 49). Multiple criteria (Gorsuch, 1983) were examined to determine the number of factors to retain and included 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). The scree test is a subjective criterion so the SE_{scree} as programmed by Watkins (2007) was used because it was reportedly the most accurate objective scree method (Nasser, Benson, & Wisenbaker, 2002).

HPA and MAP were included because they are considered more accurate and less likely to overfactor (Frazier & Youngstrom, 2007; Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986), although in the presence of a strong general factor HPA tends to underfactor (Crawford et al., 2010). HPA indicates meaningful factors when eigenvalues from the WISC-V standardization sample data were larger than eigenvalues produced by random data containing the same number of participants and factors. Random data eigenvalues for HPA were produced using the Monte Carlo principal components analysis for the Parallel Analysis computer program (Watkins, 2000) with 100 replications to provide stable eigenvalue estimates. Retained factors were subjected to promax (oblique) rotation ($k = 4$; Gorsuch, 1983). Setting k to 4 produced greater hyperplane count compared to $k = 2$ with the present data. Salient factor pattern coefficients were defined as those ≥.30 (Child, 2006). Factor solutions were examined for interpretability and theoretical plausibility (Fabrigar et al., 1999) with the empirical requirement that each factor should be marked by two or more salient loadings and no salient cross-loadings (Gorsuch, 1983). Subtest *g* loadings (first unrotated factor coefficients) were evaluated based on Kaufman's (1994) criteria (≥.70 = good, .50 –.69 = fair, <.50 = poor).

Cognitive ability subtest scores reflect combinations of both first-order and second-order factor variance and, because of this, Carroll (1993, 1995, 1997, 2003) argued that variance from the higher-order factor must be extracted first to residualize the lower-order factors, leaving them orthogonal to the higher-order factor. The Schmid and Leiman (1957) procedure has been recommended as the statistical method to accomplish this residualization (Carroll, 1993, 1995, 1997, 2003; Carretta & Ree, 2001; Gustafsson & Snow, 1997; McClain, 1996; Ree, Carretta, & Green, 2003; Thompson, 2004). It is a reparameterization of a higher-order model and an approximate bifactor solution (Reise, 2012). Accordingly, first-order factors were orthogonalized by removing all variance associated with the second-order dimension using the Schmid and Leiman (1957) procedure as programmed in the MacOrtho computer program (Watkins, 2004). This transforms "an oblique factor analysis solution containing a hierarchy of higher-order factors into an orthogonal solution which not only preserves the desired interpretation characteristics of the oblique solution, but also discloses the hierarchical structuring of the variables" (Schmid & Leiman, 1957, p. 53).

The Schmid–Leiman (SL) orthogonalization procedure may be constrained by proportionality (Yung et al., 1999) and may be problematic with nonzero cross-loadings (Reise, 2012). Reise also noted two additional and more recent alternative exploratory bifactor methods that do not include proportionality constraints: analytic bifactor (Jennrich & Bentler, 2011) and target bifactor (Reise, Moore, & Maydeu-Olivares, 2011). However, the present application of the SL orthogonalization procedure was selected for direct comparison to WISC-V results obtained by Canivez et al. (2016) with the total WISC-V standardization sample and comparisons to the numerous studies of SL application with Wechsler scales (Canivez & Watkins, 2010a; 2010b; Golay & Lecerf, 2011; Lecerf & Canivez, 2017; Watkins, 2006; Watkins et al., 2017) and with other intelligence tests (Canivez, 2008, 2011; Canivez & McGill, 2016; Canivez et al., 2009; Dombrowski, 2013, 2014a, 2014b; Dombrowski & Watkins, 2013; Dombrowski et al., 2009; Dombrowski, McGill, & Canivez, 2017a, 2017b; Nelson & Canivez, 2012; Nelson et al., 2007; Strickland, Watkins, & Caterino, 2015). For convenience, this method is labeled the SL bifactor (Reise, 2012).

Omega-hierarchical and omega-hierarchical subscale coefficients (Reise, 2012; Rodriguez et al., 2016) were estimated as model-based reliability estimates of the latent factors (Gignac & Watkins, 2013). Chen, Hayes, Carver, Laurenceau, & Zhang (2012) noted that "for multidimensional constructs, the alpha coefficient is complexly determined, and McDonald's (1999) omega-hierarchical (ω_H) provides a better estimate for the composite score and thus should be used" (p. 228). These same problems are inherent with other internal consistency estimates such as split-half or KR-20. Omega-hierarchical (ω_H) is the model-based reliability estimate for the hierarchical general intelligence factor independent of the variance of group factors. Omega-hierarchical subscale (ω_{HS}) is the model-based reliability estimate of a group factor with all other group and general factors removed (Reise, 2012). Omega estimates (ω_H and ω_{HS}) may be obtained from EFA SL bifactor solutions and were produced using the Omega program (Watkins, 2013), which was 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-hierarchical coefficients should at a minimum exceed .50, but .75 would be preferred (Reise, 2012; Reise, Bonifay, & Haviland, 2013).

2 | RESULTS

2.1 | Factor extraction criteria comparisons

Figures A1–A4 (Appendix A in online supplemental materials) show scree plots from HPA for the four age groups. Table 1 summarizes results from the multiple factor extraction criteria (eigenvalues >1, scree test, standard error of scree, HPA, MAP, theory) for determining the number factors to extract and retain. As shown in Table 1, only the publisher recommended/theory justified extraction of five factors. All other criteria across the four age groups mostly recommended extraction of only one to three factors.

2.2 | Five-factor exploratory and hierarchical analyses

It has been suggested that it is better to overextract than underextract (Gorsuch, 1997; Wood, Tataryn, & Gorsuch, 1996) so EFA began with extracting five factors to examine subtest associations based on the publisher's suggested structure and to allow examination of the performance of smaller factors. Tables B1 through B8 (Appendix B in online supplemental materials) show exploratory factor analyses results (odd-numbered Tables B1–B7) and exploratory SL bifactor model results (even-numbered Tables B2–B8) for the four age groups. In each of the four age groups, extraction of five factors produced psychometrically inadequate results as the fifth factor included only one salient factor

TABLE 1 Number of WISC-V factors suggested for extraction across five different criteria by age group

	WISC-V Age Groups				
Extraction Criterion	**6–8**	**9–11**	**12–14**	**15–16**	**6–16**
Eigenvalue >1	3	3	2	3	2
Scree test (visually examined)	2	2	2	2	2
Standard error of scree (SE_{scree})	2	4	3	3	3
Horn's parallel analysis (HPA)	2	2	2	2	2
Minimum average partials (MAP)	1	1	1	2	1
Prior Wechsler structure/theory	4	4	4	4	4
Publisher (theory) proposed	5	5	5	5	5

pattern coefficient (Cancellation [ages 6–8], Arithmetic [ages 9–11], Picture Concepts [ages 12–14 and 15–16]) and factors cannot be defined by only one indicator (see odd-numbered Tables B1–B7 in online supplemental materials). Further, contrary to the publisher's desire to split the Perceptual Reasoning factor into separate Visual Spatial (Block Design, Visual Puzzles) and Fluid Reasoning (Matrix Reasoning, Figure Weights) factors, extraction of five factors still resulted in Block Design, Visual Puzzles, Matrix Reasoning, and Figure Weights having salient factor pattern loadings on the same (Perceptual Reasoning) factor. Exploratory SL bifactor model results (see even-numbered Tables B2–B8 in online supplemental materials) also show the dominance of the general intelligence factor for all subtests except Coding, Symbol Search, and Cancellation (Processing Speed subtests), known to be poor indicators of general intelligence.

2.3 | Four-factor exploratory and hierarchical analyses

2.3.1 | Ages 6–8 first-order EFA: Four-factor extraction

Table 2 shows results of four-factor extraction with promax rotation for the 6- to 8-year-olds. The *g* loadings ranged from .175 (Cancellation) to .746 (Information) and all were within the fair-to-good range (except Coding and Cancellation). Picture Concepts failed to exhibit salient pattern loadings on any group factor. Table 2 shows robust Verbal Comprehension (Similarities, Vocabulary, Information, Comprehension), Working Memory (Digit Span, Picture Span, Letter–Number Sequencing), Perceptual Reasoning (Block Design, Visual Puzzles, Matrix Reasoning, Figure Weights), and Processing Speed (Coding, Symbol Search, Cancellation) factors with theoretically consistent subtest associations. Picture Concepts, a fair indicator of general intelligence, was not adequately associated with any of the four group factors; although its highest pattern coefficient was on the Perceptual Reasoning factor. There were no subtests with salient cross-loadings. The moderate-to-high factor correlations shown in Table 2 (.372 to .710) imply a higher-order or hierarchical structure that required explication (Gorsuch, 1983) and the Schmid–Leiman procedure was applied to better understand variance apportionment among general and group factors.

