Psychological Assessment
2003, Vol. 15, No. 2, 149–162

1040-3590/03/$12.00 DOI: 10.1037/1040-3590.15.2.149

The Joint WAIS–III and WMS–III Factor Structure: Development and Cross-Validation of a Six-Factor Model of Cognitive Functioning

David S. Tulsky
Kessler Medical Rehabilitation Research and Education Corporation and University of Medicine and Dentistry of New Jersey—New Jersey Medical School

Larry R. Price
Southwest Texas State University

During the standardization of the Wechsler Adult Intelligence Scale (3rd ed.; WAIS–III) and the Wechsler Memory Scale (3rd ed.; WMS–III) the participants in the normative study completed both scales. This "co-norming" methodology set the stage for full integration of the 2 tests and the development of an expanded structure of cognitive functioning. Until now, however, the WAIS–III and WMS–III had not been examined together in a factor analytic study. This article presents a series of confirmatory factor analyses to determine the joint WAIS–III and WMS–III factor structure. Using a structural equation modeling approach, a 6-factor model that included verbal, perceptual, processing speed, working memory, auditory memory, and visual memory constructs provided the best model fit to the data. Allowing select subtests to load simultaneously on 2 factors improved model fit and indicated that some subtests are multifaceted. The results were then replicated in a large cross-validation sample (N = 858).

When a four-factor model of intelligence was introduced in the Wechsler Adult Intelligence Scale, Third Edition (WAIS–III; Wechsler, 1997a), the structure of intellect was better delineated and described than it had been in previous editions (e.g., Wechsler Bellevue Form I, Wechsler, 1939; WAIS, Wechsler, 1955; WAIS—Revised [WAIS–R], Wechsler, 1981). Recent reports examining the clinical utility of the four-factor structure of the WAIS–III have demonstrated that individuals with neuropsychological deficits perform worse on some of the factor scores than on others and that their performance often deviates from the performance patterns of individuals without such deficits (Hawkins, 1998; Martin, Donders, & Thompson, 2000). Furthermore, recent research has suggested that more multifaceted models of intelligence have greater utility when describing cognitive functioning than do simple structures of verbal, performance, and full-scale IQ (e.g., see Carroll, 1993; Horn & Noll, 1997; Smith et al., 1992). On the basis of these reports, Tulsky, Ivnik, Price, and Wilkins (2003) argued that the traditional structures of verbal, performance, and full-scale IQs have become outdated and should be replaced with more contemporary, refined indices of cognitive functioning.

David S. Tulsky, Kessler Medical Rehabilitation Research and Education Corporation, West Orange, New Jersey, and Department of Physical Medicine and Rehabilitation, University of Medicine and Dentistry of New Jersey—New Jersey Medical School; Larry R. Price, Department of Education, Southwest Texas State University.

We thank Glenn Larrabee and Scott R. Millis for their critical review and thoughtful comments and suggestions for changes to the manuscript. Scott R. Millis provided invaluable statistical consultation. We also thank The Psychological Corporation for granting access to the WAIS–III and WMS–III standardization data sets.

Correspondence concerning this article should be addressed to David S. Tulsky, Kessler Medical Rehabilitation Research and Education Corporation, 1199 Pleasant Valley Way, West Orange, New Jersey 07052. E-mail: dtulsky@kmrrec.org

During the standardization of the WAIS–III and the Wechsler Memory Scale, Third Edition (WMS–III; Wechsler, 1997b) a "co-norming" methodology was used, and all of the WMS–III standardization research participants completed all of the subtests of both the WAIS–III and the WMS–III. This sampling methodology established a large, representative sample of individuals who had completed both tests. This sample offers great potential to investigate the relationship between the WAIS–III and the WMS–III and to identify the cognitive abilities that are measured by joint administration of these two scales. Despite this opportunity, the WAIS–III and WMS–III were developed and published as separate instruments, and there has not been a formal study to investigate the joint structure between these two instruments. This has been cited as the one "significant omission" in the joint WAIS–III and WMS–III publications (Larrabee, 1999, p. 477).

During the development of the WAIS–III and WMS–III, emphasis was placed on preserving the historical distinction between intelligence and memory (Tulsky, Chiaravalloti, Palmer, & Chelune, 2003; Tulsky, Saklofske, & Zhu, 2003). The distinction between memory and intelligence in cognitive testing dates back to Binet and Simon (1905/1916), and the constructs have been treated independently in clinical settings through most of the 20th century (see Tulsky, Saklofske, & Ricker, 2003).

The purpose of this article is to investigate the underlying structure of the subtests that are represented on the WAIS–III and WMS–III so that an expanded model of cognitive functioning can be advanced for clinical use. Developing a single battery measuring an integrated model of cognitive functioning across the WAIS–III and WMS–III is relevant for practitioners using both scales in clinical practice. This article provides the empirical rationale for the creation of such a clinical battery. The actual development of the battery along with new normative information is described in detail in a companion chapter (see Tulsky, Ivnik, et al., 2003).

Background

Historically, joint factor analytic studies of the WAIS–III and WMS–III have been scarce. In the first study of its kind, Larrabee, Kane, and Schuck (1983) examined a sample of 256 individuals who had been referred for a neuropsychological evaluation and completed the WAIS and WMS as part of a larger assessment battery. They performed an exploratory maximum-likelihood factor analysis and determined that there were six factors: Perceptual Organization, Verbal Comprehension, Attention Concentration, Verbal Learning and Recall, Information and Orientation, and an extra factor that was uninterpretable (Larrabee et al., 1983). Ryan, Rosenberg, and Heilbronner (1984) extended this work and conducted factor analyses on the WAIS–R and WMS and the WAIS and WMS. In both cases, the authors found that there were four factors: Perceptual Organization, Attention Concentration, Verbal Comprehension, and Verbal Learning and Recall. Similar findings were obtained whether the WAIS or WAIS–R was used.

Four more recent joint studies have been performed with the WAIS–R and WMS–R. Leonberger, Nicks, Larrabee, and Goldfader (1992) found a five-factor model (Spatial Reasoning, Verbal Comprehension, Attention/Concentration, Memory, and Psychomotor Speed) when the WAIS–R and WMS–R were analyzed in a joint study including variables from the Halstead Reitan Neuropsychological Battery. Similarly, Larrabee and Curtiss (1992; see Larrabee, 2000) found support for a six-factor model in a study containing the WAIS–R and WMS as well as several other neuropsychological variables. They concluded that the WMS memory variables caused the identification of other factors in addition to verbal, perceptual, and attention. Smith et al. (1992) used confirmatory techniques to develop a shorter "core" battery for individuals within the Mayo Clinic's Older American Normative Studies. The authors investigated the factor structure of the WAIS–R, WMS–R, and Auditory Verbal Learning Test on a sample of 415 older adult examinees who had not been diagnosed with any neuropsychological problems (age range, 55–97; $M = 71.5$ years). Their analyses supported a five-factor solution, which guided their follow-up study and clinical practice.

Bowden, Carstairs, and Shores (1999) performed a confirmatory factor analysis of the Australian WAIS–R and WMS–R. In their study, they tested a sample of 399 young adults (18–34 years old) and found that a five-factor model composed of Verbal Comprehension, Perceptual Organization, Attention Concentration, Verbal Memory, and Visual Memory provided the best fit to the data. On the basis of their results, they found that the Digit Symbol subtest seemed to be related to Visual Memory. Furthermore, a model was specified and tested to see whether Digit Symbol would load on a separate factor; this analysis yielded results that were close, but not superior, to the five-factor model described above. They also found that second-order models that had more defined constructs of intelligence (e.g., verbal, performance, attention) under a global *g* factor and more defined constructs of memory (e.g., immediate and delayed) under a general memory factor could not be identified owing to nonconvergence of the models.

Until now, the WAIS–III and WMS–III have not been factor analyzed simultaneously in a study. Because the revised scales contain additional subtests, and are now the current versions of the scales, such analyses would help elucidate the relationship between the scales and could be used to develop new models of expanded domains of functioning for clinical use.

