Exploratory and Higher-Order Factor Analyses of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) Adolescent Subsample

Gary L. Canivez  
Eastern Illinois University

Marley W. Watkins  
Arizona State University

The factor structure of the Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV; Wechsler, 2008a) with the adolescent participants (ages 16–19 years; N = 400) in the standardization sample was assessed using exploratory factor analysis, multiple factor extraction criteria, and higher-order exploratory factor analyses. Results from exploratory factor analyses were not included in the WAIS-IV Technical and Interpretation Manual (Wechsler, 2008b) and are necessary for determining convergence or divergence with the reported confirmatory factor analyses. As found with the total WAIS-IV standardization sample (Canivez & Watkins, in press), the present results with the adolescent subsample found all WAIS-IV subtests (10- and 15-subtest configurations) were properly associated with their four theoretically proposed first-order factors, but only one factor extraction criterion (standard error of scree) recommended extraction of four factors. Hierarchical exploratory analyses with the Schmid and Leiman (1957) procedure found that the second-order g factor accounted for major portions of total and common variance, while the four first-order factors accounted for small portions of total and common variance. It was concluded that the WAIS-IV provides strong measurement of general intelligence in adolescents and clinical interpretation should be primarily at that level.

Keywords: WAIS-IV, exploratory factor analysis, factor extraction criteria, Schmid-Leiman higher-order analysis, structural validity

The Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV; Wechsler, 2008a) is the latest version of one of the most frequently used intelligence tests for adults and older adolescents (Goh, Teslow, & Fuller, 1981; Hutton, Dubes, & Muir, 1992; Stinnett, Havey, & Oehler-Stinnett, 1995; Watkins, Campbell, Nieberding, & Hallmark, 1995). It includes 15 subtests (10 core and 5 supplemental), four first-order factor index scores (Verbal Comprehension [VC], Perceptual Reasoning [PR], Working Memory [WM], and Processing Speed [PS]), and the higher-order Full Scale score (FSIQ). Verbal and Performance IQs are no longer available and the Object Assembly and Picture Arrangement subtests were deleted, thus reducing subtests with manipulative objects. Three new subtests were created (Visual Puzzles, Figure Weights, Cancellation), and item coverage and range were increased. Like other recently published intelligence tests such as the Wechsler Intelligence Scale for Children–Fourth Edition (WISC-IV; Wechsler, 2003), the Stanford-Binet Intelligence Scales–Fifth Edition (SB-5; Roid, 2003a), Kaufman Assessment Battery for Children–Second Edition (KABC-II; Kaufman & Kaufman, 2004; Reynolds Intellectual Assessment Scales (RIAS; Reynolds & Kamphaus, 2003), and Wide Range Intelligence Test (WRIT; Glutting, Adams, & Sheslow, 2000); the WAIS-IV content and structure reflect current conceptualizations of intelligence articulated by Carroll, Cattell, and Horn (Car-
The WAIS-IV provides a FSIQ, factor index scores, index score discrepancies, ipsative subtest comparisons (strengths/weaknesses), and pairwise subtest comparisons (Wechsler, 2008b) for interpretation. However, interpreting each of these various test scores and comparisons requires due consideration of psychometric evidence presented in the test manual and the extant literature (American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education, 1999). One important aspect of test score validity derives from the latent structure of the test and an important continuing debate is the degree to which intelligence tests fundamentally measure fewer versus more dimensions. Largely supported by confirmatory factor analysis (CFA), intelligence test authors and publishers claim to measure multiple dimensions of intellectual ability beyond the general intelligence factor and, based on this assertion, proffer broad interpretations. With respect to the factor index scores of the WAIS-IV, it was noted in the Technical and Interpretive Manual (Wechsler, 2008b) that analyses of these four index scores is recommended as the primary level of clinical interpretation, especially in cases with considerable variability across the index and or subtest scores” (p. 127).