2.3.2 | Ages 6–8 SL bifactor analyses: Four first-order factors

Results for the Schmid and Leiman orthogonalization of the higher-order factor analysis are shown in Table 3. All subtests were properly associated (higher residual variance) with their theoretically proposed factor after removing *g* variance. The *g* factor accounted for 33.2% of the total variance and 66.4% of the common variance.

The general factor also accounted for between 2.3% (Cancellation) and 49.7% (Digit Span) of individual subtest variability. At the first-order level, VC accounted for an additional 4.6% of the total variance and 9.1% of the common variance, WM accounted for an additional 3.2% of the total variance and 6.5% of the common variance, PR accounted for an additional 3.3% of the total variance and 6.6% of the common variance, and PS accounted for an additional 5.7% of the total variance and 11.4% of the common variance. The general and group factors combined to measure 50.0% of the variance in WISC-V scores, resulting in 50.0% unique variance (combination of specific and error variance).

TABLE 2 Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) exploratory factor analysis: Four oblique factor solution for the standardization sample 6- to 8-year-olds ($n = 600$)

	General[a]	F1: Verbal Comprehension		F2: Working Memory		F3: Perceptual Reasoning		F4: Processing Speed		
WISC-V subtest	*S*	*P*	*S*	*P*	*S*	*P*	*S*	*P*	*S*	h^2
Similarities	.721	**.643**	.759	.169	.620	–.001	.560	–.009	.318	.589
Vocabulary	.703	**.885**	.828	–.086	.530	.055	.577	–.095	.213	.699
Information	.746	**.766**	.806	–.005	.598	.004	.585	.109	.393	.660
Comprehension	.626	**.571**	.660	.150	.541	–.034	.475	.018	.293	.447
Block Design	.611	.026	.491	–.021	.483	**.608**	.675	.145	.406	.472
Visual Puzzles	.667	–.013	.548	.001	.521	**.827**	.795	–.055	.297	.635
Matrix Reasoning	.704	.035	.572	.253	.628	**.487**	.700	.042	.397	.536
Figure Weights	.550	.093	.473	.197	.487	**.365**	.545	–.042	.253	.328
Picture Concepts	.537	.098	.452	.190	.481	.281	.504	.062	.318	.297
Arithmetic	.680	.121	.565	**.629**	.718	–.023	.501	.035	.397	.523
Digit Span	.732	–.012	.576	**.788**	.796	.075	.570	–.064	.374	.639
Picture Span	.507	–.029	.392	**.500**	.540	.103	.413	–.015	.278	.297
Letter–Number Sequencing	.692	.051	.561	**.788**	.766	–.035	.503	–.065	.348	.592
Coding	.368	–.089	.197	.101	.361	–.078	.239	**.721**	.707	.510
Symbol Search	.517	.036	.353	.073	.471	–.005	.380	**.722**	.771	.602
Cancellation	.175	.058	.124	–.279	.083	.161	.189	**.408**	.354	.159
Eigenvalue		6.50		1.44		1.04		.93		
% Variance		37.73		5.94		3.23		2.99		
Promax-based factor correlations		F1: VC		F2: WM		F3: PR		F4: PS		
F1: Verbal Comprehension (VC)		–								
F2: Working Memory (WM)		.710		–						
F3: Perceptual Reasoning (PR)		.701		.674		–				
F4: Processing Speed (PS)		.372		.520		.431				

[a]Factor structure coefficients from first unrotated factor (*g* loadings) are correlations between the subtest and the general factor. *S* = structure coefficient, *P* = pattern coefficient, h^2 = communality. Salient pattern coefficients are shown in bold (pattern coefficient ≥ .30).

TABLE 3 Sources of variance in the Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) for the standardization sample 6- to 8-year-olds (n = 600) according to an exploratory SL bifactor model with four first-order factors

	General		F1: Verbal Comprehension		F2: Working Memory		F3: Perceptual Reasoning		F4: Processing Speed			
WISC-V subtest	b	S^2	b	S^2	b	S^2	b	S^2	b	S^2	h^2	u^2
Similarities	.665	.442	**.374**	**.140**	.082	.007	−.001	.000	−.008	.000	.589	.411
Vocabulary	.640	.410	**.515**	**.265**	−.042	.002	.032	.001	−.080	.006	.684	.316
Information	.680	.462	**.445**	**.198**	−.002	.000	.002	.000	.092	.008	.669	.331
Comprehension	.578	.334	**.332**	**.110**	.073	.005	−.020	.000	.015	.000	.450	.550
Block Design	.575	.331	.015	.000	−.010	.000	**.353**	**.125**	.123	.015	.471	.529
Visual Puzzles	.634	.402	−.008	.000	.000	.000	**.481**	**.231**	−.047	.002	.636	.364
Matrix Reasoning	.668	.446	.020	.000	.123	.015	**.283**	**.080**	.036	.001	.543	.457
Figure Weights	.522	.272	.054	.003	.096	.009	**.212**	**.045**	−.036	.001	.331	.669
Picture Concepts	.507	.257	.057	.003	.093	.009	**.163**	**.027**	.053	.003	.298	.702
Arithmetic	.647	.419	.070	.005	**.307**	**.094**	−.013	.000	.030	.001	.519	.481
Digit Span	.705	.497	−.007	.000	**.384**	**.147**	.044	.002	−.054	.003	.649	.351
Picture Span	.489	.239	−.017	.000	**.244**	**.060**	.060	.004	−.013	.000	.303	.697
Letter–Number Sequencing	.666	.444	.030	.001	**.384**	**.147**	−.020	.000	−.055	.003	.595	.405
Coding	.335	.112	−.052	.003	.049	.002	−.045	.002	**.611**	**.373**	.493	.507
Symbol Search	.472	.223	.021	.000	.036	.001	−.003	.000	**.612**	**.375**	.599	.401
Cancellation	.151	.023	.034	.001	−.136	.018	.094	.009	**.346**	**.120**	.171	.829
Total variance		.332		.046		.032		.033		.057	.500	.500
Common variance		.664		.091		.065		.066		.114		
	ω_H = .821		ω_{HS} = .253		ω_{HS} = .174		ω_{HS} = .165		ω_{HS} = .478			

Note. b = loading of subtest on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_H = omega–hierarchical, ω_{HS} = omega-hierarchical subscale. Bold type shows coefficients and variance estimates consistent with the theoretically proposed factor.

Table 3 also shows ω_H and ω_{HS} that were estimated based on the SL results. The ω_H coefficient for general intelligence (.821) was high and sufficient for scale interpretation; however, the ω_{HS} coefficients for the four group factors (VC, WM, PR, PS) were considerably lower (.174–.478). Thus, for the four group factors, with the possible exception of PS, unit-weighted composite scores based on these indicators would likely possess too little true score variance for clinical interpretation (Reise, 2012; Reise et al., 2013) for the 6- to 8-year-old age group.

2.3.3 | Ages 9–11 first-order EFA: Four-factor extraction

Table 4 shows results of four-factor extraction with promax rotation for 9- to 11-year-olds. The *g* loadings ranged from .226 (Cancellation) to .803 (Vocabulary) and all were within the fair-to-good range (except Coding, Symbol Search, Cancellation). Picture Concepts failed to exhibit salient pattern loadings on any group factor. Table 4 shows robust Verbal Comprehension (Similarities, Vocabulary, Information, Comprehension), Perceptual Reasoning (Block Design, Visual Puzzles, Matrix Reasoning, Figure Weights), Working Memory (Arithmetic, Digit Span, Picture Span, Letter–Number Sequencing), and Processing Speed (Coding, Symbol Search, Cancellation) factors with theoretically consistent subtest associations. Picture Concepts was again a fair indicator of general intelligence but was not adequately associated with any of the four group factors; although its highest pattern coefficient was on Perceptual Reasoning. There were no subtests with salient cross-loadings. The moderate-to-high factor correlations shown in Table 4 (.392 to .724) imply a higher-order or hierarchical structure that requires explication (Gorsuch, 1983) and the Schmid–Leiman procedure was applied to better understand variance apportionment among general and group factors.

