Published in final edited form as:

Assessment. 2014 April ; 21(2): 131–142. doi:10.1177/1073191113486690.

Evaluating the Correspondence of Different Cognitive Batteries

Timothy A. Salthouse[1]

[1]University of Virginia, Charlottesville, VA, USA

Abstract

It is widely accepted that abilities are a meaningful level of abstraction for distinguishing among individuals with respect to their levels of cognitive functioning. However, relatively little is known about the extent to which different combinations of tests reflect the same cognitive abilities, or about the relation of cognitive abilities in one test battery with specific tests in another battery. Data from two cognitive batteries were analyzed to determine the correspondence of ability factors in the two batteries, and to evaluate the relative influence of cognitive abilities from one battery on the subtest scores in the other battery. Although the batteries involved different combinations of tests, correlations between the theoretically similar ability factors in the two batteries were very high (i.e., $r > .84$). Furthermore, with only a few exceptions, the primary influences on the subtest scores in one battery were from the theoretically relevant ability factor in the other battery.

Keywords

abilities; equivalence; validity

A very large number of cognitive tests have been developed, and many have been combined into batteries to provide a broad assessment of cognitive functioning. However, because different test batteries involve different combinations of tests, an important question is whether the same cognitive ability constructs are being assessed in each battery. A key assumption in psychological measurement (e.g., Campbell & Fiske, 1959; Messick, 1989) is that a theoretical construct can be considered to be valid to the extent that the construct is not specific to particular methods of measurement (i.e., exhibits convergent validity), and is not redundant with other constructs (i.e., exhibits discriminant validity). In the domain of cognitive functioning, this assumption implies that the meaningfulness of a cognitive ability construct is determined, at least in part, by the degree to which it is correlated with theoretically similar constructs based on alternative methods of assessment.

Corresponding Author: Timothy A. Salthouse, Department of Psychology, University of Virginia, 102 Gilmer Hall, Charlottesville, VA 22904, USA., salthouse@virginia.edu.

The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

Declaration of Conflicting Interests
The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

One way the question of correspondence of constructs across batteries can be investigated is with exploratory factor analyses in which measures from different test batteries are included in the same analysis to determine if they load on the same factor. Although this method can be informative about which measures group together, interpretations of the groupings can be somewhat subjective. Another approach consists of examining correlations of composite scores or latent variables hypothesized to represent similar cognitive abilities in separate batteries. Moreover, if the constructs are arranged in a hierarchical structure, correlations can be examined at different levels in the hierarchy, including the highest order factor representing general cognitive ability, or *g* (e.g., Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004; Johnson, te Nijenhuis, & Bouchard, 2008; Keith, Kranzler, & Flanagan, 2001).

Although valuable, a potential limitation of research examining across-battery correspondence at the level of ability factors is that it implicitly assumes that each measure is a valid indicator of its respective construct. The current study explicitly tested this assumption by investigating unique influences of the ability factors from one cognitive battery on individual subtest scores from the other battery. Because it was originally designed to investigate influences on target variables in the context of age-related differences in relevant cognitive abilities, the analytical procedure has been termed *contextual analysis* (e.g., Salthouse, 2005; Salthouse, Pink, & Tucker-Drob, 2008; Salthouse, Siedlecki, & Krueger, 2006). The analytical framework, in which different cognitive abilities serve as simultaneous predictors of a target variable, is illustrated in the left panel of Figure 1. Coefficients for the relations portrayed by dotted lines are of greatest interest for the current purpose as they indicate the relations between the ability factors from one test battery on individual subtest scores in the other test battery. If the target variable reflects a single factor it would be expected to have a strong unique relation with the relevant factor from the other battery, and little or no relations with other factors. In contrast, a discovery that the variable was unrelated to any of the factors, or had moderate to strong relations with several factors, would suggest that it reflects other ability factors instead of, or addition to, the hypothesized factor.

When a predictor is considered in isolation, its regression coefficient reflects all the influences shared with other variables in addition to influences unique to that variable. An advantage of the simultaneous analysis portrayed in the left panel of Figure 1 is that the coefficients reflect unique relations of each predictor, independent of the influences of the other predictors in the analysis. However, a disadvantage of simultaneous analysis is that collinearity can result if the predictors are strongly correlated, which can reduce the meaningfulness of the regression coefficients. One solution to the collinearity problem is to conduct separate analyses in which only one of the correlated predictors is in each analysis. This approach was used in the current study with predictors which had correlations with one another of .80 or greater.

Another method of dealing with the collinearity problem is to represent the shared influences with an orthogonal common factor related to all observed variables but not to any of the ability factors, and then examining relations of the ability factors and the orthogonal common factor on the target variable. A model of this type is portrayed in the right panel of

Figure 1. The major advantage of the orthogonal common factor approach is that shared influences that might be responsible for high correlations among ability predictors are explicitly represented as a separate factor, but it has the limitation that the common factor may absorb influences of variables and factors that are strongly related to it. In an attempt to converge on the most meaningful conclusion regarding ability influences on specific target variables, both the correlated predictors and orthogonal common factor models portrayed in Figure 1 were examined in this study.

The two batteries investigated in the current project were a combination of the most recent versions of the Wechsler cognitive (Wechsler Adult Intelligence Scale–IV [WAIS-IV], Wechsler, 2008) and memory (Wechsler Memory Scale–IV [WMS-IV], Wechsler, 2009) test batteries, and a cognitive battery used in an ongoing longitudinal study (the Virginia Cognitive Aging Project [VCAP]). These test batteries are well suited to address questions of comparability because, at least when the WAIS-IV and WMS-IV batteries are combined, both batteries assess a mixture of cognitive and memory abilities, and each had three or more tests representing each ability. Salthouse (2009) conducted parallel analyses of WAIS-IV/WMS-IV and VCAP data, and found similar structural patterns in the two batteries. However, because the data were derived from independent samples, no direct comparisons were possible in that report. The current study included data from a linkage sample of 90 adults who performed both the WAIS-IV/WMS-IV and VCAP batteries in addition to data from the WAIS-IV and WMS-IV normative samples, and data from the VCAP sample (as of the end of 2011).

To summarize, the primary research questions in this study were whether the ability factors hypothesized to represent similar abilities in two different test batteries were highly correlated with one another, and whether individual subtests from one battery were most strongly related to the relevant ability factor in the other test battery. Negative answers to the questions would raise concerns about the validity of the factors in one or both test batteries, whereas positive answers would increase confidence in the validity of the cognitive abilities and the specific tests in each battery.