The WAIS-IV Technical and Interpretive Manual presented final CFA structural models with standardized coefficients for the 10 core subtests (ages 16–90) and 15 core and supplementary subtests (ages 16–69) in Figures 5.1 and 5.2, respectively, to illustrate the hierarchical structure of the WAIS-IV. Goodness-of-fit statistics presented for total samples showed superior model fit for the WAIS-IV hierarchical structure that allowed the Arithmetic subtest to load on both WM and VC factors, although the standardized coefficients for the VC to Arithmetic paths appeared generally small. While fit indices for this model were superior to the model with Arithmetic loading solely on WM, improvements appeared modest. In addition, Tables 5.2 and 5.3 presented separate CFA analyses results across 5 age groups, including the adolescent subsample (ages 16–19), illustrating similarities and differences. With respect to the WAIS-IV adolescent subsample (N = 400), no model improvement was observed for the 10 core subtest CFA when allowing Arithmetic to load on both WM and VC. Goodness-of-fit statistics from the 15 core and supplemental subtest models supported the superiority of the model that allowed the Arithmetic subtest to load on both the WM and VC factors. CFA analyses for the adolescent subsample thus supported the hierarchical model with general intelligence at the highest level and four first-order factors consistent with theory and construction of the WAIS-IV, whether or not Arithmetic loaded on VC.

Unfortunately, the WAIS-IV Technical and Interpretive Manual presented only CFA results in support of the latent factor structure and provided no exploratory factor analytic (EFA) procedures or results. Many consider EFA and CFA to be complimentary procedures, answering different questions. However, a recent trend has been for test authors and publishers to present only CFA results to support the latent structure of tests (Elliott, 2007; McGrew & Woodcock, 2001; Roid, 2003b; Wechsler, 2008b). In contrast, a number of previous and current editions of tests included both EFA and CFA results (Bracken & McCallum, 1998; Elliott, 1990; Glutting, Adams, & Sheslow, 2000; Kaufman & Kaufman, 1993; Naglieri & Das, 1997; Wechsler, 1991, 2002a, 2002b; Wechsler & Naglieri, 2006). When EFA and CFA are in agreement, there is greater confidence in the latent structure of the test (Gorsuch, 1983).

The problem that Frazier and Youngstrom (2007) illustrated regarding the disagreement between the number of latent factors reported in contemporary intelligence tests is that CFA procedures and the most liberal EFA factor extraction criteria (eigenvalues &gt;1 and scree) suggest greater numbers of factors than EFA procedures that included the most psychometrically sound methods for determining the correct number of factors to extract and retain (parallel analysis and minimum average partials). Without presentation of EFA procedures and results with standardization sample data, school psychologists are unable to consider convergence or divergence of WAIS-IV CFA and EFA results. Such information is important in determining relative importance of various scores for interpretation, particularly when the Technical and Interpretive Manual designates the four factor
Several investigations of major intelligence tests using EFA procedures have recently been published and challenge the optimistic conclusions of CFA results illustrated in the respective test technical manuals. Two studies of the SB-5 standardization sample (Canivez, 2008; DiStefano & Dombrowski, 2006) obtained significantly different results than those reported in the SB-5 technical manual (Roid, 2003b). Both studies concluded that the SB-5 measured only one dimension \((g)\) and found no evidence to support the existence of the five factors reported by Roid (2003b). Three studies of the WISC-IV (Bodin, Pardini, Burns, & Stevens, 2009; Watkins, 2006; Watkins, Wilson, Kotz, Carbone, & Babula, 2006) showed that most variance was associated with general intelligence and substantially smaller amounts of variance were related to the first-order factors. These studies concluded that interpretation of the WISC-IV should focus on the global FSIQ score because it accounts for most of the common and total variance and additional research showed the superior predictive validity of FSIQ (Glutting, Watkins, Konold, & McDermott, 2006; Glutting et al., 1997). Further, the limited unique variance captured by the four first-order factors may be responsible for the limited incremental predictive validity of factor scores observed in the WISC–III and WISC-IV. Two studies of the RIAS also indicated that it fundamentally measures a single general intelligence factor (Dombrowski, Watkins, & Brogan, 2009; Nelson, Canivez, Lindstrom, & Hatt, 2007), which was the primary goal of its authors (Reynolds & Kamphaus, 2003). A recent joint investigation of the WRIT and Wechsler Abbreviated Scale of Intelligence (WASI; Psychological Corporation, 1999) also found substantial variability associated with general intelligence and smaller portions of variance apportioned to the first-order factors; supporting primary interpretation of the FSIQ/GIQ (Canivez, Konold, Collins, & Wilson, 2009).