2.3.4 | Ages 9–11 SL bifactor analyses: Four first-order factors

Results for the Schmid and Leiman orthogonalization of the higher-order factor analysis are shown in Table 5. All subtests were properly associated (higher residual variance) with their theoretically proposed factor after removing *g* variance except Picture Concepts, which had equivalent residual loadings with Perceptual Reasoning and Verbal Comprehension. The *g* factor accounted for 33.6% of the total variance and 64.1% of the common variance.

The general factor also accounted for between 4.0% (Cancellation) and 52.4% (Vocabulary) of individual subtest variability. At the first-order level, VC accounted for an additional 5.4% of the total variance and 10.4% of the common variance, PR accounted for an additional 3.3% of the total variance and 6.4% of the common variance, WM accounted for an additional 3.6% of the total variance and 6.9% of the common variance, and PS accounted for an additional 6.4% of the total variance and 12.3% of the common variance. The general and group factors combined to measure 52.4% of the variance in WISC-V scores resulting in 47.6% unique variance (combination of specific and error variance).

Also presented in Table 5 are ω_H and ω_{HS} coefficients that were estimated based on the SL results. The ω_H coefficient for general intelligence (.817) was high and sufficient for scale interpretation; however, the ω_{HS} coefficients for the four group factors (VC, WM, PR, PS) were considerably lower (.064–.517). Thus, unit-weighted composite scores for the four group factors, with the possible exception of PS, would likely possess too little true-score variance for clinical interpretation (Reise, 2012; Reise et al., 2013) for the 9- to 11-year-old age group.

2.3.5 | Ages 12–14 first-order EFA: Four-factor extraction

Table 6 shows results of four-factor extraction with promax rotation for 12- to 14-year-olds. The *g* loadings ranged from .252 (Cancellation) to .806 (Vocabulary) and all were within the fair-to-good range (except Coding, Symbol Search, Cancellation). Picture Concepts and Arithmetic had salient factor pattern coefficients on the Verbal Comprehension factor but no other factors. Table 6 shows robust Verbal Comprehension (Similarities, Vocabulary, Information, Comprehension), Working Memory (Digit Span, Picture Span, Letter–Number Sequencing), Perceptual Reasoning (Block Design, Visual Puzzles, Matrix Reasoning, Figure Weights), and Processing Speed (Coding, Symbol Search, Cancellation) factors with theoretically consistent subtest associations. Oddly, Picture Concepts and Arithmetic migrated away from their theoretically consistent factors to the Verbal Comprehension factor. No salient cross-loadings were observed. The moderate-to-high factor correlations presented in Table 6 (.399 to .732) imply a higher-order or hierarchical

TABLE 4 Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) exploratory factor analysis: Four oblique factor solution for the standardization sample 9- to 11-year-olds ($n = 600$)

	General[a]	F1: Verbal Comprehension		F2: Perceptual Reasoning		F3: Working Memory		F4: Processing Speed		
WISC-V subtest	*S*	*P*	*S*	*P*	*S*	*P*	*S*	*P*	*S*	h^2
Similarities	.744	**.805**	.815	.034	.602	−.035	.539	.022	.336	.665
Vocabulary	.803	**.872**	.883	.043	.652	−.014	.588	−.027	.327	.781
Information	.749	**.807**	.823	.035	.608	.005	.555	−.032	.301	.678
Comprehension	.646	**.668**	.694	−.064	.495	.087	.510	.038	.313	.488
Block Design	.683	−.031	.541	**.780**	.770	−.053	.511	.110	.417	.603
Visual Puzzles	.653	.119	.568	**.673**	.718	−.024	.488	−.056	.276	.524
Matrix Reasoning	.532	−.021	.421	**.475**	.562	.114	.450	.056	.311	.326
Figure Weights	.627	.020	.522	**.715**	.710	.027	.488	−.084	.252	.510
Picture Concepts	.507	.239	.470	.262	.479	.031	.395	.052	.276	.265
Arithmetic	.712	.165	.608	.220	.630	**.387**	.677	.060	.408	.526
Digit Span	.647	−.080	.491	.084	.541	**.802**	.770	−.075	.315	.601
Picture Span	.562	.031	.446	−.059	.429	**.666**	.659	.025	.330	.436
Letter–Number Sequencing	.677	.100	.560	−.026	.532	**.729**	.766	−.028	.350	.591
Coding	.461	−.044	.302	−.023	.345	.124	.421	**.714**	.746	.563
Symbol Search	.455	.040	.317	−.017	.340	−.038	.362	**.801**	.792	.628
Cancellation	.226	.008	.151	.051	.185	−.119	.148	**.474**	.443	.203
Eigenvalue		6.61		1.52		1.07		.99		
% Variance		38.59		6.60		3.92		3.31		
Promax-Based Factor Correlations		F1: VC		F2: PR		F3: WM		F4: PS		
F1: Verbal Comprehension (VC)		–								
F2: Perceptual Reasoning (PR)		.724		–						
F3: Working Memory (WM)		.672		.682		–				
F4: Processing Speed (PS)		.392		.441		.479		–		

[a]Factor structure coefficients from first unrotated factor (*g* loadings) are correlations between the subtest and the general factor. *S* = structure coefficient, *P* = pattern coefficient, h^2 = communality. Salient pattern coefficients are shown in bold (pattern coefficient ≥ .30).

TABLE 5 Sources of variance in the Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) for the standardization sample 9- to 11-year-olds ($n = 600$) according to an exploratory SL bifactor model with four first-order factors

	General		F1: Verbal Comprehension		F2: Perceptual Reasoning		F3: Working Memory		F4: Processing Speed			
WISC-V subtest	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	*b*	S^2	h^2	u^2
Similarities	.670	.449	**.463**	**.214**	.018	.000	−.020	.000	.019	.000	.664	.336
Vocabulary	.724	.524	**.502**	**.252**	.023	.001	−.008	.000	−.023	.001	.777	.223
Information	.677	.458	**.464**	**.215**	.018	.000	.003	.000	−.027	.001	.675	.325
Comprehension	.584	.341	**.384**	**.147**	−.034	.001	.049	.002	.032	.001	.493	.507
Block Design	.652	.425	−.018	.000	**.409**	**.167**	−.030	.001	.094	.009	.602	.398
Visual Puzzles	.621	.386	.068	.005	**.353**	**.125**	−.013	.000	−.048	.002	.517	.483
Matrix Reasoning	.511	.261	−.012	.000	**.249**	**.062**	.064	.004	.048	.002	.330	.670
Figure Weights	.603	.364	.012	.000	**.375**	**.141**	.015	.000	−.071	.005	.510	.490
Picture Concepts	.471	.222	*.138*	.019	**.137**	**.019**	.017	.000	.044	.002	.262	.738
Arithmetic	.674	.454	.095	.009	.115	.013	**.218**	**.048**	.051	.003	.527	.473
Digit Span	.630	.397	−.046	.002	.044	.002	**.451**	**.203**	−.064	.004	.608	.392
Picture Span	.539	.291	.018	.000	−.031	.001	**.374**	**.140**	.021	.000	.432	.568
Letter–Number Sequencing	.648	.420	.058	.003	−.014	.000	**.410**	**.168**	−.024	.001	.592	.408
Coding	.423	.179	−.025	.001	−.012	.000	.070	.005	**.608**	**.370**	.554	.446
Symbol Search	.408	.166	.023	.001	−.009	.000	−.021	.000	**.682**	**.465**	.633	.367
Cancellation	.201	.040	.005	.000	.027	.001	−.067	.004	**.403**	**.162**	.208	.792
Total Variance		.336		.054		.033		.036		.064	.524	.476
Common Variance		.641		.104		.064		.069		.123		
	$\omega_H = .817$		$\omega_{HS} = .280$		$\omega_{HS} = .174$		$\omega_{HS} = .207$		$\omega_{HS} = .517$			

Note. b = loading of subtest on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_H = omega-hierarchical, ω_{HS} = omega-hierarchical subscale. Bold type shows coefficients and variance estimates consistent with the theoretically proposed factor. Italic type shows coefficients and variance estimates associated with an alternate factor (where cross-loading b was larger than for the theoretically assigned factor). For theoretical and practical reasons, Picture Concepts was assigned to Perceptual Reasoning for omega subscale estimation.