We performed three studies. Two developed and tested the joint factor structure of the subtests of the WAIS–III and WMS–III; the third cross-validated the structure in an independent sample. In Study 1, confirmatory factor analytic techniques (CFA) were used, with a structural equation modeling approach. Confirmatory methodology has unique advantages over exploratory analyses because measurement models are developed in an a priori fashion, and specific factor structures can be tested to see whether they "fit" the data. Previous research conducted on the joint structure, along with the theoretical suppositions about the possible factor structure, guided this research. Another advantage of modeling the joint WAIS–III and WMS–III structure through CFA is that both immediate and delayed variables can be included in the same model, by allowing their error terms to covary.

One limitation of CFA occurs when models are misspecified or alternative structures are not represented or tested. This is especially salient with the WAIS–III and WMS–III because the subtests are multifaceted and, in exploratory research, have loadings that tend to "split" over multiple factors (see Larrabee, 1999; Tulsky, Ledbetter, & Zhu, 1998). In this case, a variable might not load with its intended factor and a better fit to the model would be obtained if an alternate confirmatory model were specified. Therefore, Study 2 was conducted to investigate how the multifaceted subtests such as Picture Arrangement, Spatial Span, Visual Reproduction, and Arithmetic interact if allowed to split between two factors.

Study 3 was conducted to verify the adequacy of the structures described above by using cross-validation with an independent validity sample. Cross-validation allows one to compare and test model fit indices obtained in one analysis with those obtained in a second analysis using a different sample. In this case, the goal was to confirm that the factor structure developed in Study 1 could be replicated in a new sample in which model shrinkage was sure to occur. Furthermore, cross-validation provides a framework for verifying that the paths between the subtests and the factors can be adequately replicated and also that the magnitude of the subtest loadings are consistent across the samples.

The WMS–III standardization sample served as the calibration sample and was used to develop the model in Studies 1 and 2. A sample of 858 individuals who had completed both the WAIS–III and the WMS–III as part of oversampling efforts but were not included in the WMS–III standardization sample served as the validity sample to cross-validate the structure in Study 3.

Study 1: Determining the Number of Factors in a Joint Structure

Method

Subjects. The WMS–III standardization sample (weighted $N = 1{,}250$) was used for the analyses. The sample included individuals 16–89 years of age. The details of the WMS–III standardization sample have been reported in the *WAIS–III—WMS–III Technical Manual* (The Psychological Corporation, 1997), with some additional clarification provided by Tulsky and Ledbetter (2000), and are not repeated here. All participants were deemed cognitively normal and had to meet the inclusionary criteria for the WAIS–III—WMS–III study (see The Psychological Corporation, 1997).

Measures. Examinees completed all of the core and optional subtests from the WAIS–III and the WMS–III. The 26 subtests included in this analysis were Vocabulary, Information, Similarities, Comprehension, Block Design, Matrix Reasoning, Picture Completion, Picture Arrangement, Symbol Search, Digit Symbol—Coding, Arithmetic, Digit Span, Letter Number Sequencing, Spatial Span, Faces (Immediate and Delayed), Family Pictures (Immediate and Delayed), Visual Reproduction (Immediate and Delayed), Logical Memory (Immediate and Delayed), Verbal Paired Associates (Immediate and Delayed), Word List (Immediate and Delayed). As reported earlier, Letter Number Sequencing was administered only one time in the testing session (see Tulsky & Zhu, 2000), as part of the WMS–III battery.

Model development and analytic procedure. Models of cognitive function were developed on the basis of previous research with the Wechsler scales as well as theoretical hypotheses about what the subtests were purported to measure (see The Psychological Corporation, 1997). Given the publication of the four-factor WAIS–III structure and the addition of memory variables, a model incorporating Verbal, Performance, Working Memory, Processing Speed, and Memory was expected to have the best fit to the data. Moreover, it was hypothesized that Memory could be separated into auditory and visual and, possibly, immediate and delayed (see Price, Tulsky, Millis, & Weiss, 2002, for a new factor analytic study of the WMS–III). Therefore, a five-, six-, or eight-factor model was expected to have the best fit to the data. Such findings would support the utility of a multifactorial model of cognitive functioning by unifying the WAIS–III and WMS–III.

To explore the unified model, different measurement models were created, fit, and compared against one another using a competing-models approach. First, each measurement model was compared with a general, one-factor model. The model specifications are outlined in Table 1.

Next, successive measurement models were evaluated according to a variety of goodness-of-fit, model parsimony, and information–theoretic measures. The indices used in the WAIS–III and WMS–III confirmatory analyses (see The Psychological Corporation, 1997) were used here. These indices were selected because they are less sensitive to sample size or to the number of degrees of freedom (see Bollen, 1989; Bollen & Long, 1993; Hu & Bentler, 1999; Marsh, Balla, & McDonald, 1988; Tanaka, 1993). Thus, the chi-square index divided by degrees of freedom (χ^2/df) was used to calculate the Tucker–Lewis Index (TLI; Tucker & Lewis, 1973), which is a comparative fit index that makes adjustment for the number of degrees of freedom in the model. The goodness-of-fit index adjusted for degrees of freedom (AGFI; Jöreskog & Sörbom, 1993), was used, as was the root-mean-square residual (RMSR) index, a measure of the degree of reproduction of the covariance matrix from the model estimates. Testing for successive improvement in model fit involved moving from single- to multiple-factor models. Evaluation of the improvement in model fit was provided by the chi-square difference test along with review of modification indices. The TLI shows the comparative fit of each model to the one-factor model.

In addition to the fit measures used in the original WMS–III analyses (The Psychological Corporation, 1997), several others were used in the follow-up analyses conducted by Millis, Malina, Bowers, and Ricker (1999). These measures include the comparative fit index (CFI; Bentler, 1990), an index that compares the proposed model to a baseline model, with values close to 1 indicating a very good fit; the normed fit index (NFI; Bentler & Bonett, 1980), which is another hierarchical fit index that compares a proposed model fit against a baseline model, with values above .90 indicating a good fit; TLI (Hu & Bentler, 1999), a nonnormed index that compensates for the effect of model complexity, in which cutoff value of around .90 indicates a good fit to the model; and the root-mean-square error of approximation (RMSEA; Browne & Cudeck, 1993; Hu & Bentler, 1999), which compensates for model complexity by dividing the *F* statistic by the number of degrees of freedom. Whereas exact fit to the model would be indicated by an RMSEA of 0, an RMSEA of less than .08 indicates a reasonable model fit and a value of .05 indicates a very close fit in relation to degrees of freedom. Finally, given the number of different models being tested, the Bayes information criterion (BIC) is an information–theoretic measure that allows the investigator to assess the amount of support or evidence for the various models. When comparing competing models that have reasonable fit to the data, BIC tends to favor simpler, more parsimonious models than does a traditional sequential *p*-value approach (Raftery, 1993). A BIC difference of 5 points offers strong evidence of model fit, and any value above 10 offers very near conclusive evidence (Raftery, 1993).

Results and Discussion

Confirmatory analyses were conducted with the AMOS 4.0 program (Arbuckle & Wothke, 1999). The results of each of these goodness-of-fit analyses are presented in Table 2. The results indicate that out of the models presented, Model 8, which is a six-factor model, has the best fit to the data. There is improvement on each of the fit indices over the other models. The AGFI is .912, the CFI is .954, and the NFI is .941, all of which are acceptable values and indicate good fit of the model to the data. The RMSR is .365, which is also an improvement over other models presented. The RMSEA is .051, which indicates that the model fit adjusted for complexity is very good. The TLI improves as parameters are added until six factors are present and the value is .946. Finally, the BIC difference fit statistics have been included in Table 3. The table shows the difference in BIC between two competing models. The results show a steady improvement and, using Raftery's (1993) criteria, provide near conclusive evidence that the six-factor model provides the most parsimonious solution of the models tested here. Cumulatively, these fit statistics support a six-factor model of cognitive functioning as measured by the WAIS–III and WMS–III. A representation of this model is presented in Figure 1.