Method

Participants

The WAIS-IV/WMS-IV data were derived from the samples used to establish the norms for those test batteries (Wechsler, 2008, 2009). The participants ranged in age from 16 to 90 years and were selected to match demographic proportions in the U.S. population. The sample sizes differed across variables because some of the tests were not administered to adults older than 69 years, and there was not complete overlap of the individuals in the samples used to establish the WAIS-IV norms and the WMS-IV norms. Further details on the sampling methods, and the representativeness of the samples, can be found in the Technical and Interpretive Manuals (i.e., Wechsler, Coalson, & Raiford, 2008; Wechsler, Holdnack, & Drozdick, 2009).

The VCAP data were collected between 2001 and 2011 with participants recruited from newspaper advertisements, flyers, and referrals from other participants. The sample sizes varied across variables, primarily because some tests were introduced in later years of the project. Participants ranged from 18 to 97 years old, and females outnumbered males because of their higher volunteering rates. Approximately 80% of the participants identified themselves as White, 10% as Black, and small percentages in other ethnicities, or reporting more than one ethnicity. More details on the samples are available in earlier reports (e.g., Salthouse, 2005; Salthouse et al., 2006; Salthouse et al., 2008).

The new linkage sample consisted of 90 healthy adults ranging from 25 to 79 years of age, with 10 males and 20 females in each 20-year age band. None of the individuals had previously participated in VCAP, and all had Mini Mental State Exam (Folstein, Folstein, & McHugh, 1975) scores greater than 23. Approximately 78% identified themselves as White, 17% as Black, and 5% reported more than one ethnicity. Each participant completed three 2-hour sessions within a period of approximately 1 month, with the VCAP battery performed on the first session, the WAIS-IV subtests on the second session, and the WMS-IV subtests on the third session.

Tests

Prior research (e.g., Salthouse, 2009) has established that a six-factor model (crystallized ability or Gc, fluid ability or Gf, working memory, verbal memory, visual memory, and processing speed)[1] provided a good fit to variables from the combined WAIS-IV and WMS-IV batteries, and that a five-factor model (vocabulary, reasoning, spatial visualization, memory, and speed) provided a good fit to the variables in the VCAP battery. These factor structures, which are illustrated with their relevant tests in Figure 2, were therefore used in the current analyses after two minor modifications from the analyses reported in Salthouse (2009). First, both immediate and delayed measures of the memory tests were included in the current analyses to broaden the coverage of the memory factors in the Wechsler battery. And second, different versions of the Vocabulary and Logical Memory tests were used in the VCAP battery for the linkage sample to avoid repetition of the very similar tests used in the WAIS-IV and WMS-IV batteries.

Detailed descriptions of the Wechsler tests are available in the WAIS-IV and WMS-IV manuals, and the VCAP tests and their sources are briefly described in the appendix. Test reliabilities, results of factor analyses, and age relations in both batteries were reported in Salthouse (2009).

Results

Means and standard deviations of the variables in each sample are reported in Table 1. The rightmost column contains effect sizes in *d* units for individual subtests. It can be seen that many of the means in the linkage sample were higher than the means from the WAIS-IV/WMS-IV sample (i.e., median *d* of .25 for the WAIS-IV variables and .41 for the WMS-IV

[1]The labels of the factors are the terms used in Salthouse (2009). In the WAIS-IV manual, the first factor is labeled verbal comprehension, the second perceptual reasoning, and the sixth processing speed. The WAIS-IV manual also refers to a working memory factor, but it is composed of the total score in the digit span task, the arithmetic task, and the letter–number sequencing task, without the symbol span and spatial addition tasks from the WMS-IV. The verbal memory and visual memory factors in Figure 2 are based on tasks from the WMS-IV.

variables), and that many of the means in the linkage sample were lower than those in the VCAP sample (i.e., median *d* of −.15). However, it is important to note that the similar values of the standard deviations imply that the samples had comparable levels of between-person variability.

Correspondence at the Ability Level

The correspondence between factors in the two test batteries was examined with correlations between composite scores, formed by averaging *z* scores of relevant variables. The correlations are reported in Table 2, with values above the diagonal from the linkage sample, values below and to the left from the VCAP sample, and values below and to the right from the WAIS-IV/WMS-IV sample. The similar values above and below the diagonal indicate that the correlations of the composite scores in the linkage sample were generally comparable to those in the VCAP and WAIS-IV/WMS-IV samples.

Correlations between theoretically similar constructs in the two batteries, indicated by the values in boldface, ranged from .84 to .88. Estimates of the reliability of the composites can be obtained from coefficient alpha with the subtest scores used to form composites serving as items. After adjusting for unreliability of the relevant constructs, the correlations between theoretically similar constructs ranged from .97 to 1.02. Correlations were also computed between composite scores formed after deleting the VCAP subtests derived from earlier versions of the Wechsler tests (i.e., Wechsler Vocabulary, Digit Symbol, and Logical Memory). These correlations were very similar to the original correlations in Table 2 as the Gc–Vocabulary correlation was .87, the correlation between the two speed constructs was .78, and the correlation between the Wechsler verbal memory and the VCAP memory constructs was .83. Correlations computed separately for males and females were also very similar to those in the table, as the correlations for theoretically parallel constructs ranged from .80 to .95.

The next analysis consisted of a structural equation model with hierarchical structures in both the Wechsler and VCAP data (see Figure 2, and Salthouse, 2009). These and other structural equation analyses carried out in this study were conducted with the AMOS (Arbuckle, 2006) statistical package. The model with the combined data had a reasonable fit (i.e., χ^2 = 6,864, *df* = 690, CFI = .92, RMSEA = .04).[2] The patterns of standardized relations to the highest order *g* factor were similar in the two batteries as the highest loadings were Gf (.94) in the Wechsler battery and reasoning (1.0) in the VCAP battery, and the lowest loadings were Gc (.63) in the Wechsler battery and vocabulary (.47) in the VCAP battery. The other loadings were .94 for working memory, .83 for processing speed, .77 for verbal memory, and .97 for visual memory in the WAIS-IV/WMS-IV battery and .87 for spatial visualization, .79 for memory, and .74 for speed in the VCAP battery. Importantly, the correlation of the *g* factor representing the highest level in the hierarchy was .91, and it was .96 when the analysis was repeated after omitting the memory variables from both batteries.