TABLE 6 Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) exploratory factor analysis: Four oblique factor solution for the standardization sample 12- to 14-year-olds ($n = 600$)

	General[a]	F1: Verbal Comprehension		F2: Working Memory		F3: Perceptual Reasoning		F4: Processing Speed		
WISC-V Subtest	*S*	*P*	*S*	*P*	*S*	*P*	*S*	*P*	*S*	h^2
Similarities	.791	**.770**	.833	.013	.620	.074	.646	−.002	.344	.697
Vocabulary	.806	**.919**	.879	−.011	.623	−.034	.624	−.018	.328	.774
Information	.771	**.807**	.828	−.054	.581	.103	.641	−.040	.300	.691
Comprehension	.692	**.794**	.746	.016	.550	−.134	.500	.096	.363	.570
Block Design	.667	.048	.563	−.021	.519	**.656**	.731	.124	.416	.547
Visual Puzzles	.702	−.001	.593	−.073	.522	**.889**	.836	−.008	.340	.702
Matrix Reasoning	.663	.228	.603	.154	.566	**.343**	.627	.032	.349	.451
Figure Weights	.708	.169	.635	.206	.613	**.464**	.702	−.058	.315	.542
Picture Concepts	.515	**.347**	.512	.047	.414	.199	.469	−.037	.211	.282
Arithmetic	.737	**.329**	.687	.262	.658	.223	.646	.013	.373	.548
Digit Span	.718	−.022	.592	**.892**	.840	−.022	.545	−.042	.392	.708
Picture Span	.629	−.050	.514	**.631**	.688	.165	.538	−.036	.335	.485
Letter–Number Sequencing	.716	.136	.620	**.783**	.808	−.128	.510	.024	.420	.662
Coding	.446	.007	.307	.082	.421	−.110	.299	**.805**	.802	.649
Symbol Search	.490	−.053	.342	.035	.433	.132	.413	**.686**	.739	.558
Cancellation	.252	.057	.185	−.161	.174	.068	.213	**.489**	.459	.222
Eigenvalue		7.36		1.51		.98		.87		
% Variance		43.50		6.56		3.62		3.11		
Promax-based factor correlations		F1: VC		F2: WM		F3: PR		F4: PS		
F1: Verbal Comprehension (VC)		–								
F2: Working Memory (WM)		.725		–						
F3: Perceptual Reasoning (PR)		.732		.674		–				
F4: Processing Speed (PS)		.399		.507		.433		–		

[a]Factor structure coefficients from first unrotated factor (*g* loadings) are correlations between the subtest and the general factor. S = structure coefficient, P = pattern coefficient, h^2 = communality. Salient pattern coefficients are shown in bold (pattern coefficient ≥ .30).

structure that required explication (Gorsuch, 1983) and the Schmid–Leiman procedure was applied to better understand variance apportionment among general and group factors.

2.3.6 | Ages 12–14 SL bifactor analyses: Four first-order factors

Results for the Schmid and Leiman orthogonalization of the higher-order factor analysis are shown in Table 7. All subtests were properly associated (higher residual variance) with their theoretically proposed factor after removing *g* variance except Picture Concepts and Arithmetic, which had somewhat higher residual loadings with the Verbal Comprehension factor. The *g* factor accounted for 38.3% of the total variance and 67.4% of the common variance.

The general factor also accounted for between 5.1% (Cancellation) and 53.6% (Vocabulary) of individual subtest variability. At the first-order level, VC accounted for an additional 5.4% of the total variance and 9.5% of the common variance, WM accounted for an additional 3.3% of the total variance and 5.8% of the common variance, PR accounted for an additional 3.5% of the total variance and 6.2% of the common variance, and PS accounted for an additional 6.3% of the total variance and 11.1% of the common variance. The general and group factors combined to measure 56.7% of the variance in WISC-V scores resulting in 43.3% unique variance (combination of specific and error variance).

Table 7 also shows ω_H and ω_{HS} coefficients that were estimated based on the SL results. Because of subtest migration of Picture Concepts and Arithmetic on Verbal Comprehension, omega-hierarchical and omega-subscale coefficients were estimated with Picture Concepts and Arithmetic loadings on Verbal Comprehension as well as with their theoretically consistent loadings on Perceptual Reasoning and Working Memory, respectively. The ω_H coefficient for general intelligence (.847, .842) was high and sufficient for scale interpretation; however, the ω_{HS} coefficients for the four group factors (VC, WM, PR, PS) were considerably lower (.149–.503, .173–.503). Thus, unit-weighted composite scores for the four group factors based on these indicators, with the possible exception of PS, likely possess too little true score variance for clinical interpretation (Reise, 2012; Reise et al., 2013) for 12- to 14-year-olds.

2.3.7 | Ages 15–16 first-order EFA: Four-factor extraction

Table 8 shows the results of four-factor extraction with promax rotation. The *g* loadings ranged from .243 (Cancellation) to .813 (Vocabulary) and all were within the fair-to-good range (except Coding, Symbol Search, and Cancellation). Picture Concepts had a salient pattern coefficient on the Verbal Comprehension factor. Arithmetic failed to exhibit salient pattern loadings on any group factor but had split loadings on Verbal Comprehension (.299), Working Memory (.291), and Perceptual Reasoning (.291), that would be salient considering a confidence interval. Figure Weights had a secondary cross-loading with Verbal Comprehension. Table 8 shows robust Verbal Comprehension (Similarities, Vocabulary, Information, Comprehension), Working Memory (Digit Span, Picture Span, Letter–Number Sequencing), Perceptual Reasoning (Block Design, Visual Puzzles, Matrix Reasoning, Figure Weights), and Processing Speed (Coding, Symbol Search, Cancellation) factors with theoretically consistent subtest associations. Picture Concepts again migrated away from its theoretically related factor to the Verbal Comprehension factor. The moderate-to-high factor correlations shown in Table 6 (.323 to .754) imply a higher-order or hierarchical structure that required explication (Gorsuch, 1983) and the Schmid–Leiman procedure was applied to better understand variance apportionment among general and group factors.

2.3.8 | Ages 15–16 SL bifactor analyses: Four first-order factors

Results for the Schmid and Leiman orthogonalization of the higher-order factor analysis are shown in Table 9. All subtests were properly associated (higher residual variance) with their theoretically proposed factor after removing *g* variance except Picture Concepts, which had higher residual loading on the Verbal Comprehension factor. The *g* factor accounted for 37.5% of the total variance and 66.7% of the common variance.

The general factor also accounted for between 5.2% (Cancellation) and 56.9% (Arithmetic) of individual subtest variability. At the first-order level, VC accounted for an additional 5.1% of the total variance and 9.1% of the common variance, WM accounted for an additional 4.2% of the total variance and 7.4% of the common variance, PS accounted for an additional 6.8% of the total variance and 12.1% of the common variance, and PR accounted for an additional 2.6%

TABLE 7 Sources of variance in the Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) for the standardization sample 12- to 14-year-olds ($n = 600$) according to an exploratory SL bifactor model with four first-order factors

	General		F1: Verbal Comprehension		F2: Working Memory		F3: Perceptual Reasoning		F4: Processing Speed			
WISC-V subtest	***b***	***S*²**	***b***	***S*²**	***b***	***S*²**	***b***	***S*²**	***b***	***S*²**	***h*²**	***u*²**
Similarities	.724	.524	**.408**	**.166**	.007	.000	.042	.002	−.002	.000	.692	.308
Vocabulary	.732	.536	**.487**	**.237**	−.006	.000	−.019	.000	−.015	.000	.774	.226
Information	.702	.493	**.428**	**.183**	−.028	.001	.058	.003	−.034	.001	.681	.319
Comprehension	.628	.394	**.421**	**.177**	.008	.000	−.076	.006	.081	.007	.584	.416
Block Design	.628	.394	.025	.001	−.011	.000	**.372**	**.138**	.105	.011	.545	.455
Visual Puzzles	.664	.441	−.001	.000	−.038	.001	**.504**	**.254**	−.007	.000	.696	.304
Matrix Reasoning	.625	.391	.121	.015	.079	.006	**.195**	**.038**	.027	.001	.450	.550
Figure Weights	.671	.450	.090	.008	.106	.011	**.263**	**.069**	−.049	.002	.541	.459
Picture Concepts	.479	.229	*.184*	*.034*	.024	.001	**.113**	**.013**	−.031	.001	.278	.722
Arithmetic	.694	.482	*.174*	*.030*	**.135**	**.018**	.127	.016	.011	.000	.546	.454
Digit Span	.705	.497	−.012	.000	**.459**	**.211**	−.012	.000	−.036	.001	.709	.291
Picture Span	.615	.378	−.027	.001	**.325**	**.106**	.094	.009	−.031	.001	.494	.506
Letter–Number Sequencing	.694	.482	.072	.005	**.403**	**.162**	−.073	.005	.020	.000	.655	.345
Coding	.412	.170	.004	.000	.042	.002	−.062	.004	**.683**	**.466**	.642	.358
Symbol Search	.457	.209	−.028	.001	.018	.000	.075	.006	**.582**	**.339**	.554	.446
Cancellation	.225	.051	.030	.001	−.083	.007	.039	.002	**.415**	**.172**	.232	.768
Total Variance		.383		.054		.033		.035		.063	.567	.433
Common Variance		.674		.095		.058		.062		.111		
PC with PR, AR with WM	$\omega_H = .847$		$\omega_{HS} = .252$		$\omega_{HS} = .163$		$\omega_{HS} = .149$		$\omega_{HS} = .503$			
PC and AR with VC	$\omega_H = .842$		$\omega_{HS} = .195$		$\omega_{HS} = .213$		$\omega_{HS} = .173$		$\omega_{HS} = .503$			

Note. b = loading of subtest on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_H = omega-hierarchical, ω_{HS} = omega-hierarchical subscale. Bold type shows coefficients and variance estimates consistent with the theoretically proposed factor. Italic type shows coefficients and variance estimates associated with an alternate factor (where cross-loading b was larger than for the theoretically assigned factor).