The WAIS-IV Technical and Interpretive Manual does not present proportions of variance accounted for by the higher-order \(g\)-factor and the four first-order factors, subtest \(g\)-loadings, subtest specificity estimates, or incremental predictive validity estimates of the four factors and subtests. Thus, school psychologists are unable to judge the relative importance of the factor index scores and subtest scores relative to the FSIQ. If the factor index scores and subtests do not capture meaningful portions of true score variance nor provide important amounts of incremental predictive validity, they may be of questionable clinical utility and should be deemphasized or perhaps eliminated in test score interpretation.

Major tests of intelligence, including the WAIS-IV, have applied Carroll’s (1993) model of the structure of cognitive abilities. Carroll’s (1993, 2003) 3-stratum theory of cognitive abilities is hierarchical, proposing some 50–60 narrow abilities (Stratum I), 8–10 broad ability factors (Stratum II), and at the apex (Stratum III), the general ability factor \((g);\) Spearman, 1904, 1927). Carroll’s model has been used by test authors to facilitate subtest and factor selection and to aid in interpretations of scores and performance. However, subtest performance on cognitive ability tests reflects combinations of both first-order (Stratum II) and second-order (Stratum III) factors and Carroll argued that the Schmid and Leiman (1957) procedure must be used to first extract variance from the higher-order factor to residualize the lower-order factors, leaving them orthogonal to the higher-order factor. Specifically, Carroll (1995) argued that:

> from the standpoint of analysis and ready interpretation, results should be shown on the basis of orthogonal factors, rather than oblique, correlated factors. I insist, however, that the orthogonal factors should be those produced by the Schmid-Leiman, 1957, orthogonalization procedure, and thus include second-stratum and possibly third-stratum factors. (p. 437)

Variability associated with a higher-order factor must be accounted for before interpreting variability associated with lower-order factors. The Schmid and Leiman procedure was recommended by Carroll (1993, 1995, 1997, 2003); McClain (1996); Gustafsson and Snow (1997); Carretta and Ree (2001); Ree, Carretta, and Green (2003); and Thompson (2004). In addition, it was used in the previously discussed investigations of the SB-5 (Canivez, 2008), WISC-IV (Watkins, 2006; Watkins et al., 2006), RIAS (Dombrowski et al., 2009; Nelson et al., 2007), and WRIT and WASI (Canivez et al., 2009).

and found that in the 10 core subtest (ages 16–90), 15 core and supplemental subtest (ages 16–69), and 12 core and supplemental subtest (ages 70–90) configurations that the second-order \((g)\) dimension accounted for large portions of common (67%–69.1%) and total (40.6%–44.7%) variance while the first-order (factor index scores) dimensions accounted for appreciably smaller portions of common and total variance. It was concluded that primary interpretation should reside at the FSIQ level rather than the factor index score level recommended by Wechsler (2008b). These analyses, however, were performed on the total standardization sample, which included only 18% adolescents (16–19) for WAIS-IV core subtests and 22% adolescents for WAIS-IV core and supplemental subtests. Given the large number of older participants, results from the total sample may not generalize to the adolescent subsample. For example, the structure of the Wechsler Adult Intelligence Scale–Third Edition (WAIS-III; Wechsler, 1991) varied across age groups (Sattler, 2001); differential Wechsler verbal versus performance profiles have been found for children, adolescents, and adults (Isen, 2010); and the structure of the WAIS-IV was not invariant across age groups (Benson, Hulac, & Kranzler, 2010). In addition, genetic studies have demonstrated that heritability increases across childhood and adolescence and peaks in adulthood (Plomin, 2004) and that heritability may be differentially responsible for general and specific factors (Plomin & Spinath, 2002).