[2]CFI refers to the comparative fit index and RMSEA refers to the root mean square error of approximation. As suggested by Kline (2005), CFI values greater than .90 and RMSEA values less than .08 can be interpreted as representing a reasonable fit to the data.

Contextual Analysis Results

The remaining analyses were based on the contextual analysis models portrayed in Figure 1. The reference abilities at the top of the figures are latent variables corresponding to the ability factors in one test battery (with the relations between latent and observed variables as specified in Figure 2), and the target variable is the score in one of the subtests from the other battery. Note that because age is related to both the ability factors and the target variable, the influence of age is controlled at the average age when considering relations of the abilities to the target variable. For the purpose of evaluating the meaning of the target variables, the relations of greatest interest in Figure 1 are represented by the dotted lines between the reference abilities in one test battery and the score on the target test in the other test battery. It is important to note that the fits of the models were primarily determined by the measurement model representing relations between observed variables and reference abilities, and not by the relations between the reference abilities and the target variable. Nevertheless, all the contextual analysis models had respectable fits, with CFI > .88 and RMSEA < .08.

Some latent variables (i.e., Gf and working memory in the Wechsler battery, and reasoning and spatial visualization in the VCAP battery) had very high correlations with one another (i.e., above .80), which resulted in high collinearity when they were entered as simultaneous predictors. Separate analyses were therefore conducted in which only one of the correlated factors served as a reference ability predictor in each analysis. Coefficients from the contextual analyses with the WAIS-IV/WMS-IV subtest scores as target variables and the VCAP reference constructs are reported in Table 3, and the coefficients with the VCAP subtest scores as target variables and the WAIS-IV/WMS-IV reference constructs are presented in Table 4. Although not reported in the table, the patterns were very similar when the analyses were repeated after deleting the VCAP tests (i.e., Vocabulary, Digit Symbol, Logical Memory) derived from earlier versions of the Wechsler subtests.

The contextual analysis results with the VCAP reference constructs predicting the subtest scores from the WAIS-IV/WMS-IV battery were generally consistent with expectations based on assumptions about what the subtests represent. That is, the strongest relations on each variable were from the ability construct it was postulated to assess. For example, the Wechsler verbal memory variables were most closely related to the VCAP memory factor, and the Wechsler speed variables were most closely related to the VCAP speed factor. Scores on subtests postulated to assess perceptual reasoning (i.e., block design, matrix reasoning, visual puzzles, and figure weights) had strong relations with the VCAP reasoning construct, with the block design and visual puzzles measures also having strong relations with the VCAP spatial visualization construct.

However, scores on the Wechsler Vocabulary and Information subtests had weak relations with the VCAP vocabulary construct, and the relations were even smaller with the Similarities and Comprehension measures. Surprisingly, there was no unique relation of either the VCAP reasoning or spatial visualization constructs, but a moderate relation of the speed construct, on the Wechsler Picture Completion variable. This pattern suggests that the Picture Completion test may not reflect the same aspects of the Gf construct as the other tests used to represent that construct.

Scores on the Wechsler subtests postulated to assess working memory (i.e., Digit Span, Arithmetic, and Letter Number Sequencing), including three component measures from the Digit Span test, and scores on the subtests assessing visual memory (i.e., Immediate and Delayed Visual Reproduction and Immediate Designs) and visual working memory (i.e., Symbol Span and Spatial Addition), all had strong relations with the VCAP reasoning construct and weaker relations with the VCAP spatial visualization construct. It is noteworthy that with the exception of the Delayed Designs measure, the visual memory measures all had stronger relations with the VCAP reasoning construct than with the VCAP memory construct.

The contextual analysis results with the WAIS-IV/WMS-IV reference constructs as predictors of the VCAP measures are reported in Table 4. As expected, the reasoning and spatial visualization measures had their strongest relations with the Wechsler Gf construct, and the strongest relations on the VCAP memory measures were with the Wechsler verbal memory construct. The speed measures had strong influences from the Wechsler processing speed construct, although there was also an influence of the Wechsler Gf construct on the pattern comparison measure. The Gf influence on the vocabulary measures was not expected, nor was the lack of a Gc relation on Picture Vocabulary in the analysis in which Gf, but not WM, was a predictor. The lack of an influence of the Wechsler verbal memory factor on the VCAP Logical Memory measure was also surprising, as was the Gf influence on the Paired Associates measure.

As noted earlier, a limitation of analyses with correlated predictors is that influences shared across multiple predictors are not explicitly evaluated. However, these influences can be assessed by specifying an additional common factor that is orthogonal to the specific ability factors, as portrayed in the right panel of Figure 1.

All the measures in each battery were significantly related to the common factor for that battery. An indication of the composition of the common factor is available from the average loadings of the measures representing different abilities on the common factor. The average standardized coefficients in the combined Wechsler battery were .54 for Gc, .75 for Gf, .67 for processing speed, .72 for working memory, .59 for verbal memory, and .78 for visual memory and in the VCAP battery they were .39 for vocabulary, .83 for reasoning, .70 for spatial visualization, .59 for memory, and .59 for speed. These results suggest that although the common factors were broad, they were most strongly influenced by the Gf and reasoning variables in their respective batteries.

Contextual analysis results with the orthogonal common factor model are reported in Table 5 with VCAP reference abilities and Wechsler subtest scores as target variables, and in Table 6 with Wechsler reference abilities and VCAP scores as target variables. The entries in Table 5 reveal that all Wechsler subtest scores except the Delayed Verbal Paired Associates score were significantly related to the VCAP common factor, and that the unique influences were generally as expected, particularly on the speed and verbal memory subtests. The Vocabulary and Information subtests had moderate influences of the VCAP vocabulary construct, but as was the case in the correlated predictor analyses in Table 3, the VCAP vocabulary relations were weak on the Similarities and Comprehension measures.

Probably because the VCAP reasoning factor was absorbed by the common factor, there were no unique relations of the VCAP reasoning factor on the perceptual reasoning subtests (i.e., Block Design, Matrix Reasoning, Visual Puzzles, Figure Weights), but the Block Design and Visual Puzzles measures had strong unique influences of the VCAP spatial visualization construct. Consistent with the correlated predictors' results, there was no significant VCAP reasoning or spatial visualization relation, but a significant speed relation, on the Picture Completion measure. Relations on the subtests postulated to represent working memory (i.e., Digit Span, Arithmetic, and Letter Number Sequencing) and visual memory (i.e., Immediate and Delayed Visual Reproduction and Designs) were inconsistent, possibly because they had considerable influences of the VCAP common factor.