TABLE 8 Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) exploratory factor analysis: Four oblique factor solution for the standardization sample 15- to 16-year-olds ($n = 400$)

	General[a]	F1: Verbal Comprehension		F2: Working Memory		F3: Processing Speed		F4: Perceptual Reasoning		
WISC-V Subtest	*S*	*P*	*S*	*P*	*S*	*P*	*S*	*P*	*S*	h^2
Similarities	.767	**.956**	.845	.000	.582	.051	.288	−.169	.574	.726
Vocabulary	.813	**.755**	.856	.046	.629	−.081	.236	.126	.692	.743
Information	.749	**.767**	.804	−.055	.546	−.054	.224	.123	.642	.654
Comprehension	.680	**.698**	.716	−.010	.513	.068	.292	.004	.553	.517
Block Design	.728	.032	.614	−.106	.515	.135	.449	**.813**	.824	.696
Visual Puzzles	.709	.148	.636	−.055	.515	−.004	.321	**.701**	.774	.607
Matrix Reasoning	.640	.171	.577	.194	.552	−.009	.280	**.361**	.615	.425
Figure Weights	.680	**.303**	.653	.076	.537	−.118	.200	**.443**	.672	.513
Picture Concepts	.529	**.313**	.510	.121	.443	.018	.229	.141	.465	.283
Arithmetic	.799	.299	.732	.291	.707	.026	.367	.291	.722	.645
Digit Span	.713	−.074	.567	**.770**	.812	.052	.392	.108	.589	.667
Picture Span	.614	.088	.517	**.625**	.686	.077	.342	−.048	.468	.478
Letter–Number Sequencing	.673	.055	.568	**.899**	.825	−.053	.277	−.137	.482	.692
Coding	.387	−.006	.245	.035	.324	**.739**	.745	−.016	.318	.555
Symbol Search	.451	.078	.313	−.036	.348	**.786**	.801	.009	.380	.645
Cancellation	.243	−.100	.141	.048	.210	**.396**	.428	.105	.230	.190
Eigenvalue		7.24		1.61		1.03		.82		
% Variance		42.78		7.20		4.14		2.36		
Promax-based factor correlations		F1: VC		F2: WM		F3: PS		F4: PR		
F1: Verbal Comprehension (VC)		–								
F2: Working Memory (WM)		.705		–						
F3: Processing Speed (PS)		.323		.412		–				
F4: Perceptual Reasoning (PR)		.754		.667		.427		–		

[a]Factor structure coefficients from first unrotated factor (*g* loadings) are correlations between the subtest and the general factor. *S* = structure coefficient, *P* = pattern coefficient, h^2 = communality. Salient pattern coefficients are shown in bold (pattern coefficient ≥ .30).

TABLE 9 Sources of variance in the Wechsler Intelligence Scale for Children-Fifth Edition (WISC-V) for the standardization sample 15- to 16-year-olds ($n = 400$) according to an exploratory SL bifactor model with four first-order factors

	General		F1: Verbal Comprehension		F2: Working Memory		F3: Processing Speed		F4: Perceptual Reasoning			
WISC-V subtest	b	S^2	b	S^2	b	S^2	b	S^2	b	S^2	h^2	u^2
Similarities	.688	.473	**.507**	**.257**	.000	.000	.045	.002	−.085	.007	.740	.260
Vocabulary	.749	.561	**.400**	**.160**	.027	.001	−.072	.005	.063	.004	.731	.269
Information	.687	.472	**.407**	**.166**	−.032	.001	−.048	.002	.061	.004	.645	.355
Comprehension	.619	.383	**.370**	**.137**	−.006	.000	.060	.004	.002	.000	.524	.476
Block Design	.707	.500	.017	.000	−.062	.004	.120	.014	**.406**	**.165**	.683	.317
Visual Puzzles	.686	.471	.078	.006	−.032	.001	−.004	.000	**.350**	**.123**	.600	.400
Matrix Reasoning	.611	.373	.091	.008	.113	.013	−.008	.000	**.180**	**.032**	.427	.573
Figure Weights	.648	.420	.161	.026	.044	.002	−.105	.011	**.222**	**.049**	.508	.492
Picture Concepts	.494	.244	*.166*	*.028*	.071	.005	.016	.000	**.070**	**.005**	.282	.718
Arithmetic	.754	.569	.158	.025	**.170**	**.029**	.023	.001	.145	.021	.644	.356
Digit Span	.680	.462	−.039	.002	**.449**	**.202**	.046	.002	.054	.003	.671	.329
Picture Span	.576	.332	.047	.002	**.364**	**.132**	.068	.005	−.024	.001	.472	.528
Letter–Number Sequencing	.634	.402	.029	.001	**.524**	**.275**	−.047	.002	−.069	.005	.684	.316
Coding	.349	.122	−.003	.000	.020	.000	**.656**	**.430**	−.008	.000	.553	.447
Symbol Search	.406	.165	.041	.002	−.021	.000	**.698**	**.487**	.004	.000	.654	.346
Cancellation	.227	.052	−.053	.003	.028	.001	**.352**	**.124**	.052	.003	.182	.818
Total Variance		.375		.051		.042		.068		.026	.562	.438
Common Variance		.667		.091		.074		.121		.047		
PC with PR	$\omega_H = .844$		$\omega_{HS} = .241$		$\omega_{HS} = .209$		$\omega_{HS} = .530$		$\omega_{HS} = .108$			
PC with VC	$\omega_H = .841$		$\omega_{HS} = .214$		$\omega_{HS} = .209$		$\omega_{HS} = .530$		$\omega_{HS} = .131$			

Note. b = loading of subtest on factor, S^2 = variance explained, h^2 = communality, u^2 = uniqueness, ω_H = omega-hierarchical, ω_{HS} = omega-hierarchical subscale. Bold type shows coefficients and variance estimates consistent with the theoretically proposed factor. Italic type shows coefficients and variance estimates associated with an alternate factor (where cross-loading b was larger than for the theoretically assigned factor).

of the total variance and 4.7% of the common variance. The general and group factors combined to measure 56.2% of the variance in WISC-V scores resulting in 43.8% unique variance (combination of specific and error variance).

Also shown in Table 9 are ω_H and ω_{HS} coefficients that were estimated based on the SL results. Because of subtest migration of Picture Concepts on Verbal Comprehension, omega-hierarchical and omega-subscale coefficients were estimated with Picture Concepts loading on Verbal Comprehension as well as with its theoretically consistent loading on Perceptual Reasoning. The ω_H coefficient for general intelligence (.844, .841) was high and sufficient for scale interpretation; however, the ω_{HS} coefficients for the four group factors (VC, WM, PR, PS) were considerably lower (.108–.530, .131–.530). Thus, for the four group factors, with the possible exception of PS, unit-weighted composite scores based on these indicators would likely possess too little true score variance for clinical interpretation (Reise, 2012; Reise et al., 2013) for the 15- to 16-year-old age group.

2.4 | One-, two-, and three-factor extraction

Examination of results when extracting fewer than four factors paralleling those of Canivez et al. (2016) resulted in structures that were not consistent with previous versions of the WISC nor other Wechsler scales. One-, two-, and three-factor models fused theoretically meaningful constructs indicative of underextraction and were judged unsatisfactory (Gorsuch, 1983; Wood et al., 1996).