All of the models that specify identification of separate immediate and delayed memory factors, either as distinct factors (i.e., Model 9) or nested within a hierarchical structure (i.e., Model 10), yielded inadmissible statistics above 1. These results are indicative of model specification errors. These specification problems again mirror the findings originally obtained by Millis et al. (1999) and the reanalysis of the actual standardization data performed by Price et al. (2002) by demonstrating that separate immediate and delayed structures cannot be obtained within the same measurement model. The findings here also agree with previous research with earlier editions of the Wechsler scales (most recently by Bowden et al., 1999, who were also unable to fit hierarchical models with separate immediate and delayed variables within their WAIS–R/WMS–R data set). How the model-specification errors might have occurred is best described by Millis et al. (1999), who pointed out that these models have specified variables on a factor that have small correlations (e.g., Faces Immediate and Family Pictures Immediate, $r = .31$; Faces Delayed and Family Pictures Delayed, $r = .28$) whereas the correlations between the immediate and delayed variables are highly correlated (e.g., Faces Immediate and Faces Delayed, $r = .67$; Family Pictures Immediate and Family Pictures Delayed, $r = .91$) and are placed on different factors. These findings make it difficult to evaluate whether the immediate and delayed breakdown provides additional benefit to the model. As it stands, this distinction is unsupported by these factor analytic results.

Table 4 provides the factor loadings obtained for each subtest. In general, most of the subtests have high loadings on their respective factors. The notable exception is the Faces subtest, which has unacceptably low loadings, indicating that it is somewhat different

Table 1
Model Specification for Confirmatory Analyses

Model and factors	Observed variable
Model 1 (1 factor)	All 26 subtests on a general factor
Model 2 (2 factors)	
Factor 1	All verbal subtests from WAIS–III
	All auditory subtests from WMS–III
	Letter Number Sequencing
Factor 2	All performance subtests from WAIS–III
	All visual subtests from WMS–III
	Spatial Span from WMS–III
Model 3 (3 factors)	
Factor 1	All subtests from Verbal Comprehension index of WAIS–III
	Comprehension from WAIS–III
	All subtests from WMS–III Auditory Immediate and Delayed indices
	Word List from WMS–III
Factor 2	All subtests from Perceptual Organization index of WAIS–III
	All subtests from Visual Immediate index of WMS–III
	All subtests from Visual Delayed index of WMS–III
	Picture Arrangement from WAIS–III
	Visual Reproduction from WMS–III
Factor 3	All subtests from WAIS–III and WMS–III Working Memory
	All subtests from WAIS–III Processing Speed
Model 4 (3 factors)	
Factor 1	All verbal subtests from WAIS–III
Factor 2	All performance subtests from WAIS–III
	Spatial Span from WMS–III
Factor 3	All subtests from Visual Immediate index of WMS–III
	All subtests from Visual Delayed index of WMS–III
	Visual Reproduction from WMS–III
	All subtests from Auditory Immediate index of WMS–III
	All subtests from Auditory Delayed index of WMS–III
	Word List from WMS–III
Model 5 (4 factors)	
Factor 1	All verbal subtests from WAIS–III
Factor 2	All performance subtests from WAIS–III
	Spatial Span from WMS–III
Factor 3	All subtests from Auditory Immediate index of WMS–III
	All subtests from Auditory Delayed index of WMS–III
	Word List from WMS–III
Factor 4	All subtests from Visual Immediate index of WMS–III
	All subtests from Visual Delayed index of WMS–III
	Visual Reproduction from WMS–III
Model 6 (4 factors)	
Factor 1	All subtests from Verbal Comprehension index of WAIS–III
	Comprehension from WAIS–III
Factor 2	All subtests from Perceptual Organization index of WAIS–III
	Picture Arrangement from WAIS–III
Factor 3	All subtests from Visual Immediate index of WMS–III
	All subtests from Visual Delayed index of WMS–III
	Visual Reproduction from WMS–III
	All subtests from Auditory Immediate index of WMS–III
	All subtests from Auditory Delayed index of WMS–III
	Word List from WMS–III
Factor 4	All subtests from WAIS–III and WMS–III Working Memory
	All subtests from WAIS–III Processing Speed
Model 7 (5 factors)	
Factor 1	All subtests from Verbal Comprehension index of WAIS–III
	Comprehension from WAIS–III
Factor 2	All subtests from Perceptual Organization index of WAIS–III
	Picture Arrangement from WAIS–III
Factor 3	All subtests from Visual Immediate index of WMS–III
	All subtests from Visual Delayed index of WMS–III
	Visual Reproduction from WMS–III
	All subtests from Auditory Immediate index of WMS–III
	All subtests from Auditory Delayed index of WMS–III
	Word List from WMS–III
Factor 4	All subtests from WAIS–III and WMS–III Working Memory
Factor 5	All subtests from WAIS–III Processing Speed

Table 1 (*continued*)

Model and factors	Observed variable
Model 8 (6 factors)	
Factor 1	All subtests from Verbal Comprehension index of WAIS–III Comprehension from WAIS–III
Factor 2	All subtests from Perceptual Organization index of WAIS–III Picture Arrangement from WAIS–III
Factor 3	All subtests from Auditory Immediate index of WMS–III All subtests from Auditory Delayed index of WMS–III Word List from WMS–III
Factor 4	All subtests from Visual Immediate index of WMS–III All subtests from Visual Delayed index of WMS–III Visual Reproduction from WMS–III
Factor 5	All subtests from WAIS–III and WMS–III Working Memory
Factor 6	Processing Speed from WAIS–III
Model 9 (8 factors)	
Factor 1	All subtests from Verbal Comprehension index of WAIS–III Comprehension from WAIS–III
Factor 2	All subtests from Perceptual Organization index of WAIS–III Picture Arrangement from WAIS–III
Factor 3	All subtests from Auditory Immediate index of WMS–III Word List Immediate from WMS–III
Factor 4	All subtests from Auditory Delayed index of WMS–III Word List Delayed from WMS–III
Factor 5	All subtests from Visual Immediate index of WMS–III Visual Reproduction Immediate from WMS–III
Factor 6	All subtests from Visual Delayed index of WMS–III Visual Reproduction Delayed from WMS–III
Factor 7	All subtests from WAIS–III and WMS–III Working Memory
Factor 8	All subtests from WAIS–III Processing Speed
Model 10 (10 factors; hierarchical model)	
Factor 1	All subtests from Verbal Comprehension index of WAIS–III Comprehension from WAIS–III
Factor 2	All subtests from Perceptual Organization index of WAIS–III Picture Arrangement from WAIS–III
Factor 3	Hierarchical auditory factor
Factor 4	Subfactor under Factor 3 All subtests from Auditory Immediate index of WMS–III Word List Immediate from WMS–III
Factor 5	Subfactor under Factor 3 All subtests from Auditory Delayed index of WMS–III Word List Delayed from WMS–III
Factor 6	Hierarchical visual factor
Factor 7	Subfactor under Factor 6 All subtests from Visual Immediate index of WMS–III Visual Reproduction Immediate from WMS–III
Factor 8	Subfactor under Factor 6 All subtests from Visual Delayed index of WMS–III Visual Reproduction Delayed from WMS–III
Factor 9	All subtests from WAIS–III and WMS–III Working Memory
Factor 10	All subtests from WAIS–III Processing Speed

Note. WAIS–III = Wechsler Adult Intelligence Scale (3rd ed.); WMS–III = Wechsler Memory Scale (3rd ed.).

from the other Visual Memory subtests. Once again, this finding mirrors those reported previously by Millis et al. (1999). In the current study, the factor loadings for the Faces I and II subtests were .32 and .37, respectively—about half of the magnitude of other factor loadings, which result in extremely low R^2 values of .10 and .14, respectively. These low R^2 values confirm the results previously reported by Millis et al., which led those authors to question whether the composite scores adequately measure unified constructs of visual memory. Such a finding is not totally unexpected, as facial recognition appears to be a unique, "hardwired" ability (see McCarthy & Warrington, 1990). These results do indicate that the Faces I and II subtests are unrelated to the other variables defining the construct that has been labeled visual memory and have led to the development of alternate index scores (see Tulsky, Ivnik, et al., 2003).