To provide necessary information for school psychologists to judge CFA results in the WAIS-IV Technical and Interpretive Manual (Wechsler, 2008b), the present study utilized the adolescent subsample \((N = 400)\) data from the WAIS-IV standardization sample to examine the factor structure using EFA procedures as conducted in Canivez and Watkins (2010). The primary research questions were (a) using multiple criteria, how many factors are recommended to be extracted and retained from the WAIS-IV adolescent standardization sample; and (b) when forcing extraction of four theoretical factors and applying the Schmid and Leiman (1957) procedure, what portions of variance are attributed to the general intelligence (Stratum III) dimension and the four broad ability factors (Stratum II)? Analyses were provided for the two principal test configurations for adolescents: the 10 Core Subtests and the 10 Core and 5 Supplemental Subtests, which parallel CFA models examined and reported in the Technical and Interpretive Manual. If multiple factors and levels of the WAIS-IV are to be interpreted for adolescents, it is imperative school psychologists know how variability is apportioned across the first- and second-order dimensions.

### Method

#### Participants

Participants were the 400 individuals between 16 and 19 years of age included in the WAIS-IV standardization sample. Detailed demographic characteristics are provided in the WAIS-IV Technical and Interpretive Manual. In short, the standardization sample was obtained using stratified proportional sampling across variables of age, gender, race/ethnicity, education level (or parent education level for ages 16–19), and geographic region. Examination of tables in the Technical and Interpretive Manual revealed a close match to the October 2005 U.S. census across stratification variables. The adolescent sample included individuals within the following racial/ethnic categories: 256 (64.0%) White/Caucasian, 53 (13.3%) Black/African American, 64 (16.0%) Hispanic/Latino, 11 (2.8%) Asian American, and 16 (4.0%) other. Parent education levels of the adolescents were as follows: 17 (4.3%) ≤8 years, 32 (8.0%) 9–11 years, 114 (28.5%) 12 years, 133 (33.3%) 13–15 years, and 104 (26.0%) ≥16 years. Geographic distribution of the adolescent sample included 75 (18.8%) Northeast, 92 (23.0%) Midwest, 128 (32.0%) South, and 105 (26.3%) West.

#### Instrument

The WAIS-IV is an individual test of general intelligence for ages 16 to 90 that originated with the 1939 Wechsler-Bellevue Intelligence Scale (Wechsler, 1939a). Consistent with Wechsler’s definition of intelligence (i.e., “global capacity,” Wechsler, 1939b, p. 229), the WAIS-IV measures general intelligence through the administration of numerous subtests, each of which is an indicator and es-
timate of intelligence. The WAIS-IV uses 10 core subtests to produce the FSIQ. The Verbal Comprehension Index (VCI) and Perceptual Reasoning Index (PRI) are each composed of 3 subtests while the Working Memory Index (WMI) and Processing Speed Index (PSI) are each composed of 2 subtests. Supplemental subtests (Comprehension, Figure Weights, Picture Completion, Letter-Number Sequencing, and Cancellation) are provided to substitute for core subtests when necessary (1 each for the VC, WM, and PS scales and 2 for the PR scale).