3 | DISCUSSION

The WISC-V *Technical and Interpretive Manual* claimed support for a five first-order and one higher-order (*g*) factor model for the 16 primary and secondary subtests. Structural validity support was based exclusively on CFA as no EFA results were included. Also absent were decomposed variance estimates (or any variance estimates) for the higher-order and lower-order factors and model-based reliability (ω_H and ω_{HS}) estimates that would provide users of the WISC-V information necessary for judging the psychometric fitness of provided scores (Canivez, 2010, 2014a; Canivez & Kush, 2013; Rodriguez et al., 2016). Given the absence of these necessary analyses and summary statistics, the present study used EFA and hierarchical EFA methods to assess the WISC-V structure to examine CFA and EFA convergence or divergence among four age groups in the WISC-V standardization sample.

Consistent with the findings from Canivez et al. (2016), who investigated the WISC-V structure with the total standardization sample, the present study also indicated there was no EFA evidence to support a five-factor representation of the WISC-V within any of the four age groups examined (see Figures A1–A4 in Appendix A and Tables B1 through B8 in Appendix B in the online supplemental materials). Forced extraction of five factors resulted in the fifth factor having only one subtest with a salient factor pattern loading and is inadequate (Preacher & MacCallum, 2003).

Also consistent with Canivez et al. (2016) was general support for most subtests' association with a four-factor model that was similar to the WISC-IV. In each of the four age groups, the Verbal Comprehension subtests (Similarities, Vocabulary, Information, Comprehension), Working Memory subtests (Digit Span, Picture Span, Letter–Number Sequencing), Perceptual Reasoning subtests (Block Design, Visual Puzzles, Matrix Reasoning, Figure Weights), and Processing Speed subtests (Coding, Symbol Search, Cancellation) were consistently associated with the theoretical constructs previously posited (i.e., WISC-IV, WAIS-IV, WPPSI-IV) despite changes in subtest content. The subtests thought to represent separate Visual Spatial (Block Design and Visual Puzzles) and Fluid Reasoning (Matrix Reasoning and Figure Weights) factors merged together in all four age groups and appear to represent the former Perceptual Reasoning factor present in the WISC-IV and WAIS-IV. It appears that FW and MR are weaker indicators of Perceptual Reasoning than are BD and VP, but they clearly did not produce a separate Fluid Reasoning factor. These results, as with those from Canivez et al. (2016), fail to support the publisher's creation of separate Visual Spatial and Fluid Reasoning factors and standardized factor index scores that represent them. Other evidence of problems with specifying separate Visual Spatial and Fluid Reasoning factors is present in the redundant loading of FR on general intelligence reported in CFA in the WISC-V *Technical and Interpretive Manual* and shown in Figure 1, as well as in Chen et al. (2015). Thus, it appears that the WISC-V has been overfactored as represented in the WISC-V *Technical and Interpretive Manual*.

Following transformation with the Schmid and Leiman (1957) procedure, the WISC-V *g* factor accounted for 5 to 6 times more variance than any single group factor and approximately twice the variance of all four group factors combined in all four age groups. To further show the general dominance of subtest measurement of general intelligence, Figure 2 shows the portions of subtest variance apportioned to the general intelligence dimension and the portions of subtest variance apportioned to the four WISC-V group factors. With the exception of the CD, SS, and CN subtests; most common subtest variance was that associated with general intelligence in each of the four age groups and that what is primarily measured is general intelligence, not the group factors.

Also, the ω_H coefficients for the *g* factor in all four age groups (.817–.847) were high and indicated large portions of true score variance attributable to unit-weighted scores based on all subtests. The ω_{HS} coefficients for the four group factors in all four age groups were considerably lower (range of .131 to .280 for the VC, PR, and WM factors), falling far below the minimum threshold of .50 suggested by Reise (2012) and Reise et al. (2013) for confident clinical interpretation. That is, they captured too little unique true score variance once *g* variance was removed. The ω_{HS} coefficients for the PS factor in all four age groups ranged from .478 to .530 and approached or met the minimum standard for possible interpretation. These results appear to support Carroll's model but not Cattell–Horn, as pointed out by Cucina and Howardson (2017).

Arithmetic was associated with Working Memory for the 6–8 and 9- to 11-year-old age groups, but migrated to Verbal Comprehension for the 12- to 14-year-old age group and was not saliently associated with any group factor in the 15- to 16-year-old age group (its variance spread evenly between VC, PR, and WM). Numerous problems with Arithmetic as a subtest in Wechsler scales have been described (Canivez & Kush, 2013; Canivez et al., 2016; Watkins & Ravert, 2013). As suggested previously (Canivez & Kush, 2013; Canivez et al., 2015; Watkins & Ravert, 2013) it is likely time for Arithmetic to be removed as an indicator of Working Memory.

As observed by Canivez et al. (2016), Picture Concepts failed to demonstrate salient loadings on any factors in the 6–8 and 9- to 11-year-old age groups and when it did saliently load on a factor it was on a theoretically inconsistent one (VC). This may be the reason the publisher does not include Picture Concepts in any regularly calculated factor-based scores (PC is only used to replace a Fluid Reasoning subtest in calculating the FSIQ because of spoiling either Matrix Reasoning or Figure Weights). Given its failure to saliently load on any latent factor, its inclusion as a substitute for Matrix Reasoning or Figure Weights for estimating the FSIQ from a Fluid Reasoning area may be questionable.

The superiority of general intelligence observed in all four age groups is identical to that found by Canivez et al. (2016) with the total WISC-V standardization sample and similar to other studies of Wechsler scales using both EFA and CFA methods (Bodin et al., 2009; Canivez, 2014a; Canivez & Watkins, 2010a, 2010b; Canivez et al., 2017a; Dombrowski, Canivez, & Watkins, 2018; Gignac & Watkins, 2013; Lecerf & Canivez, 2017; McGill & Canivez, 2016, 2017; Nelson et al., 2013; Watkins, 2006; 2010; Watkins & Beaujean, 2014; Watkins et al., 2006, 2013, 2017) and other intelligence tests (Canivez, 2008; Canivez & McGill, 2016; Canivez et al., 2009; Cucina & Howardson, 2017; DiStefano & Dombrowski, 2006; Dombrowski, 2013, 2014a, 2014b; Dombrowski & Watkins, 2013; Dombrowski et al., 2009; Dombrowski, Golay, McGill & Canivez, 2018a; Dombrowski, McGill, & Canivez, 2017a, 2017b, 2018b; Dombrowski, McGill, Canivez & Peterson, 2018c; Nelson & Canivez, 2012; Nelson et al., 2007). These results are also consistent with the broader professional literature on the importance and dominance of general intelligence (Deary, 2013; Jensen, 1998; Lubinski, 2000; Ree et al., 2003).

As would be predicted by Frazier and Youngstrom (2007), too little true score variance was associated with the four WISC-V group factors, with the possible exception of PS, to warrant confident clinical interpretation (Reise, 2012; Reise et al., 2013). Gustafsson (1984) noted that, "individual differences in cognitive performance can be understood in terms of several sources of variance, some of which are broad and some of which are narrow" (p. 67) and Gorsuch (1983) explained that, "in science, the concern is with generalizing as far as possible and as accurately as possible. Only when the broad and not so broad generalities do not apply to a given solution does one move to the narrowest, most specific level of generality" (p. 249). Most of the WISC-V variance was contributed by a broad general factor so the WISC-V general factor is "of definite interest" (Gorsuch, 1983, p. 253) but the "lower order factors may be of little

