Center for Demography and Ecology

University of Wisconsin-Madison

The Dimensionality and Measurement of Cognitive Functioning at age 65 in the Wisconsin Longitudinal Study

James A. Yonker
Robert M. Hauser
Jeremy Freese

CDE Working Paper No. 2007-06

The Dimensionality and Measurement of Cognitive Functioning at Age 65 in the Wisconsin Longitudinal Study[1]

James A. Yonker, Robert M. Hauser, and Jeremy Freese

Center for Demography of Health and Aging
University of Wisconsin-Madison

October 8, 2007

[1]This research uses data from the Wisconsin Longitudinal Study (WLS) of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775 and AG-21079), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://www.ssc.wisc.edu/wlsresearch/. The opinions expressed herein are those of the authors. We wish to thank Joseph R. Savard for his graphical expertise and Emily P. Nell and Jesse M. Philbin for their assistance editing and proofing preliminary versions of this paper.

Abstract

The 2003-05 telephone surveys of high school graduates in the Wisconsin Longitudinal Study administered six cognitive assessments: immediate and delayed word recall, digit ordering, a subset of the WAIS-R similarities test, and letter and category frequency. We have analyzed these data separately among male and female participants in the WLS. We find that a structural model with a single, second order factor for general cognitive functioning fits the data well. The first order factors are memory/attention (word recall and digit ordering), abstract reasoning (WAIS-R), and verbal (letter and category) fluency. In addition, the memory/attention factor loads much more heavily on the general cognition factor among men than among women. We recommend this model be used in other analyses of cognitive functioning in the WLS.

1 Introduction

The Wisconsin Longitudinal Study (WLS) recently released public-use data for the 2003-05 round of interviews, when the primary respondents were ∼65 years old. A substantial addition to this wave of data collection was the expansion of the telephone-administered set of cognitive tests, five more than in the previous wave (1993) when the primary respondents were in their mid-fifties. The purpose of this report is to model the relationships among these six assessments in light of their structure suggested by theory, manifest content, and previous research.

Investigating the causes and consequences of changes in cognitive functioning across the life-course has become a major focus of much research on aging. Fundamental to this aim is measuring and modeling the relationships among sets of tasks thought to assess various interrelated dimensions of cognitive functioning. This endeavor has both substantive and methodological implications for future research.

2 Method

The WLS is a long-term panel study of 4,994 men and 5,323 women who graduated from Wisconsin high schools in 1957. The study design and history have been discussed in detail elsewhere (Hauser 2005; Sewell, Hauser, Springer, and Hauser 2004). Briefly, WLS participants are a one-third, simple random sample of all 1957 Wisconsin high school graduates. The WLS provides a comprehensive record of social background, adolescent aspirations, schooling, military service, family formation, labor market experiences, health, personality, and social participation of the original respondents. The feasibility of using telephone interviews to administer and code cognitive assessments of aging respondents has been discussed elsewhere (Brandt, Spencer, and Folstein 1988; Herzog and Wallace 1997; Herzog and Rodgers 1999; Nesselroade, Pedersen, McClearn, Plomin, and Bergeman 1988), and the technology used in the WLS is described by Hauser and Willis (2005).

2.1 Measures

Immediate and Delayed Word Recall The interviewers told the respondent to recall as many words as possible from a list. The interviewer then read ten high-frequency words, one every two seconds (Brandt, Spencer, and Folstein 1988). Each correct response was recorded by the interviewer, and a short series of questions (approximately 10 minutes) followed, without indication to the participant that the assessment was not yet complete. The intermediate questions served as a distracter task to prevent active rehearsal of the word list. The interviewer then asked the participant to recall as many words from the list as possible and noted each correct response. Table 1 shows descriptive statistics by gender for each of the cognitive measures.

Digit Ordering The digit ordering task was a modification of the WAIS-III digit backward subtest, thought to measure fluid intelligence (Wechsler 1997). Participants started with a set of three digits, read to them at one-second intervals, and had to order them from smallest to largest. The interviewer continued the exercise by adding an additional digit to the task following each correct response, up to eight digits. (Digits were not the same in each round.) Finally, participants exited the task by either 1) incorrectly ordering digits at a given level twice consecutively or 2) correctly ordering all eight digits in the final round. The interviewer recorded the highest number of digits correctly ordered by the respondent. Respondents who incorrectly ordered the first set of three digits were given a score of two.

WAIS-R Similarities Participants were administered either a six- or nine-item subset of the Wechsler Adult Intelligence Scale (WAIS-R) similarities items (Wechsler 1981). Subjects were asked how two things are alike (e.g. an orange and a banana, a poem and a statue), and scores of 0, 1, or 2 were assigned based on the correctness and categorical abstractness of the response in accordance with WAIS-R protocol. For example, in the case of the orange-banana item, “They are both fruits” gets a score of 2, while “They both have skins” gets a score of 1. This task is believed to be primarily a measure of crystallized intelligence.

Letter and Category Fluency The letter and category fluency tasks tap the constructs of verbal and declarative semantic memory. In the letter fluency task, the interviewer asked the participant to name as many words as possible starting with either the letter “L” or the letter “F” in a one-minute period (Borkowski, Benton, and Spreen 1967; Monsch, Bondi, Butters, Paulsen, Salmon, Brugger, and Swenson 1994). The category fluency task is quite similar; participants were asked to name either as many foods or as many animals as possible in one minute (Drachman and Leavitt 1972; Kozora and Cullum 1995; Schooler, Mulatu, and Oates 1999). Responses were recorded digitally and project staff transcribed the correct responses. Proper nouns were excluded from scoring for the letter fluency task, as were repetitions and permutations of words, such as “like” and “liking.” Clearly inappropriate responses were also omitted from the category fluency task and discounted during scoring. A scoring program tallied the number of valid responses in each 15-second interval. In the analysis, scores on these variables were grouped into intervals roughly equal to their standard deviations: 4 (letter fluency) and 6 (category fluency). For example, raw letter fluency scores from 1 to 4 were coded as 1, 5 through 8 coded as 2, etc. Category fluency scores were coded similarly (e.g. 1 to 6 coded as 1, 7 though 12 coded as 2).

