Structural Validity of the WAIS-III Among Postsecondary Students

Marley W. Watkins  
James M. Kuterbach  
Rebecca J. Morgan  
Julie L. FitzGerald  
Rachel M. Neuhard  
April G. Arthur  
Leah B. Bucknavage  
The Pennsylvania State University

Abstract

The recent influx of students with disabilities into postsecondary education has generated a concomitant increase in the demand for psychoeducational assessments that include a measure of cognitive ability, either to identify ability-achievement discrepancies or to rule out alternate or comorbid diagnoses. The most commonly recommended cognitive ability measure for adults is the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III). However, evidence regarding the psychometric fitness of the WAIS-III for postsecondary assessments is needed. Of particular interest is its structural validity among these students. This study applied exploratory factor analysis to the WAIS-III scores of 183 students at a large Mid-Atlantic university who were referred for determination of postsecondary disability eligibility. The same four-factor model proposed by Wechsler (1997) for the general population was also appropriate for these students. Thus, these results support the use of the WAIS-III with postsecondary students with suspected disabilities.

Although students with disabilities attain postsecondary education at rates lower than their peers without disabilities, they are increasingly entering colleges and universities. For example, around 3% of college students reported disabilities in 1978 (Henderson, 1992), but this rate had risen to 6% in 1996 and 9% in 2002 (NCES, 2000, 2003). Based on the latest compilation of self-reports, there are currently more than 1.5 million postsecondary students with disabilities (NCES, 2003). This influx of students with disabilities into postsecondary education has generated a concomitant increase in the demand for psychoeducational assessments to substantiate the presence of disabilities. While eligibility for special educational services in secondary schools is governed by the Individuals with Disability Education Act (IDEA), postsecondary eligibility is guided by Section 504 of the Vocational Rehabilitation Act of 1973 and the Americans with Disabilities Act (ADA) of 1990. Students eligible for special educational services under IDEA will not necessarily be eligible for postsecondary services (Hatzes, Reiff, & Bramel, 2002). Additionally, Section 504 and the ADA require students to provide postsecondary institutions with documentation to support their requests for access to special accommodations (Hatzes et al., 2002).

Recognizing the inchoate state of postsecondary disability assessments, the Association on Higher Education and Disability (AHEAD) published guidelines for the documentation of a learning disability in adolescents and adults in 1997. Although specifically related to identification of learning disabilities, the AHEAD guidelines established a precedent for evaluating eligibility for students with other disabilities (Sitlington, 2003). Subsequently, many postsecondary institutions adopted local guidelines for documentation of learning disabilities (LD), attention-deficit hyperactivity disorder (ADHD), and psychiatric conditions (McGuire, 1998).

Most of these postsecondary disability guidelines recommended inclusion of a measure of cognitive ability, either to identify ability-achievement discrepancies or to rule out alternate or comorbid diagnoses (Gordon & Murphy, 1998). For example, the AHEAD (1997) guidelines specified that a “complete intellectual assessment” be conducted and, in an appendix, nominated several individual IQ tests for use with postsecondary students. These “approved” instruments included the Wechsler, Stanford-Binet, and Woodcock-Johnson scales. Other guidelines identified “preferred” tests. For example, the University of Connecticut guidelines stated that the “Wechsler Adult Intelligence Scale” was the preferred cognitive ability instrument (McGuire, 1998).
Consonant with the cognitive test preferences found in disability guidelines, a survey of postsecondary disability service providers found that the Wechsler adult scale was the most frequently used ability test with postsecondary students (Ofiesh & McAfee, 2000). In fact, the Wechsler scales are generally popular among psychologists. For example, surveys of test usage have found them to be the most widely used with adolescents, among clinical psychologists, and among neuropsychologists (Archer & Newsom, 2000; Belter & Piotrowski, 2001; Lees-Haley, Smith, Williams, & Dunn, 1996). Thus, the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997) is currently the most popular test of cognitive ability for adolescents and adults.