The results in Table 6 indicate that only the Picture Vocabulary variable among the VCAP vocabulary variables had a significant relation with the Wechsler common factor. This measure was only weakly related to the Wechsler Gc construct, but the other VCAP vocabulary measures had significant influences of both the Gc and Gf constructs. As expected, the VCAP reasoning and spatial visualization measures had significant unique influences of the Wechsler Gf construct, the VCAP memory measures had significant influences of the Wechsler verbal memory construct, and the VCAP speed measures had significant influences of the Wechsler processing speed construct. In addition, and consistent with the correlated predictors' results, the Paired Associates memory measure and the Pattern Comparison speed measure also had significant influences of the Wechsler Gf construct.

Discussion

Investigation of cross-battery correspondence is important to determine the comparability of the cognitive ability constructs assessed with different tests. That is, scientific progress would be impeded if what were postulated to be theoretically similar abilities assessed in different cognitive batteries were found to have only weak relations with one another. The results of this study revealed a strong correspondence between two cognitive test batteries at the level of composite scores representing distinct cognitive abilities, with correlations ranging from .84 to .88, and essentially 1.0 after adjustment for unreliability. Furthermore, there was nearly complete overlap of the *g* factor derived from hierarchical analyses in the two batteries, which is consistent with earlier reports that different combinations of tests yield correlations between general factors very close to 1.0 (e.g., Johnson et al., 2004; Johnson et al., 2008; Keith et al., 2001; Stauffer, Ree, & Carretta, 1996). These results indicate that when examined across multiple measures, the two test batteries are assessing nearly identical dimensions of individual differences for relevant abilities, extending to the highest level in the ability hierarchy.

The primary novel contribution in this study was the detailed examination of the relation of individual variables to theoretically relevant abilities across the two test batteries with the two types of contextual analyses portrayed in Figure 1. The correlated predictors' procedure portrayed in the left panel of the figure has the advantage of indicating unique influences of each ability, but it has the disadvantage that the regression coefficients may not be meaningful when there is high collinearity among the predictors. The orthogonal common factor model portrayed in the right panel of the figure specifies a separate factor to account for shared influences, but at the cost of absorbing factors (e.g., reasoning) whose constituent measures have substantial influences of the common factor.

Despite the different strengths and weaknesses of the procedures, the overall pattern of results was similar with both types of contextual analyses as most of the subtest scores in both the WAIS-IV/WMS-IV and VCAP batteries had relations consistent with the expectations that the variables used to define each factor were primarily determined by the theoretically relevant ability. For example, there were strong relations of the VCAP reasoning and spatial visualization constructs on the Wechsler Gf measures, of the VCAP memory construct on the Wechsler verbal memory measures, and of the VCAP speed construct on the Wechsler processing speed measures. The Wechsler tests postulated to assess verbal comprehension (i.e., Vocabulary, Information, Similarities, and Comprehension) were influenced by the VCAP vocabulary construct but they also had influences of the common factor. The other subtests had a mixture of influences with moderate to strong influences from the common factor.

The major exception to the expected pattern in both sets of contextual analyses was the Picture Completion subtest in the WAIS-IV battery which had a near zero relation with the VCAP reasoning and spatial visualization constructs, and a moderate relation with the VCAP speed construct. Although there is a time limit of 20 seconds per item, the picture completion test is typically not considered a speeded test because 20 seconds has been assumed to be a sufficient amount of time to respond. Nevertheless, the speed influence, together with the lack of either a reasoning or spatial visualization relation, suggests that it may reflect something distinct than the other tests used to define the Wechsler Gf factor. This interpretation is consistent with the findings of weak loadings of the Picture Completion test on the Gf factor (or the perceptual reasoning factor in the Wechsler terminology) in other analyses (e.g., Benson, Hulac, & Kranzler, 2010; Salthouse, 2009). However, before concluding that the test represents something distinct, it would be desirable to examine the Picture Completion subtest with other combinations of cognitive tests to determine its relations with alternative sets of abilities.

Two other results from Tables 3 and 5 are noteworthy because they are relevant to the interpretation of the factors in the WAIS-IV/WMS-IV battery. One result is the strong relations of the Block Design and Visual Puzzles measures with the VCAP spatial visualization measure. This finding suggests that the WAIS-IV/WMS-IV Gf factor (i.e., perceptual reasoning in WAIS IV terminology) is a mixture of reasoning (as represented by the Matrix Reasoning and Figure Weights measures) and spatial visualization (as represented by the Block Design and Visual Puzzles measures). Although the reasoning and spatial visualization factors are highly correlated in the VCAP data (i.e., correlations of .73 and .74 in Table 2), the WAIS-IV/WMS-IV perceptual reasoning factor appears to be a hybrid of two conceptually distinct factors.

The second interesting result from Tables 3 and 5 is that the visual memory measures from the WAIS-IV/WMS-IV battery had stronger relations with the VCAP reasoning factor than with the VCAP memory factor. This may be partly attributable to shared method variance, as the VCAP reasoning tests were all visually presented whereas the VCAP memory tests were all presented auditorially. However, it is also possible that encoding and maintaining visual–spatial information requires some of the same processes involved in reasoning, such as detecting and integrating relations among elements. Future research with different types of visual memory tests and reasoning tests involving different modalities and materials would be valuable in distinguishing these possibilities.

To summarize, in addition to reporting comparisons at the level of ability factors as in earlier studies, a new analytical procedure was used to investigate the correspondence of cognitive test batteries between factors in one battery and individual variables in the other battery. Examination of the relations between batteries at this more detailed level provides valuable information about construct validity in terms of the patterns of relations of the indicator variables with relevant ability constructs. Although with most of the measures the relations were consistent with expectations, the results of the analyses lead to questions about the appropriateness of the picture completion subtest as a measure of Gf (or perceptual reasoning) in the WAIS-IV battery, and the composition of the perceptual reasoning factor.

Acknowledgments

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Award Number R37 AG-024270 from the National Institute on Aging.

References

Arbuckle, JL. AMOS [Computer software]. Chicago, IL: SPSS; 2006.

Bennett, GK.; Seashore, HG.; Wesman, AG. Differential aptitude test. San Antonio, TX: Psychological Corporation; 1997.