Also, although there appears to be better model fit with a six-factor solution, this might not be the best model. Individual subtests may have a higher loading on an alternative factor or a split loading between two factors. Hence, it is possible that restricting the loading of a subtest on a single factor was not appropriate or that the loading of a particular subtest may have been restricted to the wrong factor. In Study 2 we tested these

Table 2
Confirmatory Factor Analyses With WAIS–III and WMS–III Subtests From Calibration Sample

	Goodness-of-fit index								
Model	χ^2	*df*	χ^2/df	ΔNFI	CFI	TLI	AGFI	RMSR	RMSEA
1 (1 factor)	3,127.02	293	10.7	.846	.859	.843	.757	.673	.088
2 (2 factors)	2,542.50	292	8.7	.875	.888	.875	.801	.681	.079
3 (3 factors)	2,264.28	290	7.8	.889	.901	.890	.823	.685	.074
4 (3 factors)	2,011.75	290	6.9	.901	.914	.904	.849	.494	.069
5 (4 factors)	1,840.16	287	6.4	.910	.922	.912	.857	.447	.066
6 (4 factors)	1,648.02	287	5.7	.919	.932	.923	.879	.445	.062
7 (5 factors)	1,362.58	283	4.8	.933	.946	.938	.899	.417	.055
8 (6 factors)	1,195.93	278	4.3	.941	.954	.946	.912	.365	.051
9 (8 factors)	Error: Model would not converge								
10 (10 factors)[a]	Error: Model would not converge								

Note. Weighted $N = 1{,}250$; number of subtests = 26. WAIS–III = Wechsler Adult Intelligence Scale (3rd ed.); WMS–III = Wechsler Memory Scale (3rd ed.); NFI = normed fit index; CFI = comparative fit index; TLI = Tucker–Lewis Index; AGFI = adjusted goodness-of-fit index; RMSR = root-mean-square residual; RMSEA = root-mean-square error of approximation.
[a] Model 10 represents a hierarchical model with immediate and delayed.

assumptions in an iterative fashion to determine whether there would be better model fit with some select subtests (e.g., Picture Arrangement, Arithmetic, Spatial Span, and Visual Reproduction), either loading on an alternative factor or having a split between two factors. Such analyses aid in determining on which factor(s) these subtests should be placed.

Study 2: Testing Alternate Models to Determine Subtest Loadings

In this study, new measurement models were developed and fit in an iterative fashion to determine whether the model fit statistics would improve when a subtest was specified to load on a different factor or was not restricted to a single factor. The starting model was based on the six-factor model that was presented in Figure 1. For each successive analysis, one subtest was specified on a different factor. When testing the next model, we reset the specifications to the structure shown in Figure 1, with the exception of the subtest being tested. The only times that more than one subtest was altered from that specified in Figure 1 were in Models 7 and 9, where the specified loadings of both the immediate and the delayed Visual Reproduction variables differed from those specified originally, and in Model 11, where all of the variables in question were allowed to have split loadings and all of these changes were made simultaneously.

Four subtests have been shown to have alternative or split loadings in previous factor analytic studies. Picture Arrangement often has split loadings between Perceptual and Verbal factors (e.g., Leckliter, Matarazzo, & Silverstein, 1986 [WAIS–R]; Wechsler, 1991 [WISC–III]; The Psychological Corporation, 1997 [WAIS–III]). Arithmetic has strong loadings on the Verbal as well as the Attention/Working Memory factor (e.g., Leckliter et al., 1986 [WAIS–R]; Wechsler, 1991 [WISC–III]; The Psychological Corporation, 1997 [WAIS–III]). Spatial Span, or Visual Memory Span as it was called in the WMS–R (Wechsler, 1987), has been demonstrated to have significant loadings on both the Perceptual Organization and the Attention/Working Memory factor (Leonberger et al., 1992; Nicks, Leonberger, Munz, & Goldfader, 1992; Tulsky et al., 1998). Finally, the Visual Reproduction immediate

Table 3
Bayes Information Criterion (BIC) Differences When Comparing Models

Model	$\text{Model}_x - \text{Model}_y$	BIC_x	BIC_y	BIC difference ($\text{BIC}_x - \text{BIC}_y$)
2	Model 1 − Model 2	3,729.58	3,155.45	574.13
3	Model 2 − Model 3	3,155.45	2,898.01	257.44
4	Model 2 − Model 4	3,155.45	2,645.48	509.97
5	Model 4 − Model 5	2,645.48	2,505.05	140.43
6	Model 4 − Model 6	2,645.48	2,312.92	332.57
7	Model 6 − Model 7	2,312.92	2,069.03	243.88
8	Model 7 − Model 8	2,069.03	1,954.33	114.70

Note.

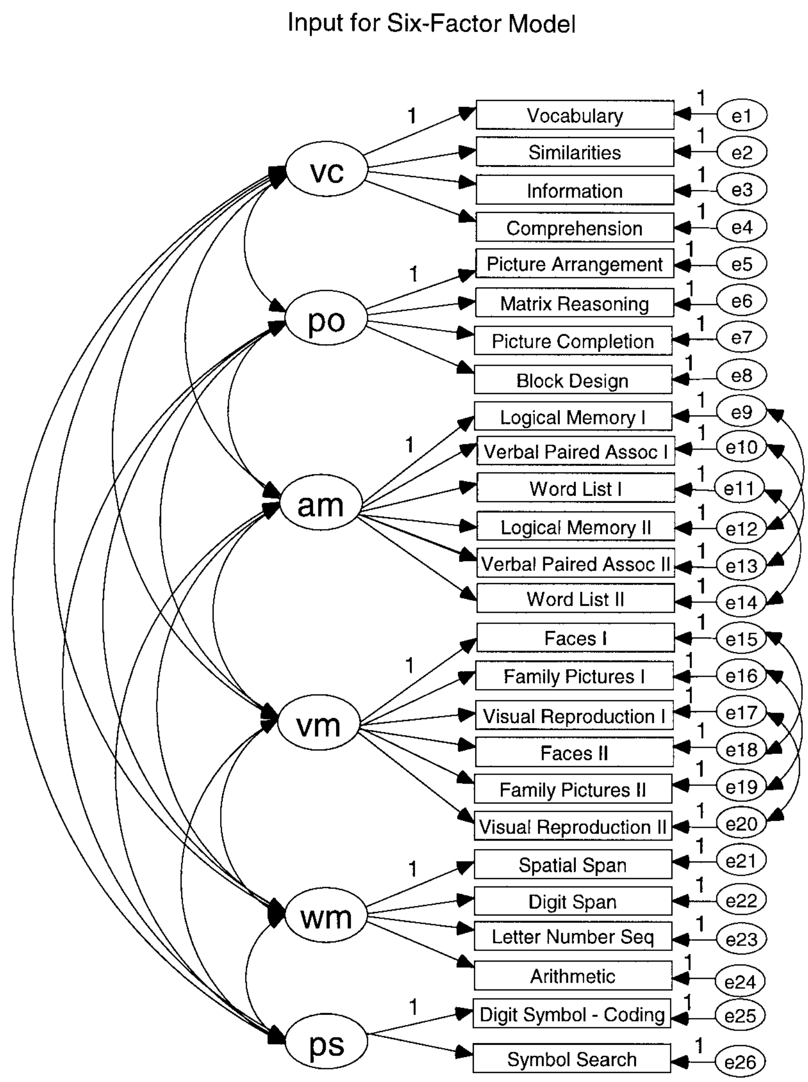

Figure 1. Path specification of the six-factor model. vc = Verbal Comprehension; po = Perceptual Organization; am = Auditory Memory; vm = Visual Memory; wm = Working Memory; ps = Processing Speed; Assoc = Associates; Seq = Sequencing.

and delayed tasks have had significant loadings on the Perceptual Organization factor (see Leonberger et al., 1992; Nicks et al., 1992) or split between Perceptual Organization and Attention/ Working Memory factors (Larrabee & Curtiss, 1992; Larrabee, Kane, Schuck, & Francis, 1985; Tulsky et al., 1998).

Method

Subjects. The WMS–III standardization sample (weighted N = 1,250) served as the sample for this study. Details have been provided previously and are not repeated here.

Analytic procedure. Model 8 from Study 1 (i.e., the six-factor model) was used to establish the structure of the relationship between the WAIS–III and WMS–III subtests. Slight variations were tested by changing the path restriction of one variable in each test. For instance, in the first iteration, Picture Arrangement was restricted to load on the Verbal Comprehension factor rather than on the Perceptual Organization factor. The results of this run would indicate whether the fit statistics provided a better fit to the model than in Model 8 from the previous study. Following these analyses, a new variation was made to Model 8 (in this case, Picture Arrangement was allowed to load on both the Verbal Comprehension and Perceptual Organization factors), the model was tested, and the fit indices between this analysis and the results from Model 8 were compared. This procedure was repeated for all of the hypothesized alternative models (see Table 5).