Given the wide spread use of the WAIS-III in postsecondary disability evaluations, professional standards mandate evidence regarding that test’s psychometric fitness for those purposes (AERA, APA, & NCME, 1999). Of the many forms of psychometric evidence, probably the most critical is construct validity (Messick, 2000). Although many forms of evidence relate to construct validity (i.e., test content, external relationships, etc.), test structure evidence is especially important. That is, does empirical analysis of a test’s components support the structure proposed by the test’s developer across a variety of test takers? If a test’s structure varies across groups of examinees, it may be measuring different constructs for each group and its scores cannot, therefore, be used interchangeably across groups. In the case of the WAIS-III, evidence of structural validity would assure users that it is measuring cognitive abilities with fidelity across a variety of examinees and, consequently, that its scores can be interpreted similarly across groups.

Typically, test structure is evaluated using factor analysis (Benson, 1998), a family of multivariate statistical methods that analyzes the patterns of correlations among a test’s subtests. By mathematically combining the subtests that correlate highly with each other into a single construct called a factor, factor analysis simplifies and clarifies the structure of the test. That is, a given test is reduced from numerous intercorrelated subtests to a smaller number of independent factors that reflect the latent constructs theoretically responsible for causing the covariation among the subtests. The resultant factor structure is the most parsimonious explanation for the observed relationships among subtests. Consequently, the stability of factor structures across groups of examinees provides evidence regarding the structural validity of the test.

For the WAIS-III, the application of factor analysis began with its normative sample (Wechsler, 1997). Four factors were found to best describe the intercorrelations among 13 WAIS-III subtests: (a) Verbal Comprehension was made up of Vocabulary, Similarities, Information, and Comprehension subtests; (b) Perceptual Organization was formed by the Picture Completion, Block Design, Matrix Reasoning, and Picture Arrangement subtests; (c) Working Memory was composed of Arithmetic, Digit Span, and Letter-Number Sequencing subtests; and Processing Speed was loaded by the Digit Symbol-Coding and Symbol Search subtests. These results were subsequently replicated in the Canadian normative sample (Saklofske, Hildebrand, & Gorsuch, 2000) and also in an independent analysis of the U.S. normative sample (Taub, 2001). The factor intercorrelations ranged from .57 to .80 in the U.S. normative sample, suggesting that a higher-order factor might explain their covariation. This higher-order general ability factor (i.e., general intelligence or g) was confirmed in another reanalysis of the WAIS-III normative sample (Arnau & Thompson, 2000).

Conflicting results were reported from two other reanalyses of the WAIS-III normative sample (Kaufman, Lichtenberger, & McLean, 2001; Ward, Ryan, & Axelrod, 2000). These researchers suggested that two or three factors might more parsimoniously explain the covariation of WAIS-III subtests. A factor analysis of WAIS-III scores of 152 Veteran’s Administration medical center patients also revealed some difficulties with the four-factor model (Ryan & Paolo, 2001). Specifically, the Working Memory factor was not congruent with normative results because it did not contain the Arithmetic subtest. Among a sample of 120 adults with schizophrenia, however, it was the Picture Arrangement test that did not load on the Perceptual Organization factor (Dickinson, Iannone, & Gold, 2002) while the three other factors were congruent with the normative sample. In contrast to these minor variations in factor structure, six factors were identified in a factor analysis of the WAIS-III scores of 328 medical patients tested for a neuropathological condition (Burton, Ryan, Axelrod, & Schellenberger, 2002).

Although a majority of factor analyses of the WAIS-III have favored a four-factor solution, alternative solutions have ranged from two to six factors. Relatively consistent results were obtained from analyses of the WAIS-III normative sample, but more variable solutions were obtained from clinical samples. As noted by Tulsky and Price (2003), “because these tests are often used with clinical populations, it is important to ascertain whether the factor structure … will be supported in various clinical populations” (p. 161).
No study has analyzed the structure of the WAIS-III among postsecondary students with suspected disabilities. Consequently, the present study applied factor-analytic methods to the WAIS-III scores of postsecondary students referred for psychoeducational evaluation.