2.2 Sample Response

The total WLS sample size is 10,317. In the most recent 2003-05 wave of data collection, 1,292 (13%) cases were confirmed dead and 447 (4%) cases were not fielded for administrative reasons, many (93%) of which were non-respondents to both the 1975 and 1993 waves. Interviews were attempted for the remaining 8,578 (83%) cases. Of these, 7,063 (82%) respondents completed the entire telephone interview, 423 (5%) completed some portion of the interview, 771 (9%) respondents refused to participate, and 321 (4%) cases were otherwise incomplete because of permanent respondent disability (39, 12%), failure to contact the respondent (279, 87%), and respondent request for data deletion (3, 1%). The total effective sample size for telephone interviews in the 2003-05 wave of data collection, including usable partial interviews, is 7,265 (85% of attempted cases).

Participants were randomly assigned to one of three sampling subgroups in order to reduce interview length and control costs. Sixteen hundred (21%) cases were assigned to complete six of the nine WAIS-R abstract reasoning items only. Two thousand fifty-eight (28%) cases were assigned to complete the immediate and delayed word recall tasks, the digit ordering task, all nine WAIS-R abstract reasoning items, and the letter fluency task. The remaining 3,607 (50%) cases were scheduled to complete all of the tasks: immediate and delayed word recall, digit ordering, all nine WAIS-R abstract reasoning items, and both letter and category fluency. Complete data are available for 1,496 (94%) of the cases in the first group, 1,669 (81%) of the cases in the second group, and 2,876 (80%) of the cases in the third group. The total analytic sample size is 6,041 cases for which all of the requested items were completed (83% of entirely or partially completed cases).

2.3 Analysis

The analytic plan for this research was a mixture of confirmatory and exploratory factor analyses. We estimated a series of three classes of models (referred to by their number of first-order latent factors) to empirically investigate the factorial structure of the cognitive battery, primarily informed by the manifest content of the six tests. Each model was estimated first with no correlated errors among observed variables, then with statistically and substantively significant correlated errors implied by patterns of residual variation, allowing for modifications of the overall structure.

Previous research and exploratory analyses suggest potential gender differences in the structure of later-life cognitive functioning, so all models were run separately for males and females. In addition, we used multiple group modeling techniques within gender, equating parameters across sampling subgroups, in order to include as many cases as possible in the analysis. A discussion of this method can be found elsewhere (Jöreskog and Sörbom 1996). Briefly, subgroups with valid data for a given variable contributed to combined parameter estimates related to that variable while subgroups with partial data on some tasks were given fixed coefficients of zero and not used to estimate these parameters. Because the subgroup sampling was completely random, this should not bias either individual parameter estimates or overall model fit (Allison 1987; Allison and Hauser 1991).

Models were estimated in LISREL 8, and we present here their specification in corresponding LISREL notation.

$$\boldsymbol{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon} \tag{1}$$

$$\boldsymbol{\eta} = \boldsymbol{\Gamma} \boldsymbol{\xi} + \boldsymbol{\zeta} \tag{2}$$

$\boldsymbol{y}$	is a $(p \times 1)$ vector of observable variables
$\boldsymbol{\Lambda}_y$	is a $(p \times m)$ matrix of first-order factor loadings
$\boldsymbol{\eta}$	is an $(m \times 1)$ vector of first-order latent factors
$\boldsymbol{\epsilon}$	is a $(p \times 1)$ vector of unique components
$\boldsymbol{\Gamma}$	is an $(m \times n)$ matrix of second-order factor loadings
$\boldsymbol{\xi}$	is an $(n \times 1)$ vector of second-order latent factors
$\boldsymbol{\zeta}$	is an $(n \times 1)$ vector of unique components

Models 1A and 1B are defined by equation 1 alone, having only one first-order latent factor and no second-order latent factors. Models 3A, 3B, and 4 are defined jointly by equations 1 and 2.

Overall model fit was evaluated by two widely accepted standards, the Bayesian Information Criterion (BIC) and the Root Mean Square Error of Approximation (RMSEA). BIC is based on information theory and is a relative measure of fit (i.e., with no absolute interpretation) that corrects for the overall degrees of freedom and number of cases, favoring parsimonious models by penalizing models with additional parameters. BIC is less sensitive to large sample sizes than traditional measures of fit based solely on χ^2 and degrees of freedom (Raftery 1995). There are several equivalent expressions of BIC. We used the following formula, where models with a larger negative BIC value are preferred.

$$BIC = \chi^2 - (df) \times \ln(n) \tag{3}$$

Raftery suggests that absolute differences in BIC values from 2 to 6 provide moderate support for the preferred model, 6 to 10 strong support, and >10 very strong support. We also used RMSEA, where acceptably fitting models have values of less than 0.05 (Loehlin 2004). The formula for RMSEA is shown in equation 4.

$$RMSEA = \sqrt{\frac{\chi^2/(df-1)}{(n-1)}} \tag{4}$$

We present summary information for each model but only present detailed parameter estimates for the best fitting model according to BIC and RMSEA values.

Because several of the variables are measured as ordered categories, we estimated these models by weighted least-squares using polychoric correlation and asymptotic variance-covariance matrices (Jöreskog and Sörbom 1996). Tables 1 and 2 present univariate descriptive statistics and observed pair-wise polychoric correlations by gender. The sample sizes in Table 1 can be used to derive the number of cases contributing to each pair-wise polychoric correlation in Table 2. For example, 2,056 males completed the immediate word recall task and 1,316 males completed the category fluency task. Therefore, 1,316 cases (the smaller of the two sample sizes) contributed to the estimated polychoric correlation between the two, $r(14,1) = 0.19$, from the matrix in Table 2.

First, we constructed a base model (Model 1A, Figure 1) where all observable variables loaded on a single latent factor. Each observable and the single latent factor also had an independent stochastic disturbance. The dashed line in the figure denotes a loading fixed at 1.0 as a normalizing constraint, that is, to fix the metric of the latent variable.[2] Modification indices between immediate recall, delayed recall, and digit ordering and between letter and category fluency implied plausible correlations between errors in observables. These substantially improved the overall model fit and were incorporated in a modification of the first model (Model 1B).

Second, we constructed two theoretically driven models. The first model included three latent factors (Model 2A, Figure 2.). Immediate word recall, delayed word recall, and digit ordering loaded on the first factor; the nine WAIS items loaded on the second factor; and

[2]Similar constraints appear in Figures 2 and 3

letter and category fluency loaded on the third factor. Each latent factor was also partly determined by an independent stochastic disturbance. These three latent factors might be reasonably described as 1) "Memory/Attention," 2) "Abstract Reasoning," and 3) "Verbal Fluency." Each latent factor then loaded on a common second-order factor. Essentially, the three-factor model was specified to account for the correlated errors in Model 1B. An additional correlated error that substantially improved overall fit was implied between immediate and delayed recall and was incorporated in a modification of this model (3B).