Benson N, Hulac DM, Kranzler JH. Independent examination of the Wechsler Adult Intelligence Scale–Fourth edition (WAIS-IV): What does the WAIS-IV measure? Psychological Assessment. 2010; 22:121–130. [PubMed: 20230158]

Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin. 1959; 56:81–105. [PubMed: 13634291]

Ekstrom, RB.; French, JW.; Harman, HH.; Dermen, D. Manual for kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service; 1976.

Folstein MF, Folstein SE, McHugh PR. “Mini-Mental State”. A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975; 12:189–198. [PubMed: 1202204]

Johnson W, Bouchard TJ, Krueger RF, McGue M, Gottesman II. Just one g: Consistent results from three test batteries. Intelligence. 2004; 32:95–107.

Johnson W, te Nijenhuis J, Bouchard TJ. Still just 1 *g*: Consistent results from five test batteries. Intelligence. 2008; 36:81–95.

Keith TZ, Kranzler JH, Flanagan DP. What does the Cognitive Assessment System (CAS) measure? Joint confirmatory factor analysis of the CAS and the Woodcock-Johnson Tests of Cognitive Ability (3rd Edition). School Psychology Review. 2001; 30:89–119.

Kline, RB. Principles and practice of structural equation modeling. 2. New York, NY: Guilford Press; 2005.

Messick, S. Validity. In: Linn, RL., editor. Educational measurement. 3. New York, NY: MacMillan; 1989. p. 13-103.

Raven, J. Advanced progressive matrices, Set II. London, England: H. K. Lewis; 1962.

Salthouse TA. Speed and knowledge as determinants of adult age differences in verbal tasks. Journal of Gerontology: Psychological Sciences. 1993; 48:P29–P36.

Salthouse TA. Relations between cognitive abilities and measures of executive functioning. Neuropsychology. 2005; 19:532–545. [PubMed: 16060828]

Salthouse TA. Decomposing age correlations on neuropsychological and cognitive variables. Journal of the International Neuropsychological Society. 2009; 15:650–661. [PubMed: 19570312]

Salthouse TA, Babcock RL. Decomposing adult age differences in working memory. Developmental Psychology. 1991; 27:763–776.

Salthouse TA, Fristoe N, Rhee SH. How localized are age-related effects on neuropsychological measures? Neuropsychology. 1996; 10:272–285.

Salthouse TA, Pink JE, Tucker-Drob EM. Contextual analysis of fluid intelligence. Intelligence. 2008; 36:464–486. [PubMed: 19137074]

Salthouse TA, Siedlecki KL, Krueger LE. An individual differences analysis of memory control. Journal of Memory and Language. 2006; 55:102–125.

Stauffer JM, Ree MJ, Carretta TR. Cognitive-components tests are not much more than *g*: An extension of Kyllonen's analyses. Journal of General Psychology. 1996; 123:193–205.

Wechsler, D. Wechsler Adult Intelligence Scale. 3. San Antonio, TX: Psychological Corporation; 1997a.

Wechsler, D. Wechsler Memory Scale: Third edition. San Antonio, TX: Psychological Corporation; 1997b.

Wechsler, D. WAIS-IV: Administration and scoring manual. San Antonio, TX: Pearson; 2008.

Wechsler, D. WMS-IV: Administration and scoring manual. San Antonio, TX: Pearson; 2009.

Wechsler, D.; Coalson, DL.; Raiford, SE. WAIS-IV: Technical and interpretive manual. San Antonio, TX: Pearson; 2008.

Wechsler, D.; Holdnack, JA.; Drozdick, LW. WMS-IV: Technical and interpretive manual. San Antonio, TX: Pearson; 2009.

Woodcock, RW.; Mather, N. Woodcock-Johnson Tests of Cognitive Ability. Standard and supplemental batteries: Examiner's manual. Allen, TX: DLM; 1990.

Zachary, RA. Shipley Institute of Living Scale–Revised. Los Angeles, CA: Western Psychological Services; 1986.

Appendix

Description of Virginia Cognitive Aging Project Variables and Sources of Tasks.

Variable	Description	Source
Wechsler Vocabulary	Provide definitions of words	Wechsler (1997a)
Picture Vocabulary	Name the pictured object	Woodcock and Mather (1990)
Antonym Vocabulary	Select the best antonym of the target word	Salthouse (1993)
Synonym Vocabulary	Select the best synonym of the target word	Salthouse (1993)
Matrix Reasoning	Determine which pattern best completes the missing cell in a matrix	Raven (1962)
Shipley Abstraction	Determine the words or numbers that are the best continuation of a sequence	Zachary (1986)
Letter Sets	Identify which of five groups of letters is different from the others	Ekstrom, French, Harman, and Dermen (1976)
Spatial Relations	Determine the correspondence between a 3-D figure and alternative 2-D figures	Bennett, Seashore, and Wesman (1997)
Paper Folding	Determine the pattern of holes that would result from a sequence of folds and a punch through folded paper	Ekstrom et al. (1976)
Form Boards	Determine which combinations of shapes are needed to fill a larger shape	Ekstrom et al. (1976)

Variable	Description	Source
Logical Memory	Number of idea units recalled across three stories	Wechsler (1997b)
Word Recall	Number of words recalled across Trials 1 to 4 of a word list	Wechsler (1997b)
Paired Associates	Number of response terms recalled when presented with a stimulus term	Salthouse, Fristoe, and Rhee (1996)
Digit Symbol	Use a code table to write the correct symbol below each digit	Wechsler (1997a)
Letter Comparison	Same/different comparison of pairs of letter strings	Salthouse and Babcock (1991)
Pattern Comparison	Same/different comparison of pairs of … line patterns	Salthouse and Babcock (1991)

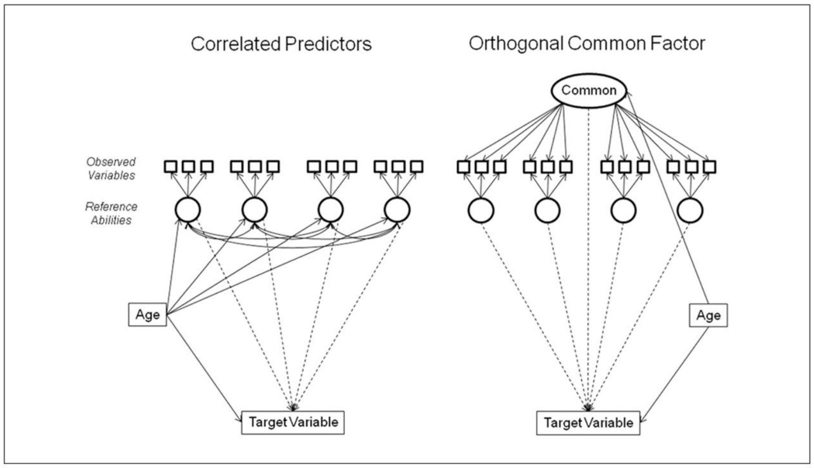

Figure 1. Schematic illustration of contextual analysis procedures with two different analytical models

Note. In each case, latent constructs corresponding to cognitive abilities (portrayed as circles) are defined by observed variables (portrayed as squares), and they are used as simultaneous predictors of the target variable (portrayed as a square). In addition, age variation in the reference abilities and the target variable is controlled by specifying relations of age to the ability constructs and to the target variable. The latent variables are correlated with one another on the model on the left, and all observed variables are related to a common factor on the model on the right.