Method

Participants

One hundred and eighty-three students (103 male and 80 female) served as participants. The ethnic background of 36% of the participants was not reported, but those who were identified were primarily of White ethnic origin (88%). The participants’ ages ranged from 16 to 46 years, with a mean of 21.1 years and standard deviation of 4.6 years. Participants were primarily enrolled in undergraduate courses (70%), but 4% were graduate students and 26% were evaluated during their final year of high school. A diagnosis was not reported for 26% of the sample; those students either did not meet the criteria or the diagnostic information was not included in the file. The remainder of the sample were identified by at least one psychiatric diagnosis: 51% with a learning disorder, 20% with ADHD, and 3% with affective disorders.

Instrument

The WAIS-III is an individually administered measure of intellectual functioning designed to assess adolescents and adults from ages 16 to 89. Its standardization sample included 2,450 individuals stratified on age, sex, education level, and geographic region according to 1995 census data. The WAIS-III contains a total of 14 subtests, but only 13 (Vocabulary, Information, Similarities, Comprehension, Block Design, Matrix Reasoning, Picture Completion, Picture Arrangement, Symbol Search, Coding, Arithmetic, Digit Span, and Letter-Number Sequencing) are necessary to compute the four factor-based index scores (i.e., Verbal Comprehension, Perceptual Organization, Working Memory, and Processing Speed). Letter-Number Sequencing and Symbol Search are not required to compute Verbal (VIQ), Performance (PIQ), and Full Scale (FSIQ) scores.

Subtest scores are standardized to a mean of 10 and a standard deviation of 3. IQ and factor indexes have a mean of 100 and standard deviation of 15. Extensive reliability and validity data are provided by Wechsler (1997). In general, the instrument demonstrated high reliability coefficients for IQ and factor index scores and strong relationships with other measures of ability and achievement. Additional, independent evidence of reliability and validity has also been reported (Blake & Impara, 2001; Groth-Marnat, 2003).

Procedure

The records of the Office of Disability Services and the school psychology clinic at a large Mid-Atlantic university were reviewed. The Office of Disability Services accepts evidence from many sources for determining eligibility, including high schools, clinics, and private evaluators. Such evaluations must be performed by an appropriately credentialed professional. The Office of Disability Services itself does not provide evaluative services. In contrast, the independent school psychology clinic provides assessments by doctoral-level school psychology students, supervised by doctoral-level certified school psychologists who may also be licensed psychologists. All 276 students with a referral for determination of postsecondary disability eligibility who received a WAIS-III as part of their psychoeducational evaluation were initially included in the sample. However, 93 students were missing at least one of the 13 WAIS-III subtests. Consequently, the final sample consisted of the 183 participants with complete WAIS-III data.

Statistical Analyses

Analyses were guided by the best practice suggestions of Fabrigar, Wegener, MacCallum, and Strahan (1999), Preacher and MacCallum (2003), and Russell (2002), among others. Given the uncertainty surrounding the structure of the WAIS-III, exploratory rather than confirmatory factor analysis was chosen (Browne, 2001). Common factor analysis was selected over principal-components analysis because the goal of the study was to identify the latent structure of the WAIS-III (Wegener & Fabrigar, 2000). Additionally, common factor analysis may produce more accurate estimates of population parameters than principal-components analysis (Widaman, 1993). Given its relative tolerance of multivariate non-normality and its superior recovery of weak factors, principal axis extraction was used (Briggs & MacCallum, 2003). Communalities were initially estimated by squared multiple correlations and were iterated twice to produce final communality estimates (Gorsuch, 2003). Following the advice of Velicer, Eaton, and Fava (2000), minimum average partials (MAP; Velicer, 1976) and parallel analysis (Horn, 1965), supplemented by a visual scree test (Cattell, 1966), were used to determine the number of factors to retain for rotation. For both theoretical and empirical reasons, it was assumed that factors would be moderately correlated (Wechsler, 1997). Thus, a Promax rotation with a $k$ value...
of 4 was selected (Tataryn, Wood, & Gorsuch, 1999). Loadings > .38 were predetermined to be salient so as to retain only those that were both statistically (p < .01) and practically significant (Stevens, 2002).