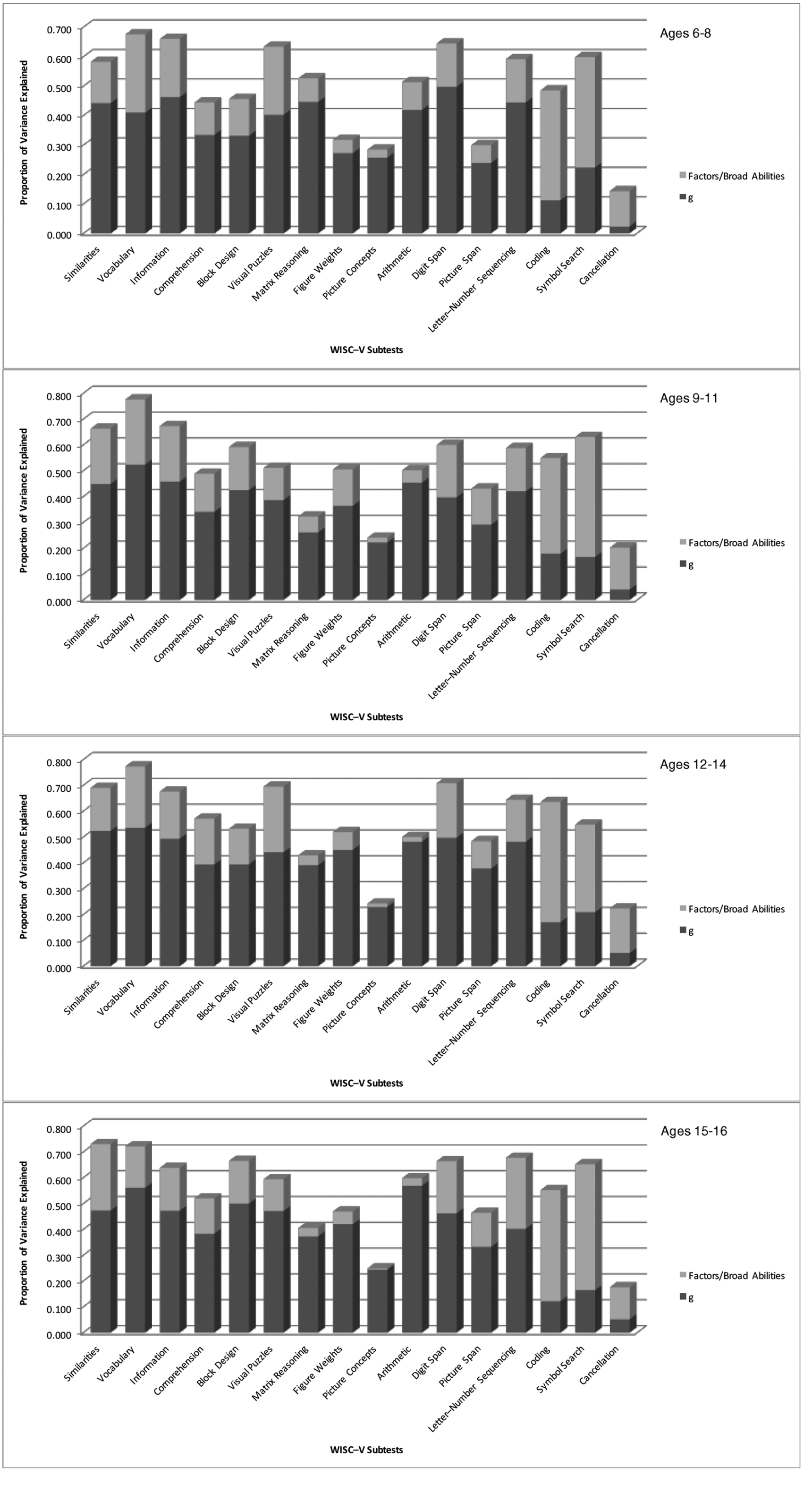

FIGURE 2 Sources of variance for the 16 WISCV primary and secondary subtests for the four age groups based on Schmid and Leiman (1957) orthogonalization of higher-order extraction with four first-order factors (VC, PR, WM, PS) based on Tables 3, 5, 7, and 9.

interest" (Wolff & Preising, 2005, p. 50). As pointed out by Cucina and Howardson (2017), such evidence supports the three-stratum theory proposed by Carroll (1993, 2003) but not the structure advanced by Cattell–Horn, which ostensibly is a two-stratum model (no *g* factor).

Given the absence of important information from the WISC-V *Technical and Interpretive Manual* as described in the present study as well as results from Canivez et al. (2016) and Canivez et al. (2017a), researchers and clinicians using the WISC-V must rely on the extant literature to adequately evaluate which WISC-V scores have sufficient reliability and validity for interpretation and use. Numerous studies have published results at odds with those provided in test technical manuals (cf. Canivez, 2008; Canivez & McGill, 2016; Canivez & Watkins, 2010a, 2010b; DiStefano & Dombrowski, 2006; Dombrowski, 2013, 2014a, 2014b; Dombrowski et al., 2009; Dombrowski, McGill, & Canivez, 2017a,b; McGill & Canivez, 2017; Watkins, 2006), but such information should have been included in those technical manuals in the first place.

Researchers and clinicians must rely on more than the test technical manuals to use test scores appropriately as they bear "the ultimate responsibility for appropriate test use and interpretation" (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014, p. 141). The present results, in addition to those of Canivez et al. (2016, 2017a), will assist users of the WISC-V to "know what their tests can do and act accordingly (Weiner, 1989, p. 829).

3.1 | Limitations

Correlations provided in the *Technical and Interpretive Manual Supplement* (Wechsler, 2014c) were analyzed because NCS Pearson declined to provide the WISC-V standardization sample raw data. Analytical methods such as exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) might be a viable alternative to traditional EFA, but ESEM requires participant raw data, which were unavailable. Thus, the correlations from the technical manual were used but are rounded to two decimals and therefore could be less precise than correlations produced from raw data. However, greater precision would not be warranted by the sample size of each age group (Bedeian, Sturman, & Streiner, 2009) and it is unlikely that the present results were substantially impacted by two-digit precision (Carroll, 1993). Another limitation is that the present study, while informative, may provide results that differ from those that might be produced by a CFA bifactor model. Reise (2012) indicated that the EFA-based SL procedure produces an approximate bifactor solution that is a reparameterization of the higher-order structure and contains proportionality constraints (Yung et al., 1999), but the SL procedure is the dominant exploratory approach to assessing bifactor structure in EFA. Use of CFA bifactor modeling as well as examination of factor invariance across these four age groups will further test the latent structure of the WISC-V and the present results will facilitate plausible CFA models to test invariance examination (Brown, 2015; Carroll, 1998). Such analyses would extend those of Reynolds and Keith (2017) by examining invariance of the bifactor structure with four group factors rather than only the first-order subtest alignment. Finally, these results may not extend to populations not well represented in the WISC-V normative sample. For example, profoundly gifted individuals may exhibit meaningful cognitive patterns that do not emerge among standardization samples (Robertson, Smeets, Lubinski, & Benbow, 2010).

4 | CONCLUSIONS

Results from this study provide important considerations for clinical interpretation of scores from the WISC-V. The results of analyses across the four age groups support interpretation of the general intelligence estimate (FSIQ). Lower-order (index scores) are generally not supported for interpretation with the possible exception of the PSI. Independent analyses of the WISC-V failed to support the test publisher's posited five-factor structure. Because there was no evidence for separate Visual Spatial and Fluid Reasoning factors in any of these four age groups or the full standardization sample (Canivez et al., 2016, 2017a), the publisher should consider producing revised norms tables for a four-factor model where the former Perceptual Reasoning factor is estimated in place of separate Visual Spatial and Fluid

Reasoning factors. The overfactoring of the WISC-V in the WISC-V *Technical and Interpretive Manual* and factor index scores for VS and FR will likely result in misinterpretation and errors in clinical decision making (Beaujean, 2015b; Dombrowski, 2015). As shown in the present study as well as with the full standardization sample (Canivez et al., 2016; 2017a; Dombrowski, Canivez, Watkins, & Beaujean, 2015), primary interpretation of the WISC-V should be at the FSIQ level and consideration of other score interpretations must be made in light of the extremely small portions of true score variance uniquely captured by the factor index scores.

ORCID

Gary L. Canivez http://orcid.org/0000-0002-5347-6534
Stefan C. Dombrowski http://orcid.org/0000-0002-8057-3751
Marley W. Watkins http://orcid.org/0000-0001-6352-7174

REFERENCES

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.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438.

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

Beaujean, A. A. (2015b). Adopting a new test edition: Psychometric and practical considerations. *Research and Practice in the Schools*, *3*, 51–57.

Beaujean, A. A. (2016). Reproducing the Wechsler Intelligence Scale for Children-Fifth Edition: Factor model results. *Journal of Psychoeducational Assessment*, *34*, 404–408.

Bedeian, A. G., Sturman, M. C., & Streiner, D. L. (2009). Decimal dust, significant digits, and the search for stars. *Organizational Research Methods*, *12*, 687–694.

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.

Briggs, N. E., & MacCallum, R. C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research*, *38*, 25–56.

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.

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.

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.

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

Canivez, G. L. (2014b). 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. (2016). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements* (pp. 247–271). Gottingen, Germany: Hogrefe.

Canivez, G. L., & 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.

Canivez, G. L., & McGill, R. J. (2016). Factor structure of the Differential Ability Scales-Second Edition: Exploratory and hierarchical factor analyses with the core subtests. *Psychological Assessment*, *28*, 1475–1488.