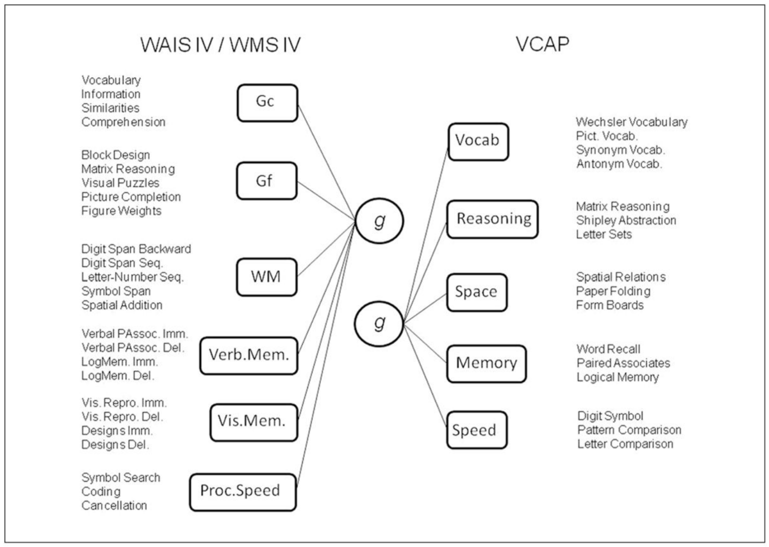

Figure 2. Subtest scores used to represent each ability construct in the two test batteries. See Salthouse (2009) for details of the analyses used to establish these structures

Note. WAIS-IV = Wechsler Adult Intelligence Scale–IV; WMS-IV = Wechsler Memory Scale–IV; VCAP = Virginia Cognitive Aging Project; Gc = crystallized ability; Gf = fluid ability; WM = working memory.

Table 1

Means (Standard Deviations) and Effect Sizes (*d*) for Comparisons Across Samples.

Variable	Linkage	WAIS-IV/WMS-IV	VCAP	*d*
N	90	900–2,200	2,973–3,796	—
Age (years)	50.8 (15.2)	45.0 (22.9)	50.9 (18.4)	—
Proportion female	0.68	0.52	0.65	—
Years of education	15.6 (3.1)	NA	15.7 (2.7)	—
Vocabulary	42.8 (10.5)	34.6 (11.1)		.78
Information	16.3 (5.9)	13.5 (5.3)		.47
Similarities	27.5 (5.7)	23.9 (5.8)		.63
Comprehension	26.6 (5.4)	22.9 (6.2)		.69
Block Design	38.5 (12.5)	38.8 (13.7)		−.02
Matrix Reasoning	17.9 (4.5)	16.0 (5.6)		.42
Visual Puzzles	14.4 (4.9)	13.6 (5.1)		.16
Picture Completion	11.3 (3.7)	12.1(4.4)		−.22
Figure Weights	14.4 (4.9)	14.5 (5.0)		−.02
Digit Span	28.9 (5.2)	26.8 (6.1)		.40
Arithmetic	15.5 (3.5)	13.6 (3.7)		.54
Letter Number Sequencing	20.3 (3.1)	19.6 (3.8)		.23
Symbol Search	31.4 (8.8)	29.5 (9.6)		.22
Coding	67.2 (16.3)	62.8 (19.7)		.27
Cancellation	35.9 (9.0)	39.6 (9.7)		−.41
Full-Scale IQ	107.4 (16.3)	100.0 (15.0)		−.49
Verbal Comprehension Index	109.2 (17.3)	100.0 (15.0)		−.61
Perceptual Reasoning Index	103.9 (15.3)	100.0 (15.0)		−.26
Working Memory Index	106.1 (14.3)	100.0 (15.0)		−.41
Perceptual Speed Index	104.2 (15.8)	100.0 (15.0)		−.28
Visual Reproduction Immediate	35.6 (5.5)	33.2 (7.7)		.44
Visual Reproduction Delayed	26.6 (9.6)	23.0 (10.6)		.38
Designs Total Immediate	69.7 (15.6)	74.6 (16.9)		−.31
Designs Total Delayed	58.5 (14.1)	60.3 (16.4)		−.13
Verbal Paired Associates Immediate	37.9 (11.5)	29.2 (12.7)		.76
Verbal Paired Associates Delayed	11.4 (2.9)	8.8 (3.8)		.90
Logical Memory Immediate	29.1 (6.5)	26.8 (7.9)		.35
Logical Memory Delayed	25.3 (7.7)	20.1 (8.2)		.68
Symbol Span	24.4 (7.4)	21.0 (8.7)		.46
Spatial Addition	11.9 (5.1)	14.0 (4.8)		−.41
Auditory Memory Index	109.8 (15.2)	100.0 (15.0)		−.65
Visual Memory Index	103.3 (15.1)	100.0 (15.0)		−.22
Immediate Memory Index	106.5 (15.4)	100.0 (15.0)		−.43
Delayed Memory Index	108.7 (15.5)	100.0 (15.0)		−.58
Vocabulary	48.9 (9.4)	50.5 (10.5)		−.17