Results

As reported in Table 1, participants’ mean WAIS-III IQ and index scores were slightly higher and somewhat less variable than those of the normative sample. Score distributions appeared to be relatively normal, with .51 the largest skew and .89 the largest kurtosis. Multiple t-tests were conducted to determine if IQ and index scores systematically differed for male and female students. A Bonferroni correction was applied in order to maintain an overall alpha level of .05. Male participants scored significantly higher than females on the VIQ, Verbal Comprehension, Working Memory, and FSIQ composite scores, but age was not significantly related to WAIS-III composite scores (r = -.06 to +.05).

Results from Bartlett’s Test of Sphericity (Bartlett, 1954) indicated that the correlation matrix was not random (x² = 1079.3; df = 78; p < .001). The Kaiser-Meyer-Olkin (KMO; Kaiser, 1974) statistic was .83, well above the minimum standard suggested by Kline (1994). Measures of sampling adequacy for each variable were also within reasonable limits. Thus, the correlation matrix was appropriate for factor analysis.

The scree and MAP criteria suggested that four factors be retained, but parallel analysis recommended only three factors. Given that it is better to over-factor than under-factor (Wood, Tataryn, & Gorsuch, 1996), four factors were extracted. The resulting solution was examined for both substantive and statistical suitability. Fit appeared to be excellent, accounting for 61% of the total variance and leaving only 11% of the residual matrix > .05. As reflected in Table 2, pattern coefficients clearly identified the four factors suggested by Wechsler (1997) with the exception of the Arithmetic subtest, which loaded on the Verbal Comprehension rather than the

<table>
<thead>
<tr>
<th>IQ/Index</th>
<th>Males (n = 103)</th>
<th>Females (n = 80)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIQ*</td>
<td>110.0 (12.2)</td>
<td>103.2 (12.4)</td>
<td>107.0 (12.7)</td>
</tr>
<tr>
<td>PIQ</td>
<td>107.6 (13.6)</td>
<td>103.7 (13.5)</td>
<td>105.9 (13.7)</td>
</tr>
<tr>
<td>FSIQ*</td>
<td>109.7 (12.6)</td>
<td>103.7 (12.5)</td>
<td>107.1 (12.9)</td>
</tr>
<tr>
<td>VC*</td>
<td>111.3 (13.1)</td>
<td>105.5 (11.4)</td>
<td>108.8 (12.7)</td>
</tr>
<tr>
<td>PO</td>
<td>110.7 (14.6)</td>
<td>105.8 (14.1)</td>
<td>108.6 (14.5)</td>
</tr>
<tr>
<td>WM*</td>
<td>102.4 (13.3)</td>
<td>96.6 (12.9)</td>
<td>99.9 (13.4)</td>
</tr>
<tr>
<td>PS</td>
<td>97.6 (14.3)</td>
<td>98.2 (13.1)</td>
<td>97.8 (13.7)</td>
</tr>
</tbody>
</table>

Note. VIQ = Verbal IQ, PIQ = Performance IQ, FSIQ = Full Scale IQ, VC = Verbal Comprehension Index, PO = Perceptual Organization Index, WM = Working Memory Index, and PS = Perceptual Speed Index.

* p < .05 experimentwise (.007 for each comparison).
Working Memory factor. Factor intercorrelations ranged from .38 between Verbal Comprehension and Processing Speed to .67 between Verbal Comprehension and Perceptual Organization, suggesting a general, second-order factor. All subtests demonstrated moderate to substantial loadings on the first unrotated principal factor (see Table 2), another reflection of the influence of an overarching generalability factor (Arnau & Thompson, 2000; Carroll, 2003). Additionally, the four-factor solution was robust across extraction (Principal Components, Maximum Likelihood) and rotation (Varimax, Oblimin) methods.

Three factors were also extracted and rotated for a statistical and substantive comparison to the four-factor solution. This three-factor solution accounted for 53% of the total variance, but left 33% of the residual matrix \( \geq 1.05 \). As reported in Table 3, the Verbal Comprehension and Perceptual Organization factors collapsed into a combined factor with the Arithmetic subtest still loading on the Verbal Comprehension factor. Factor intercorrelations ranged from .39 for factors I and III to .48 for factors I and II. Although parsimonious, this solution was marked by two major flaws: (a) the large number of sizeable residuals suggested that additional factors should be extracted (Gorsuch, 2003), and (b) it was not congruent with the bulk of published WAIS-III factor analyses.