For each variation of Model 8 that was tested, the fit indices (as outlined in Study 1) were compared with the fit indices that were obtained for Model 8. The BIC fit statistic was calculated by subtracting the BIC obtained for iteration from the BIC that had been obtained for Model 8. The final model integrated all of the “improvements” made through the successive iterations. Changes in the specified factor loadings for the Picture Arrangement, Arithmetic, Spatial Span, and Visual Reproduction subtests were made simultaneously and tested in a final run (i.e., Iteration 11). The specification decisions were made on the basis of the results of Iterations 1–10.

Table 4
Standardized Solutions by Confirmatory Factor Analysis for the Six-Factor Model

Measure	Verbal Comprehension	Perceptual Organization	Working Memory	Processing Speed	Auditory Memory	Visual Memory
Vocabulary	.91					
Information	.84					
Similarities	.82					
Comprehension	.83					
Matrix Reasoning		.74				
Block Design		.71				
Picture Completion		.67				
Picture Arrangement		.66				
Letter Number Sequencing			.74			
Digit Span			.64			
Arithmetic			.79			
Spatial Span			.61			
Symbol Search				.86		
Digit Symbol				.74		
Logical Memory I					.73	
Logical Memory II					.75	
Verbal Paired Associates I					.69	
Verbal Paired Associates II					.66	
Word List I					.70	
Word List II					.60	
Faces I						.32
Faces II						.37
Family Pictures I						.57
Family Pictures II						.61
Visual Reproduction I						.68
Visual Reproduction II						.63

Note. Values in this table are standardized regression rates (factor loadings); factor covariances and subtest error terms are not represented.

Results and Discussion

Confirmatory analyses of the measurement models were conducted with the AMOS 4.0 program (Arbuckle & Wothke, 1999) on the modified measurement models. The goodness-of-fit statistics, measures of model parsimony, and information–theoretic indices for each of these modified models as well as the original model are presented in Table 6. For Iterations 1, 3, 7, and 8, the model fit statistics did not improve upon the model that had been specified originally. This finding verifies that the Picture Arrangement, Arithmetic, and Visual Reproduction subtests were correctly placed on the factors that they were intended to measure during model specification. The results are less clear for Spatial Span, as they point to the equally strong spatial loading of this subtest (Iteration 5).

These findings do not, however, ensure accurate model specification, as the subtests might be related to more than one factor. When each of the variables in question was allowed to load on a second factor, as in Iterations 2, 4, 6, 9, and 10, most of the fit statistics improved. In some cases, the results were relatively large, as in Iteration 9, where the Visual Reproduction I and II subtests were allowed to have split loadings between the Visual Memory and Perceptual Organization subtests. Furthermore, the BIC difference statistic indicated significant improvement for Iterations 4, 9, and 10. These results confirm the multifactorial nature of some of the WAIS–III and WMS–III tasks. Table 7 lists the factor loadings for the Picture Arrangement, Arithmetic, Spatial Span, and Visual Reproduction subtests from Models 2, 4, 6, 9, and 10 (when split loadings were permitted). The magnitude of the secondary loadings can help the user determine the degree to which the subtests have relationships to two factors, and this is strongest for the Arithmetic, Spatial Span, and Visual Reproduction Immediate factors. It would appear that the WAIS–III and WMS–III are similar to their predecessors in containing these subtests that tap into complex abilities and are not easily captured on a single factor score.

As a final step, Iteration 11 allowed all of the subtests in question (e.g., Picture Arrangement, Arithmetic, Spatial Span, and Visual Reproduction I and II) to have unrestricted paths on two factors. The structure of this model and the respective factor loadings are presented in Figure 2. Fit statistics for Iteration 11 are included in Table 5 and show markedly improved model fit statistics on all measures. The AGFI increased to .926, the CFI to .965, and the TLI to .958; the RMSEA was reduced to .045; and the BIC difference statistic improved by 169.2. These improvements in model fit help support the premise that Picture Arrangement, Spatial Span, Visual Reproduction Immediate, and Arithmetic are more complicated tasks and do not measure only one factor. Restricting them to load on one factor reduces the overall model fit.

Study 3: Cross-Validation of the Six-Factor Models in an Independent Sample

Studies 1 and 2 presented in this report described a series of confirmatory factor analyses examining the joint structure of the WAIS–III and WMS–III based on the WMS–III standardization

Table 5
Variations to Model 8 Performed in Successive Iterations

Iteration	Change made to Model 8
1	Picture Arrangement was removed from the Perceptual Organization factor. Picture Arrangement was forced to load only on the Verbal Comprehension factor.
2	Picture Arrangement was not restricted to the Perceptual Organization factor. Picture Arrangement was allowed to load on both the Verbal Comprehension and Perceptual Organization factors.
3	Arithmetic was removed from the Working Memory factor. Arithmetic was forced to load only on the Verbal Comprehension factor.
4	Arithmetic was not restricted to the Working Memory factor. Arithmetic was allowed to load on both the Verbal Comprehension and Working Memory factors.
5	Spatial Span was removed from the Working Memory factor. Spatial Span was forced to load only on the Perceptual Organization factor.
6	Spatial Span was not restricted to the Working Memory factor. Spatial Span was allowed to load on both the Working Memory and Perceptual Organization factors.
7	Visual Reproduction I and II were removed from the Visual Memory factor. Visual Reproduction I and II were forced to load only on the Perceptual Organization factor.
8	Visual Reproduction I was removed from the Visual Memory factor. Visual Reproduction I was forced to load only on the Perceptual Organization factor. Visual Reproduction II was forced to load only on the Visual Memory factor.
9	Visual Reproduction I and II were not restricted to the Visual Memory factor. Visual Reproduction I and II were allowed to load on both the Visual Memory and Perceptual Organization factors.
10	Visual Reproduction I was not restricted to the Visual Memory factor. Visual Reproduction I was allowed to load on the Visual Memory and Perceptual Organization factors. Visual Reproduction II was forced to load only on the Visual Memory factor.
11	On the basis of the results obtained above, changes to Model 8 were made for the Picture Arrangement, Arithmetic, Spatial Span, and Visual Reproduction subtests simultaneously.

sample. The studies demonstrated that a six-factor model provides the best fit indices to describe the joint factor structure. Cross-validation provides an examination of fit indices in independent samples, thereby confirming, or disputing, the results from the previous analyses. Therefore, an important next step is cross-validating findings from Studies 1 and 2 in an independent validity sample.

Method

Subjects. A sample of 858 examinees who completed the standardization editions of the WAIS–III and WMS–III but were not included in the WMS–III standardization sample served as the validity sample. Of these individuals, 279 were included in the WAIS–III standardization sample but

Table 6
Confirmatory Factor Analyses With Alternative Models Testing Subtest Fit Within the Six-Factor Model

		Goodness-of-fit index									
Subtest	Iteration	χ^2	*df*	χ^2/df	NFI	CFI	TLI	AGFI	RMSR	RMSEA	BIC difference
Model 8 (original model)		1,195.93	278	4.3	.941	.954	.946	.912	.365	.051	
Pic Arrg	1: No split loadings	1,279.08	278	4.6	.937	.950	.942	.906	.400	.054	−83.20
Pic Arrg	2: With split loadings[a]	1,174.18	277	4.2	.942	.955	.947	.913	.362	.051	11.40
Arithmetic	3: No split loadings	1,298.76	278	4.7	.936	.949	.940	.903	.381	.054	−102.83
Arithmetic	4: With split loadings[a]	1,106.42	277	4.0	.946	.959	.951	.916	.352	.049	79.12
Spatial Span	5: No split loadings[a]	1,175.85	278	4.2	.942	.955	.948	.912	.365	.051	20.09
Spatial Span	6: With split loadings[a]	1,161.46	277	4.3	.943	.956	.948	.913	.364	.051	24.09
Vis Repr I/II	7: No split loadings	1,248.13	278	4.5	.939	.952	.943	.909	.396	.053	−52.20
Vis Repr I	8: No split loadings	1,237.33	278	4.5	.939	.952	.944	.910	.382	.053	−41.40
Vis Repr I/II	9: With split loadings[a]	1,137.08	276	4.1	.944	.957	.949	.916	.338	.050	38.10
Vis Repr I	10: With split loadings[a]	1,151.78	277	4.2	.943	.956	.949	.915	.348	.050	33.80
All changes	11	974.83	273	3.6	.952	.965	.958	.926	.317	.045	169.16

Note. Weighted $N = 1{,}250$ for the calibration sample. NFI = normed fit index; CFI = comparative fit index; TLI = Tucker–Lewis Index; AGFI = adjusted goodness-of-fit index; RMSR = root-mean-square residual; RMSEA = root-mean-square error of approximation; BIC = Bayes information criterion; Pic Arrg = Picture Arrangement; Vis Repr = Visual Reproduction.
[a] These iterations resulted in positive changes that improved model fit; however, none showed as much model improvement as Model 11, which allowed all variables in question to have split loadings.