Next, we constructed a second theoretically driven model with four latent factors (Model 4, see Figure 3), which is similar to the three latent factor model. Immediate and delayed word recall loaded on the first factor; digit ordering loaded on the second; the nine WAIS items loaded on the third; and letter and category fluency loaded on the fourth. These four factors were termed 1) "Memory," 2) "Attention," 3) "Abstract Reasoning," and 4) "Verbal Fluency." Each of these four factors loaded on a common second-order factor.

We compared the relative fits of these five models using BIC and RMSEA and chose the best fitting model for further investigation, particularly with respect to gender. We also estimated a series of nested models that equated combinations of factor loadings and residual variances for both genders to investigate potential gender differences in factor loadings. First, we combined males and females in a single model but did not equate any of the parameters between gender groups (Model 3B0). Then we constrained all first-order factor loadings to be equal for both genders (Model 3B1). Next, we constrained all first- and second-order factor loadings to be equal for both genders (Model 3B2). Third, we freed the second-order memory/attention loading, keeping all other first- and second-order loadings equal across gender (Model 3B3).

3 Findings

Table 3 presents summary statistics for all estimated models. In our notation, superscripts denote gender (M or F) and subscripts denote a model or a pair of models being compared (e.g., "1A" indicates Model 1A; "3A:3B" indicates Model 3A versus Model 3B).

The single latent factor models (1A) for males and females fit equally well, $RMSEA^M_{1A} = 0.0435 = RMSEA^F_{1A}$. However, the corrections for correlated errors included in the modification of this model (1B) substantially improved fit ($\Delta BIC^M_{1B:1A} = 644.32; \Delta BIC^F_{1B:1A} = 910.34$).

Model 3A fit well for both males and females ($RMSEA^M_{3A} = 0.0274; RMSEA^F_{3A} = 0.0213$), but Model 1B would still be strongly preferred over Model 3A for both genders ($\Delta BIC^M_{3A:1B} = 71.04; \Delta BIC^F_{3A:1B} = 37.75$). However, including a correction for a correlated error between immediate and delayed recall in Model 3B improved relative fit significantly ($\Delta BIC^M_{3B:3A} = 114.59; \Delta BIC^F_{3B:3A} = 61.86$), implying strong preference for Model 3B over all preceding models (e.g., $\Delta BIC^M_{3B:1B} = 43.54; \Delta BIC^F_{3B:1B} = 24.11$).

Similar to the structural analogy between Models 1B and 3A, Model 4 explicitly separated the first latent factor from Model 3B into two factors, immediate and delayed recall loading on the first and digit ordering loading on the second. This model also fit well ($RMSEA^M_4 = 0.0236; RMSEA^F_4 = 0.0230$), but Model 3B was still strongly preferred for both males and females ($\Delta BIC^M_{3B:4} = 15.68; \Delta BIC^F_{3B:4} = 109.27$). Interestingly, in terms of RMSEA and

BIC, Model 4 was the second most preferred model for males, but the fourth most preferred model for females, after Models 3B, 1B, and 3A. Note that while all of the models had RMSEA values less than 0.05, indicating acceptable overall fit, the three latent factor model with a correlated error (3B) fit best for both males and females. Our remaining discussion focuses on this model.

Preferred Model Table 4 presents standardized factor loadings for Model 3B for males and females. Figures 4 and 5 present these in graphical form, including stochastic disturbance parameters. For males, digit ordering loaded highest on the memory/attention factor ($\lambda^M_{3,1} = 0.59$), delayed recall loaded the lowest ($\lambda^M_{2,1} = 0.43$), and immediate recall was in between ($\lambda^M_{1,1} = 0.54$). A similar pattern was observed for females, but with digit ordering loading slightly higher and immediate recall slightly lower on the memory/attention factor than for males ($\lambda^F_{3,1} = 0.68; \lambda^F_{2,1} = 0.42; \lambda^F_{1,1} = 0.48$). The fourth WAIS-R item loaded highest on the abstract reasoning factor for males ($\lambda^M_{7,2} = 0.84$) and the last WAIS-R item the lowest ($\lambda^M_{12,2} = 0.27$), with the remaining WAIS-R items in the 0.50 to 0.70 range ($\lambda^M_{4,2} = 0.49; \lambda^M_{5,2} = 0.65; \lambda^M_{6,2} = 0.55; \lambda^M_{8,2} = 0.40; \lambda^M_{9,2} = 0.54; \lambda^M_{10,2} = 0.67; \lambda^M_{11,2} = 0.60$). For females, the fourth ($\lambda^F_{7,2} = 0.64$), seventh ($\lambda^F_{10,2} = 0.65$), and eighth ($\lambda^F_{11,2} = 0.64$) WAIS-R items loaded the highest on the abstract reasoning factor, though not as highly as for males. However, like males, the last WAIS-R item loaded the lowest ($\lambda^F_{12,2} = 0.29$). The remaining WAIS-R items were in the 0.40 to 0.60 range ($\lambda^F_{4,2} = 0.52; \lambda^F_{5,2} = 0.52; \lambda^F_{6,2} = 0.41; \lambda^F_{8,2} = 0.48; \lambda^F_{9,2} = 0.58$). These loadings were generally lower for females than males. Finally, letter and category fluency loaded roughly equally on the verbal fluency factor, but higher for males ($\lambda^M_{13,3} = 0.60; \lambda^M_{14,3} = 0.59$) than females ($\lambda^F_{13,3} = 0.55; \lambda^F_{14,3} = 0.54$).

The abstract reasoning factor loaded on the second-order cognitive ability factor similarly for males ($\gamma^M_{2,1} = 0.73$) and females ($\gamma^F_{2,1} = 0.72$). However, the loading for the memory/attention factor was substantially higher for males ($\gamma^M_{1,1} = 0.78$) than females ($\gamma^F_{1,1} = 0.47$) and the loading for the verbal fluency factor somewhat lower for males ($\gamma^M_{3,1} = 0.84$) than females ($\gamma^F_{3,1} = 0.91$). Table 5 presents the implied correlations among the latent factors. Note that the correlations involving the memory/attention factor are all substantially higher for males ($r^M_{2,1} = 0.57; r^M_{3,1} = 0.65; r^M_{4,1} = 0.78$) than females ($r^F_{2,1} = 0.34; r^F_{3,1} = 0.43; r^F_{4,1} = 0.47$). The other correlations were roughly comparable ($r^M_{3,2} = 0.61, r^F_{3,2} = 0.66; r^M_{4,2} = 0.73, r^F_{4,2} = 0.72; r^M_{4,3} = 0.84, r^F_{4,3} = 0.91$).