Given these statistical and substantive considerations, the four-factor solution was accepted as the most adequate for this sample of postsecondary students. To

### Table 2

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Unrotated first factor</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>Communality</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO</td>
<td>.76</td>
<td>.86</td>
<td>-.14</td>
<td>.12</td>
<td>.03</td>
<td>.70</td>
</tr>
<tr>
<td>SM</td>
<td>.70</td>
<td>.69</td>
<td>.20</td>
<td>-.11</td>
<td>-.02</td>
<td>.60</td>
</tr>
<tr>
<td>IN</td>
<td>.70</td>
<td>.75</td>
<td>.00</td>
<td>.09</td>
<td>-.07</td>
<td>.59</td>
</tr>
<tr>
<td>CM</td>
<td>.69</td>
<td>.91</td>
<td>-.03</td>
<td>-.16</td>
<td>-.03</td>
<td>.69</td>
</tr>
<tr>
<td>PC</td>
<td>.55</td>
<td>.22</td>
<td>.40</td>
<td>-.14</td>
<td>.18</td>
<td>.36</td>
</tr>
<tr>
<td>BD</td>
<td>.63</td>
<td>.11</td>
<td>.42</td>
<td>.08</td>
<td>.21</td>
<td>.44</td>
</tr>
<tr>
<td>MR</td>
<td>.66</td>
<td>-.05</td>
<td>.79</td>
<td>.08</td>
<td>-.02</td>
<td>.63</td>
</tr>
<tr>
<td>PA</td>
<td>.52</td>
<td>-.02</td>
<td>.77</td>
<td>-.03</td>
<td>-.15</td>
<td>.49</td>
</tr>
<tr>
<td>AR</td>
<td>.70</td>
<td>.38</td>
<td>.11</td>
<td>.34</td>
<td>.05</td>
<td>.51</td>
</tr>
<tr>
<td>DS</td>
<td>.51</td>
<td>-.07</td>
<td>.08</td>
<td>.78</td>
<td>.00</td>
<td>.62</td>
</tr>
<tr>
<td>LN</td>
<td>.47</td>
<td>.00</td>
<td>-.07</td>
<td>.83</td>
<td>-.03</td>
<td>.63</td>
</tr>
<tr>
<td>CD</td>
<td>.40</td>
<td>-.04</td>
<td>-.02</td>
<td>-.05</td>
<td>.83</td>
<td>.63</td>
</tr>
<tr>
<td>SS</td>
<td>.48</td>
<td>-.02</td>
<td>-.06</td>
<td>.04</td>
<td>.87</td>
<td>.73</td>
</tr>
</tbody>
</table>

**Note.** VO = Vocabulary, SM = Similarities, IN = Information, CM = Comprehension, PC = Picture Completion, BD = Block Design, MR = Matrix Reasoning, PA = Picture Arrangement, AR = Arithmetic, DS = Digit Span, LN = Letter-Number Sequencing, CD = Digit Symbol-Coding, and SS = Symbol Search. Salient coefficients (\( \geq .38 \)) are indicated in bold.
test that conclusion against the WAIS-III normative sample, the congruence coefficient ($r_c$), an index of factorial similarity, was calculated for each factor. Jensen (1998) reported than an $r_c > +.90$ indicates “a high degree of factor similarity; a value greater than +.95 is generally interpreted as practical identity of the factors” (p. 99). Based upon these guidelines, the Perceptual Organization factor from this sample of postsecondary students was found to display a high degree of factor similarity with the WAIS-III normative sample ($r_c = .90$) and the Verbal Comprehension, Working Memory, and Processing Speed factors were practically identical ($r_c = .97, .95, \text{and} .95, \text{respectively}$) to the WAIS-III normative sample.