Table 7
Standardized Solutions by Confirmatory Factor Analyses When Subtests Are Allowed to Have Split Loadings

		Factor					
Iteration	Subtest	VC	PO	WM	PS	AM	VM
2	Picture Arrangement	.22	.47				
4	Arithmetic	.36		.49			
6	Spatial Span		.38	.25			
9	Visual Reproduction I		.39				.34
	Visual Reproduction II		.21				.46
10	Visual Reproduction I		.29				.43
11	Picture Arrangement	.25	.44				
	Arithmetic	.38		.46			
	Spatial Span		.38	.29			
	Visual Reproduction I		.40				.35
	Visual Reproduction II		.21				.46

Note. VC = Verbal Comprehension; PO = Perceptual Organization; WM = Working Memory; PS = Processing Speed; AM = Auditory Memory; VM = Visual Memory.

not the WMS–III sample, 173 were "extra" cases not included in either of the standardization samples, 402 were collected as part of an education oversample (for future development of demographically adjusted normative information), and 4 were part of validity and bias samples. For these individuals, there was no report of neuropsychological injury, and they all met the inclusionary criteria to participate in the larger study (see The Psychological Corporation, 1997). Though testing order on the WAIS–III and WMS–III was counterbalanced, in this sample, a slightly higher percentage of examinees completed the WAIS–III first in the sequence (i.e., 475 individuals completed the WAIS–III first, whereas 383 completed the WMS–III first), which should not have affected scores significantly as there is some evidence that testing order does not play a significant confounding factor (see Zhu & Tulsky, 2000).

The sample ranged in age from 16 to 88 years ($M = 36.5$; $SD = 21.7$). The full-scale IQ score for this sample averaged 95.5 ($SD = 14.8$). The demographic characteristics of this sample are provided in Table 8.

Measures. Examinees completed 26 subtests from the WAIS–III and the WMS–III: Vocabulary, Information, Similarities, Comprehension, Block Design, Matrix Reasoning, Picture Completion, Picture Arrangement, Symbol Search, Digit Symbol—Coding, Arithmetic, Digit Span, Letter Number Sequencing, Spatial Span, Faces (Immediate and Delayed), Family Pictures (Immediate and Delayed), Visual Reproduction (Immediate and Delayed), Logical Memory (Immediate and Delayed), Verbal Paired Associates (Immediate and Delayed), and Word List (Immediate and Delayed). As reported earlier, the Letter Number Sequencing was administered only one time in the testing session (see Tulsky & Zhu, 2000), as part of the WMS–III battery.

Cross-validation procedure. Evaluation of the adequacy of model replication was conducted using a parameter-invariance methodology. This method involves testing the invariance of factor loadings between the calibration and validation samples and is accomplished by setting the standardized regression weights in the new sample to those that had been obtained with the calibration sample (e.g., as had been reported in Table 4). Comparison of model fit was evaluated by examining differences between the calibration and validation samples in relation to their absolute, incremental, parsimonious, and information–theoretic measures. If the model cross-validates well, there should be little or no difference between these measures of fit for the calibration and validation samples.

Results and Discussion

Confirmatory analyses of the measurement models were conducted using the AMOS 4.0 software program (Arbuckle & Wothke, 1999). The models differed only in whether select subtests were restricted to load on a single factor or whether they could be split between two factors. Model 1 restricted the loadings to a single factor, whereas Model 2 specified split loadings. Each of these models was tested for the two samples described above.

Table 9 lists the various model fit indices for both Model 1 and Model 2 in both the calibration and validation samples. The results indicate that both models fit the data reasonably well. For Model 1, the six-factor solution with the subtests restricted to a single factor, the AGFI value is .904, NFI is .935, and CFI is .955. Next, examination of the RMSR and RMSEA provided additional information for establishing which model best fit the data. Although the RMSR values were higher than desired (.465), the RMSEA values of .050 also provide evidence of acceptable model fit. The TLI value of .950 also demonstrates that the model fits the data well and is very close in magnitude to the original estimate. As in the original analyses, the fit indices improve significantly when Model 2 (i.e., the six-factor model with split loadings) is tested. The AGFI value increased to .925, the NFI is .949, and the CFI is .968. Examination of the RMSR (.400) and RMSEA (.042) provide additional support that the six-factor solution replicates in a new sample and that when loadings on the multifaceted subtests are not restricted to a single factor, the model fit statistics improve. The TLI value increases to .966, further supporting improvement in model fit. Collectively, these indices demonstrate that the models fit the data and are very close in magnitude to the fit statistics that were obtained in the calibration sample.

General Discussion

With the co-norming methodology used in the WAIS–III and WMS–III standardization projects, a true integration between the

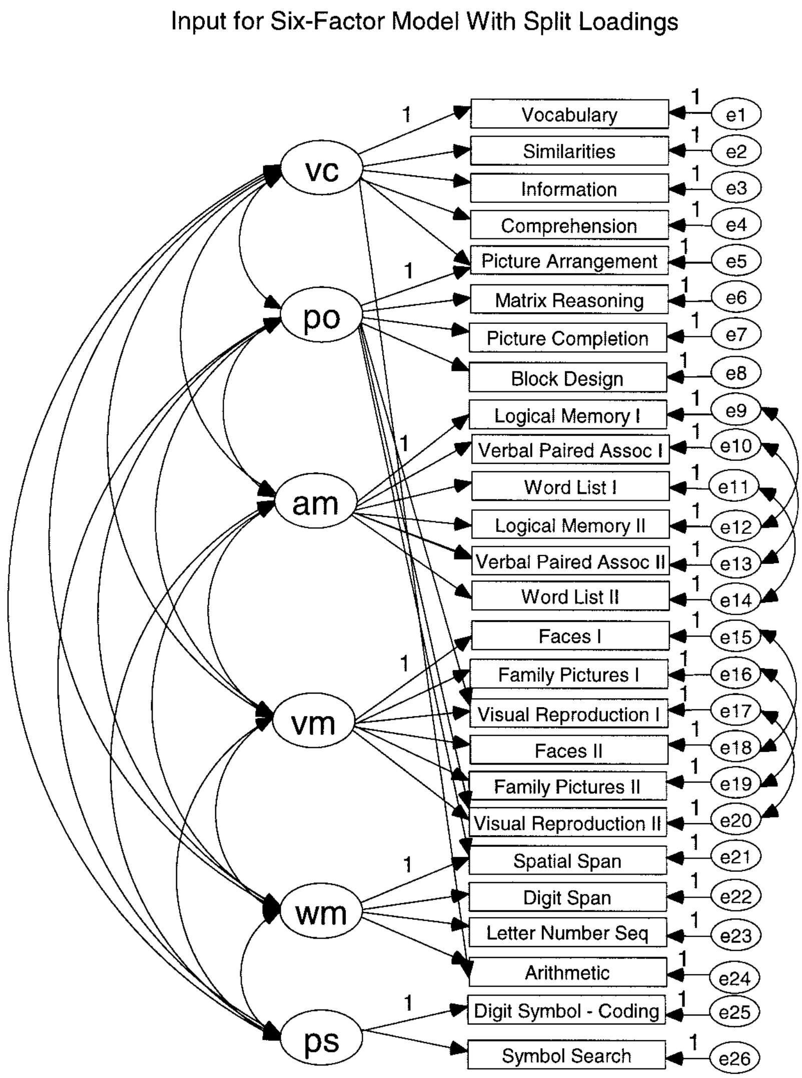

Figure 2. Path specification of the six-factor model when certain subtests were allowed to load on more than one factor. This figure is used to show the path specifications for the subtests, error terms, and covariances between factors. vc = Verbal Comprehension; po = Perceptual Organization; am = Auditory Memory; vm = Visual Memory; wm = Working Memory; ps = Processing Speed; Assoc = Associates; Seq = Sequencing.