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.

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.

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., 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.

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.

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

Canivez, G. L., Watkins, M. W., & Dombrowski, S. C. (2017a). Structural validity of the Wechsler Intelligence Scale for Children-Fifth Edition: Confirmatory factor analyses with the 16 primary and secondary subtests. *Psychological Assessment*, *29*, 458–472.

Canivez, G. L., Watkins, M. W., Good, R., James, K., & James, T. (2017b). Construct validity of the WISC-IV UK with a referred Irish sample: Wechsler and CHC model comparisons with 15 subtests. *British Journal of Educational Psychology*, *87*, 383–407.

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

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

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

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

Carroll, J. B. (1998). Human cognitive abilities: A critique. In J. J. McArdle & R. W. Woodcock (Eds.), *Human cognitive abilities in theory and practice* (pp. 5–23). Mahwah, NJ: Erlbaum.

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.

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

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.

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.

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.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Child, D. (2006). *The essentials of factor analysis*(3rd ed.). New York, NY: Continuum.

Crawford, A. V., Green, S. B., Levy, R., Lo, W.-J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement*, *70*, 885–901.

Cucina, J., & Byle, K. (2017). The bifactor model fits better than the higher-order model in more than 90% of comparisons for mental abilities test batteries. *Journal of Intelligence*, *5*, 27–47.

Cucina, J. M., & Howardson, G. N. (2017). Woodcock–Johnson-III, Kaufman Adolescent and Adult Intelligence Test (KAIT), Kaufman Assessment Battery for Children (KABC), and Differential Ability Scales (DAS) support Carroll but not Cattell–Horn. *Psychological Assessment*, *29*, 1001–1015.

Deary, I. J. (2013). Intelligence. *Current Biology*, *23*, 673–676.

DeVellis, R. F. (2017). *Scale development: Theory and applications*(4th ed.). Thousand Oaks, CA: Sage.

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

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

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.

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.

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

Dombrowski, S. C., Canivez, G. L., & Watkins, M. W. (2018). Factor structure of the 10 WISCV primary subtests across four standardization age groups. *Contemporary School Psychology*, *22*, 90–104.

Dombrowski, S. C., McGill, R. J., & Canivez, G. L. (2017a). Exploratory and hierarchical factor analysis of the WJ IV Cognitive at school age. *Psychological Assessment*, *29*, 394–407.

Dombrowski, S. C., McGill, R. J., & Canivez, G. L. (2017b). Exploratory and hierarchical factor analysis of the WJ IV Full Test battery. *School Psychology Quarterly*, https://doi.org/10.1037/spq0000221

Dombrowski, S. C., Golay, P., McGill, R. J., & Canivez, G. L. (2018a). Investigating the theoretical structure of the DAS-II core battery at school age using Bayesian structural equation modeling. *Psychology in the Schools*, *55*, 190–207.

Dombrowski, S. C., McGill, R. J., & Canivez, G. L. (2018b). An alternative conceptualization of the theoretical structure of the WJ IV Cognitive at school age: A confirmatory factor analytic investigation. *Archives of Scientific Psychology*, *6*, 1–13.

Dombrowski, S. C., McGill, R. J., Canivez, G. L., & Peterson, C. H. (2018c). Investigating the theoretical structure of the Differential Ability Scales-Second Edition through hierarchical exploratory factor analysis. *Journal of Psychoeducational Assessment*, https://doi.org/10.1177/0734282918760724

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.

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.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272–299.

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.

Gignac, G. E. (2005). Revisiting the factor structure of the WAIS-R: Insights through nested factor modeling. *Assessment*, *12*, 320–329.

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.

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.

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

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

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. *Journal of Special Education*, *40*, 103–114.

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.

Goldberg, L. R., & Velicer, W. F. (2006). Principles of exploratory factor analysis. In S. Strack (Ed.), *Differentiating normal and abnormal personality*(2nd ed., pp. 209–237). New York, NY: Springer.

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

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment*, *68*, 532–560.

Gustafsson, J. E. (1984). A unifying model for structure of intellectual abilities. *Intelligence*, *8*, 179–203.

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

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

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, 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*(3rd ed., pp. 73–98). 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.

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*, 537–549.

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

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151.

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

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children-Second Edition*. Circle Pines, MN: AGS.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*(3rd ed.). New York, NY: Guilford.

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.

Lecerf, T., & Canivez, G. L. (2017). Complementary exploratory and confirmatory factor analyses of the French WISC-V: Analyses based on the standardization sample. *Psychological Assessment*, https://doi.org/10.1037/pas0000526

Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling*, *13*, 59–72.

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.

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

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

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

McGill, R. J., & Canivez, G. L. (2017). Confirmatory factor analyses of the WISC-IV Spanish core and supplemental subtests: Validation evidence of the Wechsler and CHC models. *International Journal of School and Educational Psychology*, https://doi.org/10.1080/21683603.2017.1327831

Millsap, R. E. (2007). Structural equation modeling made difficult. *Personality and Individual Differences*, *42*, 875–881.

Morin, A. J., Arens, A. K., Tran, A., & Caci, H. (2016). Exploring sources of construct-relevant multidimensionality in psychiatric measurement: A tutorial and illustration using the Composite Scale of Morningness. *International Journal of Methods in Psychiatric Research*, *25*(4), 277–288.

Mundfrom, D. J., & Shaw, D. G. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, *5*, 159–168.

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

Nelson, J. M., & Canivez, G. L. (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.

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.

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.

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics*, *2*, 13–43.

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

Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, *32*, 329–353.

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.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696.

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.

Reise, S. P., Moore, T. M., & Maydeu-Olivares, A. (2011). Target rotations and assessing the impact of model violations on the parameters of unidimensional item response theory models. *Educational and Psychological Measurement*, *71*, 684–711.

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.

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

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

Reynolds, M. R., & Keith, T. Z. (2017). Multi-group and hierarchical confirmatory factor analysis of the Wechsler Intelligence Scale for Children-Fifth Edition. *Intelligence*, *62*, 31–47.

Robertson, K. F., Smeets, S., Lubinski, D., & Benbow, C. P. (2010). Beyond the threshold hypothesis: Even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence. *Current Directions in Psychological Science*, *19*, 346–351.

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

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

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

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

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

Strauss, E., Spreen, O., & Hunter, M. (2000). Implications of test revisions for research. *Psychological Assessment*, *12*, 237–244.

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.

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

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.

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

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

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

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

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

Watkins, M. W. (2010). Structure of the Wechsler Intelligence Scale for Children-Fourth Edition among a national sample of referred students. *Psychological Assessment*, *22*, 782–787.

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

Watkins, M. W. (2017). The reliability of multidimensional neuropsychological measures: From alpha to omega. *The Clinical Neuropsychologist*, *31*, 1113–1126.

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.

Watkins, M. W., & Ravert, C. M. (2013). Subtests, factors, and constructs: What is being measured by tests of intelligence? In J. C. Kush (Ed.), *Intelligence quotient: Testing, role of genetics and the environment and social outcomes* (pp. 55–68). Hauppauge, NY: Nova.

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.

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.

Watkins, M. W., Dombrowski, S. C., & Canivez, G. L. (2017). Reliability and factorial validity of the Canadian Wechsler Intelligence Scale for Children-Fifth Edition. *International Journal of School and Educational Psychology*, https://doi.org/10.1080/21683603.2017.1342580

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

Wechsler, D. (1949). *Wechsler intelligence scale for children*. New York, NY: The Psychological Corporation.

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.

Wechsler, D. (2014c). *Technical and interpretive manual supplement: Special group validity studies with other measures and additional tables*. San Antonio, TX: NCS Pearson.

Weiner, I. B. (1989). On competence and ethicality in psychodiagnostic assessment. *Journal of Personality Assessment*, *53*, 827–831.

Wolff, H.-G., & Preising, K. (2005). Exploring item and higher order factor structure with the Schmid–Leiman solution: Syntax codes for SPSS and SAS. *Behavior Research Methods*, *37*, 48–58.

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.

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

Yung, Y.–F., Thissen, D., & McLeod, L. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*, 113–128.

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.

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.

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

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

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Canivez GL, Dombrowski SC, Watkins MW. Factor structure of the WISC-V in four standardization age groups: Exploratory and hierarchical factor analyses with the 16 primary and secondary subtests. *Psychol Schs*. 2018;55:741–769. https://doi.org/10.1002/pits.22138