Gender Equality Constraints Table 6 shows summary statistics for the series of nested models investigating factor loading differences between genders. Model 3B0 had no parameter equality constraints. Model 3B1 constrained all first-order factor loadings to be equal across gender. While this resulted in a nominally significant decrease in fit ($\Delta\chi^2_{3B1:3B0}(11) = 51.79, p < 0.0005$), dramatic improvement in BIC suggests preference for this more parsimonious model ($\Delta BIC_{3B1:3B0} = 43.99$). Model 3B2 constrained all first- and second-order factor loadings to be equal across gender. This again resulted in a nominally significant decrease in fit compared to the baseline model ($\Delta\chi^2_{3B2:3B0}(13) = 68.59, p < 0.0005$) and the first-order loading constrained model ($\Delta\chi^2_{3B2:3B1}(2) = 16.81, p < 0.0005$). However, while BIC suggested preference for Model 3B2 over Model 3B0 ($\Delta BIC_{3B2:3B0} = 44.59$), there was no distinction in relative fit between Models 3B2 and 3B1 ($\Delta BIC_{3B2:3B1} = 0.60$). Finally, Model 3B3 allowed just the second-order memory/attention loading to vary between genders. This modification increased relative fit both nominally ($\Delta\chi^2_{3B2:3B3}(11) = 16.41, p < 0.0005$) and according to BIC ($\Delta BIC_{3B2:3B3} = 7.71$) compared to Model 3B2.

4 Discussion

We set out to determine an appropriate structural specification for the six cognitive assessments in the 2003-05 wave of WLS data collection. Though all of our models fit reasonably well, our results suggest that a three latent factor model with one correlated error represents the data better than, and is to be preferred over, both a simpler single latent factor model with and without correlated errors and a more complex four latent factor model. For women and for men, there appears to be a single, second-order common factor for overall cognitive functioning. We believe that other analyses of cognitive functioning in the 2003-05 WLS graduate sample should use this model (3B).

However, though data for males and females both suggest this was the best fitting model with respect to the alternatives, the loading differences between males and females should be noted. According to BIC, a substantial gain in relative fit was achieved by constraining all first-order factor loadings to be equal across gender. Constraining all first- and second-order factor loadings resulted in no fit difference, but then allowing just the second-order memory/attention loading to vary between genders again resulted in substantial gains in relative fit.

This suggests several conclusions regarding gender equality in factor loadings. First, it appears that all first-order factor loadings and the second-order abstract reasoning and verbal fluency loadings can be regarded as equal. However, the memory/attention factor loaded significantly higher on the second-order factor for males than females, resulting in an increase in relative fit when it was freed. Because of this, we suspect that the underlying relationships between these measures (and with the second-order general cognitive ability factor) may be markedly different for the genders. In fact, constraining all first- and second-order loadings to be equal across gender except for the memory/attention loading produced the best fitting model in our investigation.

While these findings suggest that it is statistically most efficient to use this partially constrained model (all loadings except memory/attention), doing so may not be substantively meaningful or technically feasible. Admittedly, for many purposes it may be more desirable to keep males and females separate in what amounts to a full gender interaction design. If this is indeed the case, then separate models taking the form of Model 3B would likely be sufficient.

Finally, several ideas for future investigations logically flow from our work. First, these models should be cross-validated with the WLS sibling sample data. Those data will provide additional insight into the factor structure of these tests in an age heterogeneous sample and, in conjunction with the graduate data, make modeling within- and between-family variation possible. As a minor point, future work should explicitly investigate the statistical and substantive implications of constraining cut points for the polychoric correlations across gender, if technically possible. Lastly, though we prefer this type of presentation because of its relative elegance, more work should be done on the substantive implications of using such a measurement model for cognitive ability in a variety of analytic contexts. That is, we have limited our current investigation to the explicitly cognitive measures in the 2003-05 round of WLS graduate data collection. Future work should investigate if and how the relationships between indicators of cognition and non-cognitive items are mediated by the factor structure we have explored here.

References

Allison, Paul D. 1987. "Estimation of Linear Models with Incomplete Data." pp. 71-103 in Sociological Methodology, 1987, edited by C. C. Clogg. Washington, D.C.: American Sociological Association.

Allison, Paul D. and Robert M. Hauser. 1991. "Reducing Bias in Estimates of Linear Models by Remeasurement of a Random Subsample." Sociological Methods and Research 19:466-92.

Borkowski, J. G., A. L. Benton, and O. Spreen. 1967. "Word Fluency and Brain Damage." Neuropsychologia 5:135-140.

Brandt, J., M. Spencer, and M. Folstein. 1988. "The Telephone Interview for Cognitive Status." Neuropsychiatry, Neuropsychology, Behavioral Neurology 1:111-17.

Drachman, David A. and Janet Leavitt. 1972. "Memory impairment in the aged: Storage versus retrieval deficit." Journal of Experimental Psychology 93:302-308.

Hauser, Robert M. 2005. "Survey Response in the Long Run: The Wisconsin Longitudinal Study." Field Methods 17:3-29.

Hauser, Robert M. and Robert J. Willis. 2005. "Survey design and methodology in the Health and Retirement Study and the Wisconsin Longitudinal Study." pp. 209-235 in Aging, Health, and Public Policy: Demographic and Economic Perspectives, edited by L. J. Waite. New York: Population Council.

Herzog, A. R. and R. B. Wallace. 1997. "Measures of cognitive functioning in the AHEAD Study." J Gerontol B Psychol Sci Soc Sci 52 Spec No:37-48.

Herzog, A. Regula and Willard L. Rodgers. 1999. "Cognitive Performance Measures in Survey Research on Older Adults." pp. 327-340 in Cognition, Aging, and Self-Reports, edited by N. Schwarz, D. Park, B. Knuper, and S. Sudman. Philadelphia, PA: Psychology Press.

Jöreskog, K. G. and Dag Sörbom. 1996. LISREL 8 user's reference guide. Chicago, IL :: Scientific Software International.

Kozora, E. and C. M. Cullum. 1995. "Generative Naming in Normal Aging: Total Output and Qualitative Changes Using Phonemic and Semantic Constraints." Clinical Neuropsychologist 9:313-320.

Loehlin, John C. 2004. Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis. Hillsdale, NJ: Lawrence Erlbaum Associates.

Monsch, A. U., M. W. Bondi, N. Butters, J. S. Paulsen, D. P. Salmon, P. Brugger, and M. R. Swenson. 1994. "A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease." Neuropsychology 8:25-30.