Procedure

WAIS-IV subtest correlation matrices for the two adolescent age groups (ages 16–17 and 18–19) in the standardization sample were obtained from the Technical and Interpretive Manual and combined by averaging observed correlations through Fisher’s z transformations (Barker, 1990; Guilford & Fruchter, 1978; Silver & Dunlap, 1987). Two correlation matrices were created to represent the two WAIS-IV subtest configurations examined with CFA in the WAIS-IV Technical and Interpretive Manual (Wechsler, 2008b): the 10 core subtests and the 15 core and supplementary subtests.

Analyses

Principal axis exploratory factor analyses (Cudeck, 2000; Fabrigar, Wegener, MacCallum, & Strahan, 1999; Tabachnick & Fidel, 2007) were used to analyze reliable common variance from each of the two WAIS-IV standardization sample correlation matrices representing the two configurations (10 subtests, 15 subtests) using SPSS 17.0 for Macintosh OSX. As recommended by Gorsuch (1983), multiple criteria for determining the number of factors to retain were examined and included eigenvalues >1 (Guttman, 1954), the visual 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 was used to visually determine the optimum number of factors to retain but is a subjective criterion. The SE_scree, reportedly the most accurate objective scree method (Nasser, Benson, & Wisenbaker, 2002), was used as programmed by Watkins (2007). HPA and MAP were included as they typically are more accurate and therefore reduce overfactoring (Frazier & Youngstrom, 2007; Thompson & Daniel, 1996; Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986). HPA indicated meaningful factors when eigenvalues from the WAIS-IV standardization sample data were larger than eigenvalues produced by random data containing the same number of participants and factors (Lautenschlager, 1989). Random data and resulting eigenvalues for HPA were produced using the Monte Carlo PCA for Parallel Analysis computer program (Watkins, 2000) with 100 replications to provide stable eigenvalue estimates. The MAP criterion was computed using the SPSS code supplied by O’Connor (2000).

The present study limited iterations in first-order principal axis factor extraction to two in estimating final communality estimates (Gorsuch, 2003). Each correlation matrix for the two WAIS-IV configurations was subjected to EFA (principal axis extraction of four factors), followed by promax (oblique) rotation (k = 4; Gorsuch, 2003) and the resulting first-order factors were orthogonalized 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). Four first-order factors were extracted to compare results to other studies of Wechsler scales (Canivez & Watkins, 2010; Watkins, 2006; Watkins et al., 2006) as well as to examine proportions of variance attributed to WAIS-IV factors identified through CFA by Wechsler (2008b) with the standardization sample.

Results

Factor Extraction Criteria

Figures 1 and 2 display scree plots from HPA for the two WAIS-IV configurations. Table 1 summarizes results from the multiple criteria (eigenvalues >1, scree test, standard error of scree, HPA, and MAP) for determining the number of factors to extract and retain in each of the WAIS-IV configurations. Of the objec-
tive criteria illustrated in Table 1, only the $SE_{\text{scree}}$ supported extraction of four factors and the eigenvalue >1 criterion supported extraction of three factors for the 15-subtest WAIS-IV configuration. The visual scree test (Figures 1 and 2) showed one strong factor but it might be argued there is support for two factors. HPA recommended extraction of only two factors while MAP suggested one factor among 10 WAIS-IV core subtests and two factors for the 15 WAIS-IV core and supplementary subtests.

Higher Order Factor Analyses

**WAIS-IV 10 core subtests.** Schmid and Leiman (1957) procedure results for the 10 WAIS-IV core subtests with the adolescent standardization sample (ages 16–19; $N = 400$) are presented in Table 2. All subtests were properly associated with their theoretically proposed factors. Correlations between the four first-order factors based on promax rotation ranged from .45 to .70, suggesting the presence of a higher-order factor (Gorsuch, 1983; Tabachnick & Fidell, 2007). The second-order $g$ factor accounted for 37.9% of the total variance and 65.0% of the common variance. The general factor also accounted for between 22% and 50% ($Mdn = 41\%$) of individual subtest variability. At the first-order level, VC accounted for an additional 7.1% of the total variance and 12.2% of the common variance, PR accounted for an additional 4.0% of the total variance and 6.9% of the common variance, WM accounted for an additional 3.3% of the total variance and 10.2% of the common variance. The first- and second-order factors combined to measure 58.4% of the variance in WAIS-IV scores resulting in 41.6% unique variance (combination of specific and error variance). Subtest specificity (variance unique to the subtest) estimates ranged from .22 to .41 ($Mdn = .28$).