Variable	Linkage	WAIS-IV/WMS-IV	VCAP	*d*
Picture Vocabulary	16.0 (5.8)	18.3 (5.4)		−.40
Synonym Vocabulary	6.4 (2.9)	7.1 (2.7)		−.24
Antonym Vocabulary	6.0 (3.1)	6.6 (2.8)		−.19
Ravens	6.9 (3.3)	7.8 (3.3)		−.27
Shipley Abstraction	12.9 (3.7)	13.2 (3.6)		−.08
Letter Sets	10.7 (2.7)	11.0 (2.9)		−.11
Spatial Relations	8.3 (5.0)	8.7 (5.0)		−.08
Paper Folding	5.8 (2.7)	6.2 (2.8)		−.15
Form Boards	8.3 (5.0)	7.1 (4.2)		.24
Word Recall	33.6 (6.9)	34.7 (6.5)		−.16
Paired Associates	3.0 (1.9)	3.0 (1.8)		.00
Logical Memory	44.1 (7.4)	44.1 (10.1)		.00
Digit Symbol	69.7 (15.6)	72.1 (18.0)		−.15
Pattern Comparison	15.0 (3.6)	16.2 (3.8)		−.33
Letter Comparison	10.3 (2.7)	10.5 (2.5)		−.07

Note. WAIS-IV = Wechsler Adult Intelligence Scale–IV; WMS-IV = Wechsler Memory Scale–IV; VCAP = Virginia Cognitive Aging Project

Table 2

Correlations of Composite Scores in VCAP and WAIS-IV/WMS-IV Samples.

	VCAP					WAIS-IV/WMS-IV					
	Vocab	**Reas**	**Space**	**Memory**	**Speed**	**Gc**	**Gf**	**WM**	**VbM**	**VisM**	**Speed**
Vocab	X	.64*	.46*	.43*	.23*	**.88***	.56*	.28*	.44*	.21	.26*
Reas	.43*	X	.74*	.65*	.55*	.72*	**.87***	.74*	.69*	.69*	.62*
Space	.32*	.73*	X	.50*	.44*	.48*	**.84***	.61*	.45*	.62*	.51*
Memory	.37*	.63*	.52*	X	.47*	.58*	.70*	.64*	**.85***	.65*	.48*
Speed	.13*	.60*	.52*	.53*	X	.38*	.68*	.62*	.35*	.56*	**.85***
Gc						X	.64*	.38*	.56*	.28*	.40*
Gf						.58*	X	.80*	.63*	.73*	.65*
WM						.48*	.72*	X	.54*	.70*	.58*
VbM						.48*	.47*	.51*	X	.62*	.42*
VisM						.27*	.67*	.62*	.48*	X	.61*
Spd						.36*	.60*	.57*	.39*	.52*	X

Note. VCAP = Virginia Cognitive Aging Project; WAIS-IV = Wechsler Adult Intelligence Scale–IV; WMS-IV = Wechsler Memory Scale–IV; Gc = crystallized ability; Gf = fluid ability; WM = working memory; VbM = verbal memory; VisM = visual memory. The constructs are defined by the variables listed in Figure 1. Values above the diagonal are from the Linkage sample, values below the diagonal in the upper left quadrant are from the VCAP sample, and values below the diagonal in the lower right quadrant are from the WAIS-IV/WMS-IV sample. All correlations are significant at $p < .01$. Values in boldface are correlations between theoretically comparable constructs in the two batteries.

* $p < .01$.

Table 3

Contextual Analysis Results on WAIS-IV/WMS-IV Subtests With VCAP Reference Abilities.

	Vocab	Reas	Space	Memory	Speed
WAIS-IV					
Vocabulary	.28/.40	.34	.01	.20/.29	−.23/−.10
Information	.44/.48*	.14	.01	.34/.38	−.22/−.17
Similarities	.11/.24	.45	.09	.13/.25	−.12/.04
Comprehension	.08/.22	.44	.08	.15/.29	−.16/−.04
Block Design	−.21/−.07	1.23*	.74*	−.11/.18	−.23/.12
Matrix Reasoning	−.29/−.05	1.03*	.39*	.05/.30	−.31/.00
Visual Puzzles	−.06/−.13	.92*	.92*	−.15/.03	−.15/.05
Picture Completion	.15/.04	−.18	.08	.40/.35	.68*/.61*
Figure Weights	.01/.20	.85*	.33	.01/.21	−.25/.01
Symbol Search	−.33/−.39*	.17	.26	.16/.19	.87*/.88*
Coding	−.19/−.18	−.14	−.13	.17/.16	.99*/.93*
Cancellation	−.20/−.28	−.08	.10	.08/.07	1.10*/1.05*
Digit Span	−.41/−.18	.89*	.29	.13/.36	−.12/.17
Arithmetic	−.08/.10	.79*	.29	−.07/.12	−.00/.25
Letter Number	−.50/−.31	.70	.23	.33/.50	−.02/.20
Digit Span For.	−.20/−.12	.67	.39	−.07/.09	.06/.25
Digit Span Bck.	−.28/−.11	.51	.09	.29/.42	−.19/−.02
Digit Span Seq.	−.49/−.21	.96*	.21	.08/.34	−.16/.17
WMS-IV					
Verb. PAssoc. Imm	−.33/−.21	.33	.04	.80*/.87*	−.41/−.29
Verb. PAssoc. Del	−.38/−.29	.29	.08	.66*/.72*	−.29/−.20
Log. Mem. Imm	−.11/−.17	−.15	−.02	1.15*/1.11*	−.28/−.34
Log. Mem. Del	−.30/−.18	.23	−.04	.86*/.92*	−.33/−.25
Visual Repro. Imm	−.35/−.15	1.00*	.41*	.02/.28	−.05/.26

	Vocab	Reas	Space	Memory	Speed
Visual Repro. Del	−.56*/−.48*	.63	.36*	.48*/.62*	−.03/.15
Designs Imm	−.57*/−.47*	.65*	.34	.54*/.70*	−.02/.17
Designs Del	−.29/−.29	.16	.13	.72*/.76*	−.01/.02
Symbol Span	−.73*/−.55*	.83*	.36*	.52*/.71*	−.04/.21
Spatial Arithmetic	−.33/−.30	.71*	.52*	.18/.34	−.05/.13

Note. VCAP = Virginia Cognitive Aging Project; WAIS-IV = Wechsler Adult Intelligence Scale IV; WMS-IV = Wechsler Memory Scale IV. The first coefficient is when reasoning but not space was in the analysis, and the second is when space but not reasoning was in the analysis.

* $p < .01$.

Table 4

Contextual Analysis Results on VCAP Tests With WAIS-IV/WMS-IV Reference Abilities.