Nesselroade, John R., Nancy L. Pedersen, Gerald E. McClearn, Robert Plomin, and C. S. Bergeman. 1988. "Factorial and Criterion Validities of Telephone-Assessed Cognitive Ability Measures: Age and Gender Comparisons in Adult Twins." Research on Aging 10:220-234.

Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." pp. 111-163 in Sociological Methodology 1995, edited by P. V. Marsden. Cambridge: Basil Blackwell.

Schooler, C., M. S. Mulatu, and G. Oates. 1999. "The continuing effects of substantively complex work on the intellectual functioning of older workers." Psychol Aging 14:483-506.

Sewell, William H., Robert M. Hauser, Kristen W. Springer, and Taissa S. Hauser. 2004. “As We Age: The Wisconsin Longitudinal Study, 1957-2001.” pp. 3-111 in Research in Social Stratification and Mobility, vol. 20, edited by K. Leicht. London: Elsevier.

Wechsler, David. 1981. WAIS-R Manual: Wechsler Adult Intelligence Scale-Revised. San Antonio: The Psychological Corporation.

———. 1997. Weschler Adult Intelligence Scale-Third Edition. New York: Psychological Corporation.

Table 1: Descriptive Statistics for Cognitive Ability Measures by Gender: Wisconsin Longitudinal Study, 2003-05

	N	Mean	Median	Std. Dev
Males				
Immediate Recall	2056	5.76	6.00	1.69
Delayed Recall	2056	3.57	3.00	2.00
Digit Ordering	2056	5.79	6.00	1.62
WAIS 1: Orange/Banana	2746	1.86	2.00	0.38
WAIS 2: Boat/Automobile	2056	1.71	2.00	0.66
WAIS 3: Eye/Ear	2746	1.37	1.00	0.69
WAIS 4: North/West	2056	1.85	2.00	0.52
WAIS 5: Egg/Seed	2746	1.04	1.00	0.72
WAIS 6: Table/Chair	2746	1.05	1.00	0.86
WAIS 7: Poem/Statue	2056	0.96	1.00	0.91
WAIS 8: Fly/Tree	2746	0.68	0.00	0.87
WAIS 9: Praise/Punishment	2746	0.66	0.00	0.86
Letter Fluency	2056	2.36	2.00	1.12
Category Fluency	1316	2.95	3.00	0.98
Females				
Immediate Recall	2489	6.55	7.00	1.68
Delayed Recall	2489	4.59	5.00	2.13
Digit Ordering	2489	6.02	6.00	1.60
WAIS 1: Orange/Banana	3295	1.92	2.00	0.30
WAIS 2: Boat/Automobile	2489	1.70	2.00	0.65
WAIS 3: Eye/Ear	3295	1.35	1.00	0.69
WAIS 4: North/West	2489	1.89	2.00	0.46
WAIS 5: Egg/Seed	3295	0.91	1.00	0.70
WAIS 6: Table/Chair	3295	1.18	1.00	0.87
WAIS 7: Poem/Statue	2489	0.94	1.00	0.92
WAIS 8: Fly/Tree	3295	0.65	0.00	0.83
WAIS 9: Praise/Punishment	3295	0.63	0.00	0.85
Letter Fluency	2489	2.62	3.00	1.11
Category Fluency	1560	3.22	3.00	1.08

Table 2: Observed Pair-Wise Polychoric Correlations for Cognitive Ability Measures by Gender: Wisconsin Longitudinal Study, 2003-05

Males	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00													
2	0.70	1.00												
3	0.31	0.24	1.00											
4	0.10	0.07	0.11	1.00										
5	0.20	0.15	0.20	0.38	1.00									
6	0.14	0.13	0.21	0.22	0.33	1.00								
7	0.25	0.20	0.32	0.40	0.55	0.44	1.00							
8	0.07	0.07	0.10	0.18	0.21	0.26	0.26	1.00						
9	0.14	0.12	0.13	0.30	0.33	0.19	0.42	0.10	1.00					
10	0.19	0.15	0.18	0.24	0.33	0.30	0.40	0.26	0.38	1.00				
11	0.13	0.10	0.15	0.24	0.32	0.28	0.44	0.28	0.37	0.43	1.00			
12	0.07	0.07	0.01	0.10	0.15	0.14	0.26	0.15	0.15	0.19	0.18	1.00		
13	0.22	0.18	0.21	0.16	0.22	0.19	0.26	0.13	0.15	0.24	0.18	0.09	1.00	
14	0.19	0.16	0.21	0.12	0.22	0.19	0.30	0.14	0.14	0.24	0.23	0.04	0.34	1.00

Females	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.00													
2	0.74	1.00												
3	0.31	0.28	1.00											
4	0.08	0.05	0.18	1.00										
5	0.09	0.12	0.15	0.29	1.00									
6	0.07	0.05	0.10	0.17	0.25	1.00								
7	0.07	0.05	0.11	0.35	0.28	0.27	1.00							
8	0.07	0.07	0.10	0.21	0.24	0.24	0.29	1.00						
9	0.06	0.06	0.15	0.37	0.31	0.16	0.38	0.17	1.00					
10	0.07	0.08	0.14	0.25	0.33	0.27	0.39	0.31	0.35	1.00				
11	0.07	0.04	0.11	0.34	0.28	0.23	0.37	0.34	0.36	0.42	1.00			
12	0.01	0.01	0.09	0.15	0.12	0.11	0.26	0.18	0.17	0.20	0.17	1.00		
13	0.10	0.07	0.16	0.13	0.17	0.17	0.27	0.17	0.20	0.25	0.22	0.11	1.00	
14	0.15	0.14	0.10	0.16	0.18	0.15	0.19	0.16	0.16	0.22	0.22	0.10	0.29	1.00

1=Immediate Recall, 2=Delayed Recall, 3=Digit Ordering, 4=Orange/Banana, 5=Boat/Automobile, 6=Eye/Ear, 7=North/West, 8=Egg/Seed, 9=Table/Chair, 10=Poem/Statue, 11=Fly/Tree, 12=Praise/Punishment, 13=Letter Fluency, 14=Category Fluency

Table 3: Model Summary Statistics by Gender: Wisconsin Longitudinal Study, 2003-05