**WAIS-IV 15 subtests.** Schmid and Leiman (1957) procedure results for the 15 WAIS-IV core and supplemental subtests with the adolescent standardization sample (ages 16–19; $N = 400$) are presented in Table 3. All subtests were correctly aligned with their theoretically proposed factors. Based on promax rotation, correlations between the four first-order factors ranged from .46 to .72, indicating the presence of a higher-order factor (Gorsuch, 1983; Tabachnick & Fidell, 2007). The second-order $g$ factor accounted for 37.1% of the total variance and 50% ($Mdn = 41\%$) of individual subtest variability. At the first-order level, VC accounted for an additional 7.1% of the total variance and 12.2% of the common variance, PR accounted for an additional 4.0% of the total variance and 6.9% of the common variance, WM accounted for an additional 3.3% of the total variance and 10.2% of the common variance. The first- and second-order factors combined to measure 58.4% of the variance in WAIS-IV scores resulting in 41.6% unique variance (combination of specific and error variance). Subtest specificity (variance unique to the subtest) estimates ranged from .22 to .41 ($Mdn = .28$).
variance and 65.5% of the common variance. The general factor also accounted for between 22% and 51% (Mdn = 39%) of individual subtest variability. At the first-order level, VC accounted for an additional 7.3% of the total variance and 12.8% of the common variance, PR accounted for an additional 2.6% of the total variance and 4.7% of the common variance, WM accounted for an additional 4.6% of the total variance and 8.2% of the common variance, and PS accounted for an additional 5.0% of the total variance and 8.8% of the common variance. The first- and second-order factors combined to measure 56.6% of the variance in WAIS-IV scores resulting in 43.4% unique variance (combination of specific and error variance). Subtest specificity (variance unique to the subtest) estimates ranged from .17 to .47 (Mdn = .30).

**Discussion**

Although the WAIS-IV Technical and Interpretive Manual presented CFA support for an hierarchical structure with g at the apex and four first-order factors in the adolescent subsample (ages 16–19), consideration of convergence or divergence of CFA and EFA results is not pos-
sible given the absence of EFA procedures in the manual. Further, the WAIS-IV Technical and Interpretive Manual also failed to provide proportions of variance accounted for by the first-order factor index scores, subtest $g$-loadings, and subtest specificity estimates.

The present study examined the WAIS-IV factor structure among the adolescent standardization subsample using EFA methods to answer two basic research questions: (a) how many factors should be extracted and retained using multiple criteria and (b) when four factors are extracted and orthogonalized using the Schmid and Leiman (1957) procedure, how was variance apportioned to the first- and second-order dimensions?

Multiple criteria for determining the number of factors to extract and retain included HPA and MAP because of their superior accuracy (Thompson & Daniel, 1996; Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986). The Schmid and Leiman (1957) procedure was used to examine the WAIS-IV hierarchical structure and to apportion variance to the first- and second-order factors as recommended by Carroll (1993, 1995, 1997, 2003); McClain (1996); Gustafsson and Snow (1997); Carretta and Ree (2001); Ree, Carretta, and Green (2003); and Thompson (2004). These analyses were necessary for school psychologists to consider the relative adequacy of different WAIS-IV scores (e.g., subtest, index, FSIQ), as well as convergence or divergence of CFA and EFA results within the adolescent subsample within the WAIS-IV standardization sample.