	Gc	Gf	WM	VbM	VisM	Speed
Wechsler Vocabulary	.56*/.99*	.60	−.68*	−.41*/−.51*	.06/.80*	−.44*/−.30
Picture Vocab.	.03/.57*	1.02*	−.42	−.19/−.33	−.08/.68*	−.35/−.14
Synonym Vocab.	.35/.77*	.64	−.52	−.26/−.37*	−.18/.48	−.15/−.00
Antonym Vocab.	.38/.84*	.84*	−.43	−.26/−.39*	−.26/.46	−.29/−.11
Matrix Reasoning	−.26/.27	1.20*	−.13	−.03/−.20	−.05/.65*	−.29/−.10
Shipley Abstraction	−.12/.30	1.15*	.18	−.07/−.22	−.30/.24	.09/.20
Letter Sets	−.36/.15	1.30*	.14	.09/−.06	−.23/.49	−.15/−.06
Spatial Relations	−.41*/.37	1.85*	−.04	.07/−.21	−.73*/.34	−.23/.00
Paper Folding	−.44*/.19	1.54*	.12	−.10/−.33	−.07/.72*	−.43*/−.27
Form Boards	−.49*/.06	1.41*	.03	−.09/−.23	−.10/.61*	−.21/.01
Word Recall	−.28/−.11	.26	−.15	.36*/.33	.34/.54	−.04/.03
Paired Associates	−.26/.03	.80*	.22	.50*/.37*	−.22/.13	−.21/−.16
Logical Memory	.39/.22	.03	.58	−.15/−.11	.59/.23	−.27/−.29
Digit Symbol	−.07/−.02	−.02	−.21	−.21/−.21	−.16/−.06	1.05*/1.11*
Pattern Comparison	−.30/.02	.71*	−.07	−.21/−.31*	−.34/.06	.49*/.61*
Letter Comparison	−.24/−.11	.23	−.07	−.12/−.14	−.06/.07	.87*/.92*

Note. Separate coefficients for analyses with Gf and with WM. VCAP = Virginia Cognitive Aging Project; WAIS-IV = Wechsler Adult Intelligence Scale–IV; WMS-IV = Wechsler Memory Scale–IV; Gc = crystallized ability; Gf = fluid ability; WM = working memory; VbM = verbal memory; VisM = visual memory.

* $p < .01$.

Table 5

Standardized Coefficients With the Orthogonal Common Factor Model in the Right Panel of Figure 1 With VCAP Reference Abilities.

	Common	Vocab	Reas	Space	Memory	Speed
WAIS-IV						
Vocabulary	.34*	.30*	.25	.10	.22	−.11
Information	.39*	.41*	.20	.17	.36*	−.08
Similarities	.43*	.15	.20	.09	.16	−.03
Comprehension	.39*	.09	.20	.08	.15	−.07
Block Design	.87*	−.08	.05	.23*	−.01	−.02
Matrix Reasoning	.61*	−.09	.15	.06	.04	−.12
Visual Puzzles	.71*	−.04	.05	.44*	−.02	.00
Picture Compl.	.77*	−.01	−.04	−.05	.19	.37*
Figure Weights	.65*	−.06	.27	.09	.10	−.06
Symbol Search	.51*	−.07	.18	.18	.09	.43*
Coding	.43*	.02	.06	−.07	.04	.42*
Cancellation	.63*	−.06	−.11	−.06	−.02	.49*
Digit Span	.54*	−.20	.26	.04	.09	−.01
Arithmetic	.58*	.01	.25	.11	.02	.07
Letter Number	.63*	−.34*	.11	−.02	.16	.05
Digit Span For.	.48*	−.09	.12	.14	.01	.10
Digit Span Bck.	.34*	−.16	.23	−.03	.16	−.07
Digit Span Seq.	.45*	−.21	.30	−.02	.05	−.05
WMS-IV						
Verb. PAssoc. Imm	.30*	.07	.13	.03	.44*	−.18
Verb. PAssoc. Del	.19	.03	.11	.06	.35*	−.13
Log. Mem. Imm	.56*	−.19	−.04	−.01	.60*	−.08
Log. Mem. Del	.31*	.02	.18	−.00	.49*	−.16

	Common	Vocab	Reas	Space	Memory	Speed
Visual Repro. Imm	.65*	−.11	.21	.09	.04	.01
Visual Repro. Del	.54*	−.17	.10	.01	.21	−.01
Designs Imm	.80*	−.25*	−.17	−.16	.15	.01
Designs Del	.57*	−.15	−.06	−.03	.34*	.01
Symbol Span	.57*	−.23*	.18	.07	.23	.01
Symbol Arith	.65*	−.24*	−.06	.13	.09	.06

Note. VCAP = Virginia Cognitive Aging Project; WAIS-IV = Wechsler Adult Intelligence Scale–IV; WMS-IV = Wechsler Memory Scale–IV.

* $p < .01$.

Table 6

Standardized Coefficients With the Orthogonal Common Factor Model in the Right Panel of Figure 1 With WAIS-IV/WMS-IV Reference Abilities.

	Common	Gc	Gf	WM	VbM	VisM	Speed
Wechsler Vocabulary	.03	.56*	.38*	−.19	−.17	.35	−.13
Picture Vocab.	.37*	.17	.41*	−.28	−.13	.02	−.14
Synonym Vocab.	.21	.35*	.39*	−.23	−.11	.08	.00
Antonym Vocab.	.22	.41*	.50*	−.18	−.11	.13	−.06
Matrix Reasoning	.53*	.03	.45*	−.21	−.04	.14	−.13
Shipley	.71*	.12	.27*	−.08	−.13	−.01	.05
Letter Sets	.66*	−.07	.28	−.06	.04	−.01	−.05
Spatial Relations	.44*	.03	.79*	−.24	.03	.01	−.01
Paper Folding	.55*	−.16	.56*	−.14	−.04	.00	−.15
Form Boards	.55*	−.08	.49*	−.13	−.12	−.02	−.09
Word Recall	.52*	−.09	.08	−.15	.22*	.17	−.03
Paired Associates	.48*	−.14	.28*	−.07	.40*	.01	−.02
Logical Memory	.63*	−.19	.21	.02	.26*	.00	.05
Digit Symbol	.47*	.04	−.07	−.10	−.18	.00	.56*
Pattern Comparison	.29*	−.06	.27*	−.10	−.16	−.17	.32*
Letter Comparison	.53*	−.03	.02	−.05	−.13	−.07	.47*

Note. WAIS-IV = Wechsler Adult Intelligence Scale–IV; WMS-IV = Wechsler Memory Scale–IV; Gc = crystallized ability; Gf = fluid ability; WM = working memory; VbM = verbal memory; VisM = visual memory.

* $p < .01$.