	χ^2	DF	RMSEA	BIC
Males (N=2746)				
One Latent Factor (1A)	977.83	189	0.0435	-518.65
One Latent Factor, w/Correlated Errors (1B)	301.85	185	0.0244	-1162.97
Three Latent Factors (3A)	380.80	186	0.0274	-1091.93
Three Latent Factors, w/Correlated Error (3B)	258.30	185	0.0226	-1206.51
Four Latent Factors (4)	281.90	186	0.0236	-1190.83
Females (N=3295)				
One Latent Factor (1A)	1172.75	189	0.0435	-358.18
One Latent Factor, w/Correlated Errors (1B)	230.01	185	0.0195	-1268.52
Three Latent Factors (3A)	275.86	186	0.0213	-1230.77
Three Latent Factors, w/Correlated Error (3B)	205.90	185	0.0184	-1292.63
Four Latent Factors (4)	323.28	186	0.0230	-1183.36

Table 4: Standardized Factor Loadings by Gender, Wisconsin Longitudinal Study, 2003-05

	Factor 1	Factor 2	Factor 3	2nd Order
Males				
Immediate Recall	0.54			
Delayed Recall	0.43			
Digit Ordering	0.59			
WAIS-1: Orange/Banana		0.49		
WAIS-2: Boat/Automobile		0.65		
WAIS-3: Eye/Ear		0.55		
WAIS-4: North/West		0.84		
WAIS-5: Egg/Seed		0.40		
WAIS-6: Table/Chair		0.54		
WAIS-7: Poem/Statue		0.67		
WAIS-8: Fly/Tree		0.60		
WAIS-9: Praise/Punishment		0.27		
Letter Fluency			0.60	
Category Fluency			0.59	
Factor 1: Memory/Attention				0.78
Factor 2: Abstract Reasoning				0.73
Factor 3: Verbal Fluency				0.84
Females				
Immediate Recall	0.48			
Delayed Recall	0.42			
Digit Ordering	0.68			
WAIS-1: Orange/Banana		0.52		
WAIS-2: Boat/Automobile		0.52		
WAIS-3: Eye/Ear		0.41		
WAIS-4: North/West		0.64		
WAIS-5: Egg/Seed		0.48		
WAIS-6: Table/Chair		0.58		
WAIS-7: Poem/Statue		0.65		
WAIS-8: Fly/Tree		0.64		
WAIS-9: Praise/Punishment		0.29		
Letter Fluency			0.55	
Category Fluency			0.54	
Factor 1: Memory/Attention				0.47
Factor 2: Abstract Reasoning				0.72
Factor 3: Verbal Fluency				0.91

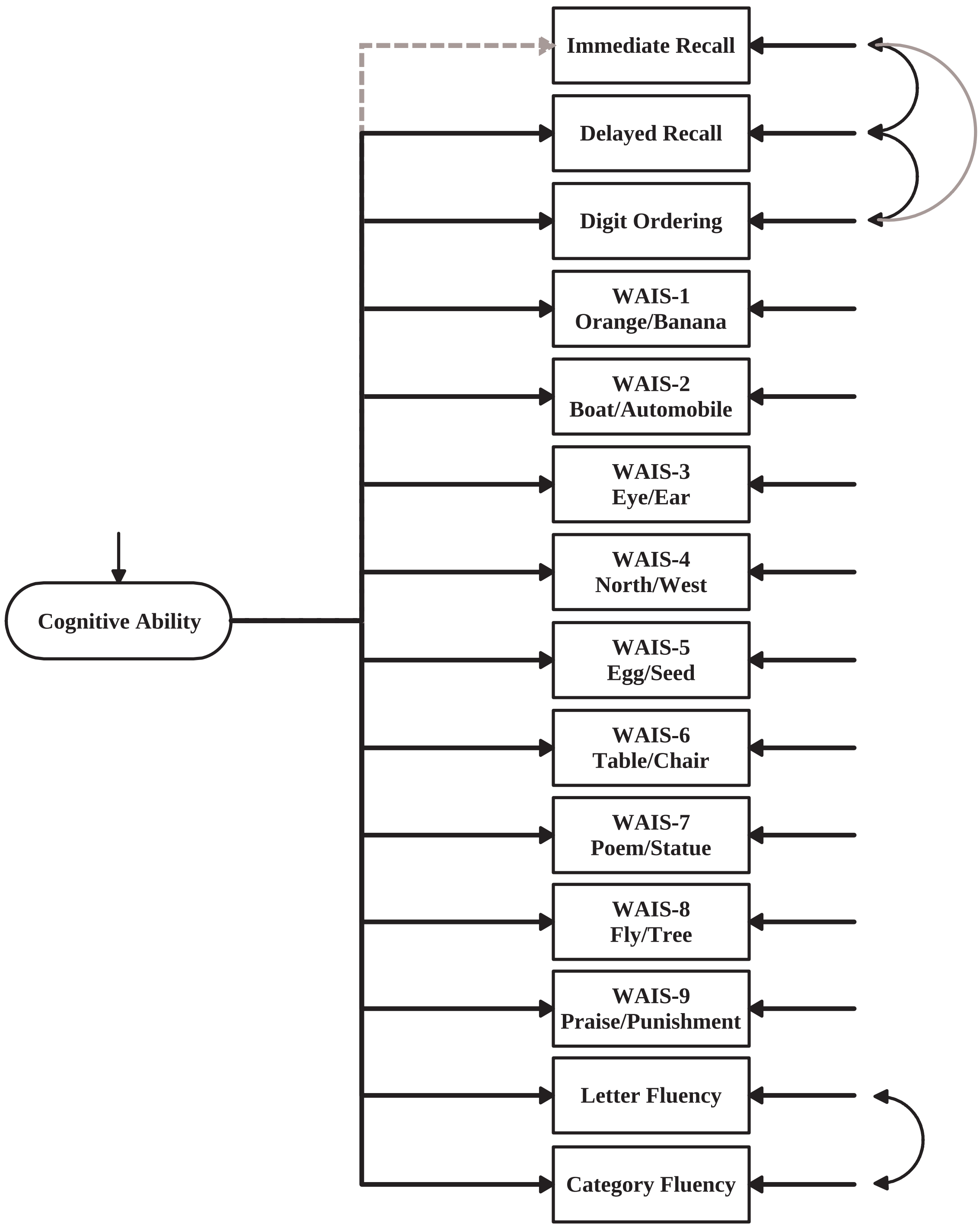
Table 5: Correlations Among Factors by Gender, Wisconsin Longitudinal Study, 2003-05