tests is possible. However, the published versions of the scales did not present such a unified framework, and an expanded model of cognitive functioning was not examined. The studies presented here are unique in that they represent the first attempt with the standardization data sets to factor analyze the WAIS–III and WMS–III simultaneously. Through these analyses, we now have increased understanding regarding the domains of cognitive functioning as measured by the WAIS–III and WMS–III. The results indicate that the WMS–III measures areas of functioning not tapped by the WAIS–III. Taken together, at least six factors of cognition, represented by Verbal Comprehension, Perceptual Organization, Working Memory, Processing Speed, Auditory Memory, and Visual Memory, can be measured. The results reported here provide the groundwork for the development of a new unified battery for the WAIS–III and WMS–III. These results have led Tulsky, Ivnik, et al. (2003) to argue for a new unified structure of the WAIS–III and WMS–III that they claim provides a clearer picture of cognitive functioning.

Studies 1 and 3 used a rigorous application of structural equation modeling that represents a significant improvement over previous methods used to examine the factor structure of the Wechsler scales. Several models of cognitive functioning were developed on the basis of previous research on these intelligence scales, and the model fit statistics across the various models were compared. In Study 1, a six-factor model was identified as having the best fit to the data, and in Study 3, cross-validation of this model was conducted using a rigorous parameter-invariance methodology that replicated the results in a new sample. This provided a rigorous analytic framework from which conclusions have been advanced, as well as providing extremely strong support for a six-factor model for the WAIS–III and WMS–III.

Table 8
Demographic Composition of the WMS–III Validity Sample 2

Age group (years)	*n*	Education level (years)	*n*	Ethnicity	*n*	Gender	*n*
16–17	177	8 or less	245	African American	191	Female	457
18–19	150	9–11	184	Hispanic	124	Male	401
20–24	56	12	169	Other	61		
25–29	74	13–15	117	White	482		
30–34	57	16+	143				
35–44	97						
45–54	66						
55–64	30						
65–69	32						
70–74	21						
75–79	40						
80–84	48						
85–89	10						

Note. WMS–III = Wechsler Memory Scale (3rd ed.).

Study 2 confirmed the multifaceted nature of some subtests of the WAIS–III and WMS–III. Increased model fit statistics were obtained when the Arithmetic, Spatial Span, Picture Arrangement, and Visual Reproduction I and II subtests were not restricted to load on a single factor. Such findings are consistent with the results obtained in previous factor analytic studies performed on the earlier editions of the Wechsler tests. Nevertheless, these results raise the question of whether the current scoring system for the Wechsler scales is appropriate for the test if subtests load on more than one factor. Current scoring methods have favored a practical and simple scoring methodology in which a "sum of scaled scores" for each index is calculated by summing equally weighted subtest scale scores, and each subtest is included on one, and only one, factor. Although it is possible to develop alternative scoring methods where subtests' contributions are differentially weighted and partial scores for subtests can contribute simultaneously to more than one factor (e.g., see Parker & Atkinson, 1995, who developed such a method for the WAIS–R), such scoring methods are difficult to implement in clinical practice. Indeed, practical considerations do play a role, and for the joint factor structure that is presented in a companion chapter (see Tulsky, Ivnik, et al., 2003), these practical issues took precedent in developing an easy-to-use scoring structure. Nevertheless, clinicians should be aware of the multifaceted nature of some of the Wechsler subtests and use this information to guide interpretation as appropriate. Additionally, these results could be of help to the test developers in future revisions of the Wechsler scales, and consideration should be given for replacing these subtests. Though such replacements may break with the tradition of the Wechsler scales, this type of change will enhance the ultimate structure of the scales by allowing better measurement of a four- or six-factor model.

Another significant finding is that, consistent with other recent publications, the factor analytic results cannot support separate immediate and delayed factors (see Millis et al., 1999; Price et al., 2002). Moreover, these results demonstrate that the Faces subtests are different from the other Visual Memory subtests (as indicated by the low factor loadings of Faces I and Faces II on the Visual Memory factor), a finding that was first reported by Millis et al. (1999). Therefore, the contribution of the Faces subtests to the Visual Memory index scores remains open to debate, and Faces may be measuring a construct unlike any other on the WAIS–III and WMS–III. Therefore, because of the continued finding that

Table 9
Goodness-of-Fit Statistics for Cross-Validation Study

Fit criteria			Model represented in Figure 1	Model represented in Figure 2
Fit index	Model improvement	Very close model fit	Validity sample (N = 858): Six-factor	Validity sample (N = 858): Split loadings
χ^2/df	Decrease		3.2	2.5
AGFI	Increase	>.900	.904	.925
NFI	Increase	>.900	.935	.949
CFI	Increase	>.900	.955	.968
TLI	Increase	>.950	.950	.966
RMSR	Decrease	<.400	.465	.400
RMSEA	Decrease	<.050	.050	.042

Note. AGFI = adjusted goodness-of-fit index; NFI = normed fit index; CFI = comparative fit index; TLI = Tucker–Lewis Index; RMSR = root-mean-square residual; RMSEA = root-mean-square error of approximation.

Faces has such a low loading on the Visual Memory factor, Tulsky, Ivnik, et al. (2003) have developed norms for an alternative Visual Memory score that includes a combination of Visual Reproduction and Family Pictures.

There are some limitations to the current work and avenues for future research. First, the analyses presented here were based on normative standardization samples. Because these tests are often used with clinical populations, it is important to ascertain whether the factor structure proposed here will be supported in various clinical populations. One specific issue revolves around support for the distinction between immediate and delayed memory constructs within clinical groups. Such distinctions have been found to be useful in clinical settings (see Heaton, Taylor, & Manly, 2003; Tulsky, Ivnik, et al., 2003) but have not received factor analytic support within the WMS–III (see Millis et al., 1999; Price et al., 2002). However, as Millis et al. pointed out, the differentiation between the immediate and delayed memory variables may be more apparent in clinical samples.

Second, these results did not examine differences that might have occurred if the analyses were broken up by demographic variables, and therefore, we do not know whether the factor structure will remain consistent across different demographic groups. Millis et al. (1999) argued that at the early stages of model development, there are insufficient empirical data to develop baseline models to test for invariance across group samples. According to their logic, it seems appropriate to use the entire standardization sample to develop a baseline model at this initial stage of examining the joint factor structure between the WAIS–III and WMS–III. However, the reader should be aware that previous studies have found differences in factor structures, as well as score magnitudes, between demographic groups. For instance, in factor analytic work on the WAIS–III, differences between older and younger samples have occurred (e.g., the distinction between the Processing Speed and Perceptual Organization factors was not supported in a group made up of the oldest adults in the WAIS–III standardization sample; see The Psychological Corporation, 1997). Moreover, other demographic variables might be important. Heaton et al. (2003) provided strong evidence suggesting that demographic variables have a strong impact on WAIS–III and WMS–III scores. They suggested that correction for demographic variables should occur after the fact and presented a procedure for demographic correction of the six-factor model that was derived in part on the results presented here. Future research should test for the specific impact of demographic variables.

Finally, although a six-factor solution seems supported by these studies, continued testing of alternative models should continue. Follow-up studies are needed to verify the structure in different samples and with different lists of variables. Alternative models could be developed and tested, and other subtests and scores from the WAIS–III and WMS–III (e.g., Mental Control, WMS–III and WAIS–III optional scores) or other tests measuring different constructs (e.g., learning variables from a test such as the California Verbal Learning Test; Delis, Kramer, Kaplan, & Ober, 1987, 2000) could be included in future factor analyses. These findings might yield different results and extend the factor structure of these scales beyond the six-factor model reported here. Another study might include validating the structure with a shortened variable list (perhaps removing the optional subtests) to see whether model fit improves. Nevertheless, the analyses and results reported here can help guide us as we search for ways to improve the practice of cognitive assessment.