Discussion

Factor analyses of the WAIS-III scores of postsecondary students evaluated for determination of eligibility for disability services indicated that the four-factor model proposed by Wechsler (1997) for the general population was also appropriate for these students. Specifically, the WAIS-III was best summarized by four factors: Verbal Comprehension, Perceptual Organization, Working Memory, and Processing Speed. This supports the conclusions from previous factor analyses of the U.S. and Canadian WAIS-III normative samples (Saklofske et al., 2000; Taub, 2001) and suggests that the structure of the WAIS-III is similar across disparate populations (Ryan & Paolo, 2001). More specifically, these results support the use of the WAIS-III with postsecondary students with suspected disabilities.

Table 3

Structure of the WAIS-III for Principal-Axis Extraction and Promax Rotation of Three Factors Among 183 Postsecondary Students Tested for a Disability

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Unrotated (\text{first factor})</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>Communality</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO</td>
<td>.75</td>
<td>.75</td>
<td>.06</td>
<td>-.01</td>
<td>.60</td>
</tr>
<tr>
<td>SM</td>
<td>.71</td>
<td>.84</td>
<td>-.12</td>
<td>-.03</td>
<td>.61</td>
</tr>
<tr>
<td>IN</td>
<td>.70</td>
<td>.75</td>
<td>.05</td>
<td>-.09</td>
<td>.56</td>
</tr>
<tr>
<td>CM</td>
<td>.68</td>
<td>.89</td>
<td>-.20</td>
<td>-.06</td>
<td>.63</td>
</tr>
<tr>
<td>PC</td>
<td>.54</td>
<td>.52</td>
<td>-.09</td>
<td>.19</td>
<td>.33</td>
</tr>
<tr>
<td>BD</td>
<td>.63</td>
<td>.43</td>
<td>.14</td>
<td>.22</td>
<td>.40</td>
</tr>
<tr>
<td>MR</td>
<td>.65</td>
<td>.52</td>
<td>.19</td>
<td>.04</td>
<td>.42</td>
</tr>
<tr>
<td>PA</td>
<td>.50</td>
<td>.51</td>
<td>.08</td>
<td>-.08</td>
<td>.28</td>
</tr>
<tr>
<td>AR</td>
<td>.70</td>
<td>.47</td>
<td>.33</td>
<td>.04</td>
<td>.51</td>
</tr>
<tr>
<td>DS</td>
<td>.51</td>
<td>-.03</td>
<td>.80</td>
<td>.00</td>
<td>.62</td>
</tr>
<tr>
<td>LN</td>
<td>.47</td>
<td>-.06</td>
<td>.82</td>
<td>-.04</td>
<td>.61</td>
</tr>
<tr>
<td>CD</td>
<td>.41</td>
<td>-.03</td>
<td>-.06</td>
<td>.82</td>
<td>.63</td>
</tr>
<tr>
<td>SS</td>
<td>.48</td>
<td>-.03</td>
<td>.03</td>
<td>.85</td>
<td>.72</td>
</tr>
</tbody>
</table>

Note. VO = Vocabulary, SM = Similarities, IN = Information, CM = Comprehension, PC = Picture Completion, BD = Block Design, MR = Matrix Reasoning, PA = Picture Arrangement, AR = Arithmetic, DS = Digit Span, LN = Letter-Number Sequencing, CD = Digit Symbol-Coding, and SS = Symbol Search. Salient coefficients (\(\geq .38\)) are indicated in bold.
However, these results must be considered within the context of this study’s limitations. Specifically, the sample of postsecondary students involved in this study was from a single university and was relatively homogeneous in terms of ethnicity and age. Results may not extend to diverse students from other postsecondary institutions. Consequently, this study should be replicated using different samples of students to ensure that the results are generalizable. Additionally, research is needed to establish concurrent and predictive validity of the WAIS-III among postsecondary students. Other forms of construct validity, such as convergent and divergent validity, should also be examined.

This study of the structural validity of the WAIS-III found support for the same four-factor model in postsecondary students that has been proposed for the general population, supporting the use of this instrument for disability evaluations at the postsecondary level. However, as noted by Benson (1998), “one study does not validate or fail to validate the scores from a test. Numerous studies may be required, utilizing different approaches, different samples, and different populations to build a body of evidence that supports, or fails to support, the validity of the scores derived from a test” (p. 10). Nevertheless, the current results constitute preliminary evidence regarding the structural validity of the WAIS-III among postsecondary students referred for an evaluation.