	Factor 1	Factor 2	Factor 3	2nd Order
Males				
Factor 1: Memory/Attention	1.00			
Factor 2: Abstract Reasoning	0.57	1.00		
Factor 3: Verbal Fluency	0.65	0.61	1.00	
Second-Order Factor: Cognitive Ability	0.78	0.73	0.84	1.00
Females				
Factor 1: Memory/Attention	1.00			
Factor 2: Abstract Reasoning	0.34	1.00		
Factor 3: Verbal Fluency	0.43	0.66	1.00	
Second-Order Factor: Cognitive Ability	0.47	0.72	0.91	1.00

Table 6: Gender Equality Constrained Model Summary Statistics: Wisconsin Longitudinal Study, 2003-05

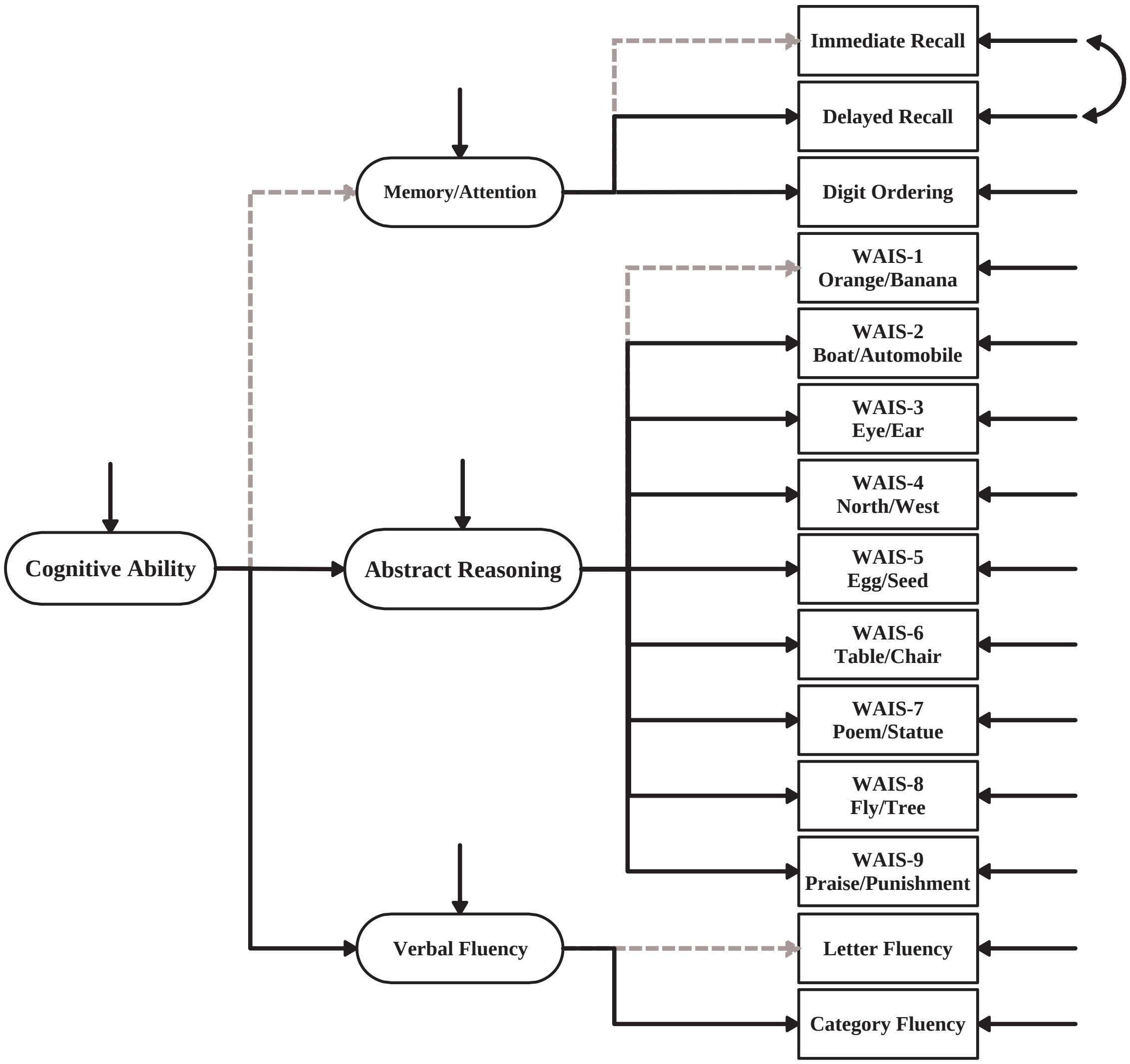
	χ^2	DF	RMSEA	BIC
Model 3B0 No Constraints	464.20	370	0.0144	-2757.14
Model 3B1 First-Order	515.98	381	0.0150	-2801.13
Model 3B2 First- and Second-Order	532.79	383	0.0152	-2801.73
Model 3B3 First- and Second-Order Minus Memory/Attention	516.38	382	0.0150	-2809.44

Figure 1: Single Latent Factor (Models 1A and 1B)



Note: Model 1B includes the correlations between errors in observables shown above; Model 1A does not.

Figure 2: Three Latent Factors, One Second-Order Factor (Models 3A and 3B)



Note: Model 3B includes the correlations between errors in observables shown above; Model 3A does not.

Figure 3: Four Latent Factors, One Second-Order Factor (Model 4)