Interpreting each of the test scores and comparisons requires due consideration of the psychometric evidence of each presented in a test manual and the extant literature (American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education, 1999). The present study found that when considering multiple factor extraction criteria across the two adolescent WAIS-IV configurations (10 and 15 subtests), only the SScree supported extraction of four factors for the 15-subtest configuration. All other criteria and configurations suggested that fewer factors be extracted. This is consistent with the results obtained by Frazier and Youngstrom (2007) and divergent from the CFA results presented in the WAIS-IV Technical and Interpretive Manual. Consistent with studies of the WISC-IV (Watkins, 2006; Watkins et al., 2006), RIAS (Dombrowski et al., 2009; Nelson et al., 2007), WRIT, and WASI (Canivez et al., 2009); and the total WAIS-IV standardization samples (Canivez & Watkins, 2006).
2010); the present study found that WAIS-IV subtests were aligned with the four theoretically proposed factors identified through CFA (Wechsler, 2008b). However, the second-order g factor accounted for a major proportion of total and common variance and the variance apportioned to the WAIS-IV first-order factors was considerably smaller.

As noted by Gustafsson (1994), “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). 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). In the present study, most of the WAIS-IV variance was contributed by a broad and general factor. In such cases, the broad factor is “of definite interest” (Gorsuch, 1983, p. 253) and “lower order factors may be of little interest” (Wolff & Preising, 2005, p. 50). The recommendation in the WAIS-IV Technical and Interpretive Manual for primary interpretation of factor index scores is inconsistent with these results.

Another consideration relates to CFA and EFA procedures that examined the 15-subtest WAIS-IV configuration, as clinicians do not typically administer all 15 WAIS-IV subtests. The five supplemental subtests available for 16–19 year olds are used only to replace core subtests. Therefore, while theoretical support is claimed for CFA results for the 15-subtest configuration, there is no provision for analysis and interpretation when all available subtests are administered (Wechsler, 2008b). Given this practice, results from the 10 core subtests seem most relevant to clinical application in school psychology practice.

The WAIS-IV appears to be an excellent measure of general intelligence for adolescents and has admirable norms, but divergent CFA and EFA results call into question the viability of the factor structure and resulting scores. However, factor analytic methods (CFA and

<table>
<thead>
<tr>
<th>WAIS-IV subtest</th>
<th>General</th>
<th>VC</th>
<th>PR</th>
<th>WM</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>%S²</td>
<td>b</td>
<td>%S²</td>
<td>B</td>
</tr>
<tr>
<td>SI</td>
<td>0.58</td>
<td>34</td>
<td>0.55</td>
<td>30</td>
<td>-0.01</td>
</tr>
<tr>
<td>VC</td>
<td>0.64</td>
<td>41</td>
<td>0.54</td>
<td>30</td>
<td>-0.01</td>
</tr>
<tr>
<td>IN</td>
<td>0.64</td>
<td>41</td>
<td>0.40</td>
<td>16</td>
<td>0.05</td>
</tr>
<tr>
<td>CO</td>
<td>0.63</td>
<td>39</td>
<td>0.54</td>
<td>29</td>
<td>0.00</td>
</tr>
<tr>
<td>BD</td>
<td>0.71</td>
<td>51</td>
<td>-0.02</td>
<td>0</td>
<td>0.35</td>
</tr>
<tr>
<td>MR</td>
<td>0.63</td>
<td>39</td>
<td>0.04</td>
<td>0</td>
<td>0.21</td>
</tr>
<tr>
<td>VP</td>
<td>0.70</td>
<td>49</td>
<td>0.00</td>
<td>0</td>
<td>0.38</td>
</tr>
<tr>
<td>FW</td>
<td>0.68</td>
<td>47</td>
<td>0.11</td>
<td>1</td>
<td>0.20</td>
</tr>
<tr>
<td>PCm</td>
<td>0.56</td>
<td>31</td>
<td>0.12</td>
<td>1</td>
<td>0.17</td>
</tr>
<tr>
<td>DS</td>
<td>0.62</td>
<td>39</td>
<td>-0.03</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>AR</td>
<td>0.69</td>
<td>47</td>
<td>0.08</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>LN</td>
<td>0.58</td>
<td>34</td>
<td>0.01</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>SS</td>
<td>0.47</td>
<td>22</td>
<td>0.00</td>
<td>0</td>
<td>0.01</td>
</tr>
<tr>
<td>CD</td>
<td>0.47</td>
<td>22</td>
<td>0.04</td>
<td>0</td>
<td>-0.04</td>
</tr>
<tr>
<td>CA</td>
<td>0.47</td>
<td>22</td>
<td>-0.08</td>
<td>1</td>
<td>0.06</td>
</tr>
<tr>
<td>% Total S²</td>
<td>37.1</td>
<td>7.3</td>
<td>2.6</td>
<td>4.6</td>
<td>5.0</td>
</tr>
<tr>
<td>% Common S²</td>
<td>65.5</td>
<td>12.8</td>
<td>4.7</td>
<td>8.2</td>
<td>8.8</td>
</tr>
</tbody>
</table>