References

Arbuckle, J. L., & Wothke, W. (1999). *AMOS 4.0 user's guide.* Chicago: SmallWaters Corporation.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Binet, A., & Simon, T. (1916). *The development of intelligence in children* (E. S. Kite, Trans.). Baltimore: Williams & Wilkins. (Original work published 1905)

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Bollen, K. A., & Long, J. S. (1993). Introduction. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 27–29). Newbury Park, CA: Sage.

Bowden, S. C., Carstairs, J. R., & Shores, E. A. (1999). Confirmatory factor analysis of combined Wechsler Adult Intelligence Scale—Revised and Wechsler Memory Scale—Revised scores in a healthy community sample. *Psychological Assessment, 11,* 339–344.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.). *Testing structural equation models* (p. 22). Newbury Park, CA: Sage.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies.* Cambridge, MA: Cambridge University Press.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *The California Verbal Learning Test—Adult Version.* San Antonio, TX: The Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *The California Verbal Learning Test—Second Edition.* San Antonio, TX: The Psychological Corporation.

Hawkins, K. A. (1998). Indicators of brain dysfunction derived from graphic representations of the WAIS–III/WMS–III technical manual samples: A preliminary approach to clinical utility. *The Clinical Neuropsychologist, 12,* 535–551.

Heaton, R. K., Taylor, M. J., & Manly, J. (2003). Demographic effects and the use of demographically corrected norms with the WAIS–III and WMS–III. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS–III and WMS–III* (pp. 181–210). San Diego, CA: Academic Press.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: G_f–G_c theory. In D. P. Flanagan, J. L. Genshart, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53–91). New York: Guilford Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6(1),* 1–55.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: User's reference guide.* Chicago: Scientific Software International.

Larrabee, G. J. (1999). Book and test review: The Wechsler Memory Scale III. *Archives of Clinical Neuropsychology, 14,* 473–477.

Larrabee, G. J. (2000). Association between IQ and neuropsychological test performance: Commentary on Tremont, Hoffman, Scott, and Adams (1998). *The Clinical Neuropsychologist, 14,* 139–145.

Larrabee, G. J., & Curtiss, G. (1992). Factor validity of various memory testing procedures [Abstract of presentation at the annual convention of the International Neuropsychological Society]. *The Clinical Neuropsychologist, 5,* 261–262.

Larrabee, G. J., Kane, R. L., & Schuck, J. R. (1983). Factor analysis of the WAIS and Wechsler Memory Scale: An analysis of the construct validity of the Wechsler Memory Scale. *Journal of Clinical Neuropsychology, 5,* 159–168.

Larrabee, G. J., Kane, R. L., Schuck, J. R., & Francis, D. E. (1985). The construct validity of various memory testing procedures. *Journal of Clinical and Experimental Neuropsychology, 7,* 239–250.

Leckliter, I. N., Matarazzo, J. D., & Silverstein, A. B. (1986). A literature review of factor analytic studies of the WAIS–R. *Journal of Clinical Psychology, 42,* 332–342.

Leonberger, F. T., Nicks, S. D., Larrabee, G. J., & Goldfader, P. R. (1992). Factor structure of the Wechsler Memory Scale—Revised within a comprehensive neuropsychological battery. *Neuropsychology, 6,* 239–249.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory factor analysis: Effects of sample size. *Psychological Bulletin, 103,* 391–411.

Martin, T. A., Donders, J., & Thompson, E. (2000). Potential of and problems with new measures of psychometric intelligence after traumatic brain injury. *Rehabilitation Psychology, 45,* 402–408.

McCarthy, R. A., & Warrington, E. K. (1990). *Clinical neuropsychology: A clinical introduction.* San Diego, CA: Academic Press.

Millis, S. R., Malina, A. C., Bowers, D. A., & Ricker, J. H. (1999). Confirmatory factor analysis of the Wechsler Memory Scale–III. *Journal of Clinical and Experimental Neuropsychology, 21,* 87–93.

Nicks, S. D., Leonberger, F. T., Munz, D. C., & Goldfader, P. R. (1992). Factor analysis of the WMS–R and the WAIS. *Archives of Clinical Neuropsychology, 7,* 387–393.

Parker, K. C. H., & Atkinson, L. (1995). Computation of Wechsler Adult Intelligence Scale—Revised factor scores: Equal and differential weights. *Psychological Assessment, 7,* 456–462.

Price, L. R., Tulsky, D. S., Millis, S. R., & Weiss, L. G. (2002). Redefining the factor structure of the Wechsler Memory Scale–III: Confirmatory factor analysis with cross-validation. *Journal of Clinical and Experimental Neuropsychology, 24,* 574–585.

The Psychological Corporation. (1997). *WAIS–III—WMS–III technical manual.* San Antonio, TX: Author.

Raftery, A. E. (1993). Bayesian model selection in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 163–180). Newbury Park, CA: Sage.

Ryan, J. J., Rosenberg, S. J., & Heilbronner, R. L. (1984). Comparative relationship of the Wechsler Adult Intelligence Scale—Revised and Wechsler Adult Intelligence Scale to the Wechsler Memory Scale. *Journal of Behavioural Assessment, 6,* 37–43.

Smith, G. E., Ivnik, R. J., Malec, J. F., Petersen, R. C., Tangalos, E. G., & Kurland, L. T. (1992). Mayo's Older Americans Normative Studies (MOANS): Factor structure of a core battery. *Psychological Assessment, 4,* 382–390.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10–12). Newbury Park, CA: Sage.

Tucker, L. R., & Lewis C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1–10.

Tulsky, D. S., Chiaravalloti, N. D., Palmer, B. W., & Chelune, G. J. (2003). The Wechsler Memory Scale, Third Edition: A new perspective. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS–III and WMS–III* (pp. 93–139). San Diego, CA: Academic Press.

Tulsky, D. S., Ivnik, R. J., Price, L. R., & Wilkins, C. (2003). Assessment of cognitive functioning with the WAIS–III and WMS–III: Development of a 6-factor model. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS–III and WMS–III* (pp. 147–179). San Diego, CA: Academic Press.

Tulsky, D. S., & Ledbetter, M. (2000). Updating to the WAIS–III and WMS–III: Considerations for research and clinical practice. *Psychological Assessment, 12,* 253–262.

Tulsky, D. S., Ledbetter, M. F., & Zhu J. (1998). Factor analyses of the WAIS–III and WMS–III [Abstract of presentation at the 106th Annual Convention of the American Psychological Association, San Francisco]. *The Clinical Neuropsychologist, 12,* 295.

Tulsky, D. S., Saklofske, D. H., & Ricker, J. H. (2003). Historical overview of intelligence and memory: Factors influencing the Wechsler scales. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS–III and WMS–III* (pp. 7–41). San Diego, CA: Academic Press.

Tulsky, D. S., Saklofske, D. H., & Zhu, J. (2003). Revising a standard: An evaluation of the origin and development of the WAIS–III. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, et al. (Eds.), *Clinical interpretation of the WAIS–III and WMS–III* (pp. 43–91). San Diego, CA: Academic Press.

Tulsky, D. S., & Zhu, J. (2000). Could test length or order affect scores on Letter Number Sequencing of the WAIS–III and WMS–III? Ruling out effects of fatigue. *The Clinical Neuropsychologist, 14,* 474–478.

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

Wechsler, D. (1955). *Wechsler Adult Intelligence Scale.* New York: The Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1987). *The Wechsler Memory Scale—Revised.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1991). *The Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997a). *WAIS–III administration and scoring manual.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale—Third Edition.* San Antonio, TX: The Psychological Corporation.

Zhu, J., & Tulsky, D. S. (2000). The effects of co-norming on the WAIS–III and WMS–III: Is there a test-order effect on IQ and memory scores? *The Clinical Neuropsychologist, 14,* 461–467.

Received June 18, 2001
Revision received February 4, 2003
Accepted February 12, 2003 ■