Note. VC = Verbal Comprehension factor; PR = Perceptual Reasoning factor; WM = Working Memory factor; PS = Processing Speed factor; b = loading of subtest on factor; h² = communality; u² = uniqueness; SI = Similarities; VC = Vocabulary; IN = Information; CO = Comprehension; BD = Block Design; MR = Matrix Reasoning; VP = Visual Puzzles; FW = Figure Weights; PCm = Picture Completion; DS = Digit Span; AR = Arithmetic; LN = Letter-Number Sequencing; SS = Symbol Search; CD = Coding; CA = Cancellation. Bold type indicates coefficients and variance estimates consistent with the theoretically proposed factor.
EFA) cannot fully address test score validity (Canivez et al., 2009). Factors capture reliable common variance but not all common variance is scientifically significant in terms of relationships with meaningful external criteria (Lubinski & Dawes, 1992). Furthermore, because latent constructs from CFA are not directly observable and latent construct scores are difficult to calculate and not readily available, they offer no direct practical clinical applications (Oh, Glutting, Watkins, Youngstrom, & McDermott, 2004). Consequently, additional methods, such as incremental predictive validity and diagnostic utility, are required to assess the relative merit of general versus narrow WAIS-IV factors.

The WAIS-IV Technical and Interpretive Manual presented correlations between WAIS-IV and the Wechsler Individual Achievement Test-Second Edition (WIAT-II; Psychological Corporation, 2002) for 93 Participants 16–19 years old where the WAIS-IV was administered first and WIAT-II administered 0 to 60 days later (M = 11 days). The WAIS-IV FSIQ had the highest correlations (with few exceptions) with WIAT-II composite (and subtest) scores ranging from .65 to .88 for the composite scores. However, examination of incremental predictive validity (Hunsley, 2003; Hunsley & Meyer, 2003) was not reported and would be needed to demonstrate that first-order factor scores provide important prediction of academic achievement beyond that predicted by the second-order Full Scale score. Previous incremental predictive validity studies with the WISC–III (Glutting et al., 1997) and WISC-IV (Glutting et al., 2006) were not favorable for factor index scores but at present there are no such studies of the WAIS-IV. If the small portions of apportioned variance to the WAIS-IV first-order factors observed in the present Schmid and Leiman (1957) analyses are able to account for meaningful portions of achievement variance beyond the second-order g factor, then there may be some utility of WAIS-IV factor scores in predicting achievement. Additional studies of incremental validity should examine how first- and second-order scores relate to other external criteria such as diagnosis and job training/performance. It is possible that WAIS-IV factor index scores will offer important prediction and classification utility beyond the Full Scale score. However, until evidence of incremental predictive validity or diagnostic utility is obtained, interpretation of WAIS-IV scores should primarily focus on the Full Scale score and caution should be exercised if moving to interpretations of subtest and index scores.

References


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