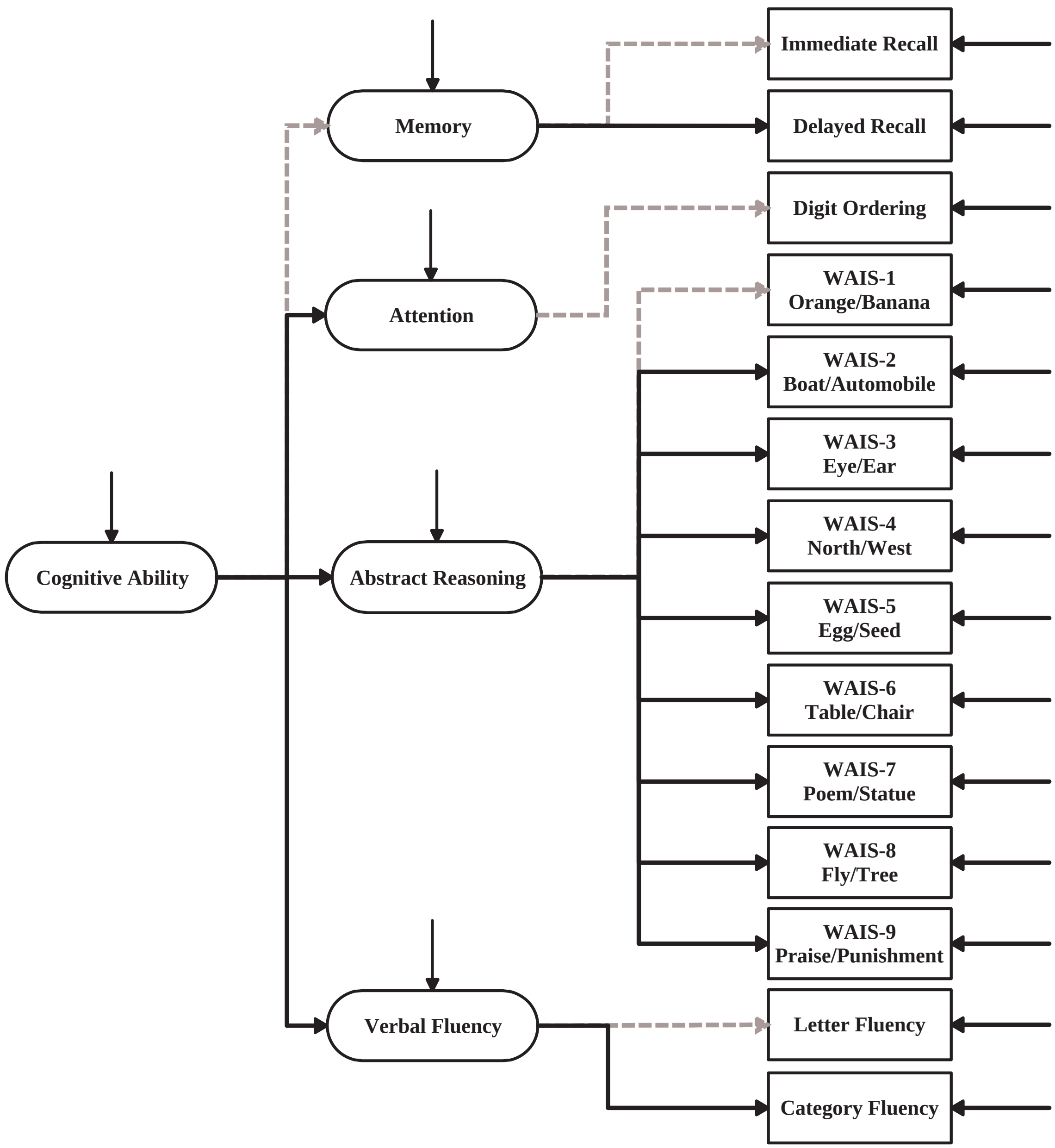


Figure 4: Estimated Standardized Solution, Males: Wisconsin Longitudinal Study, 2003-05

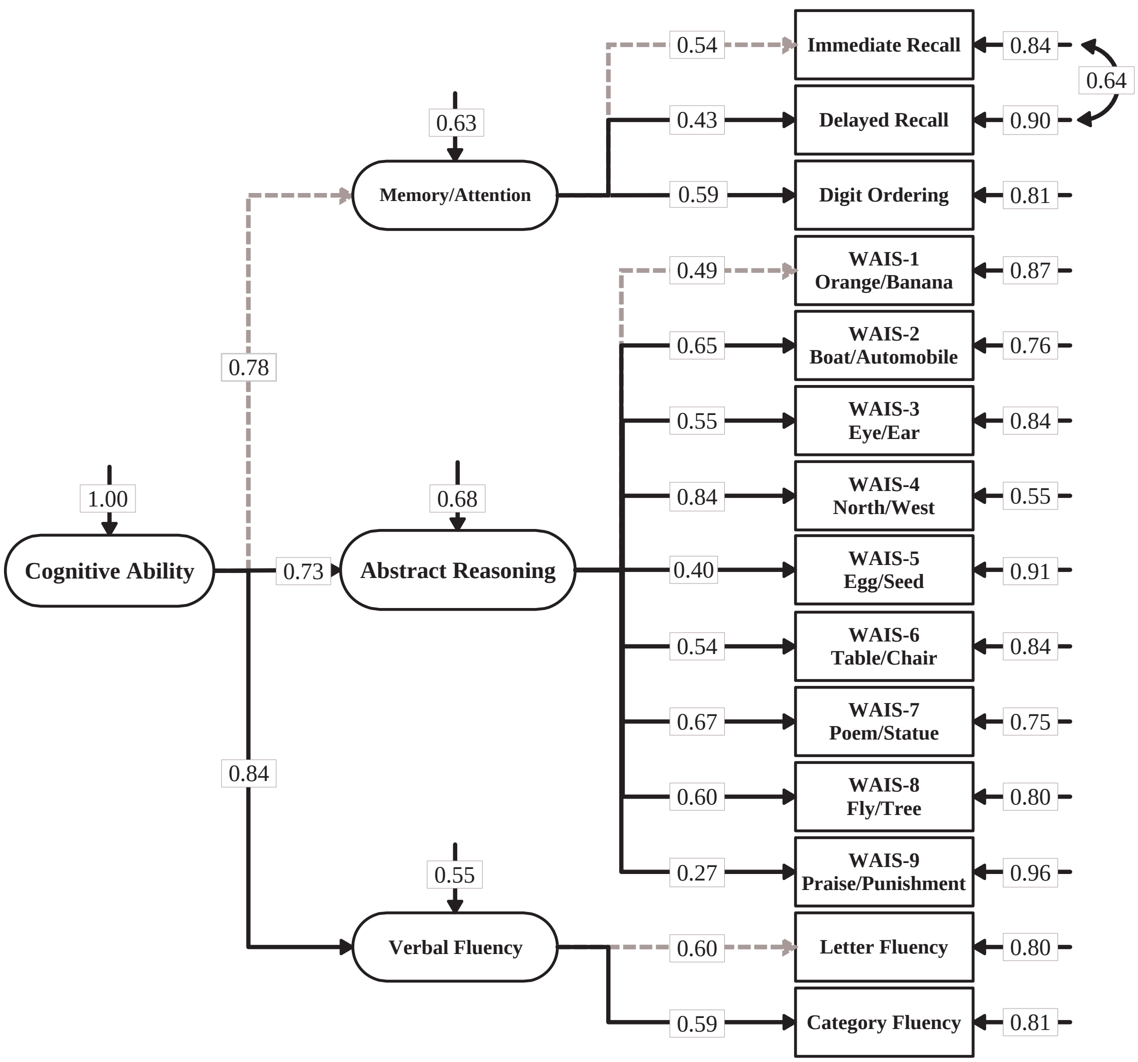
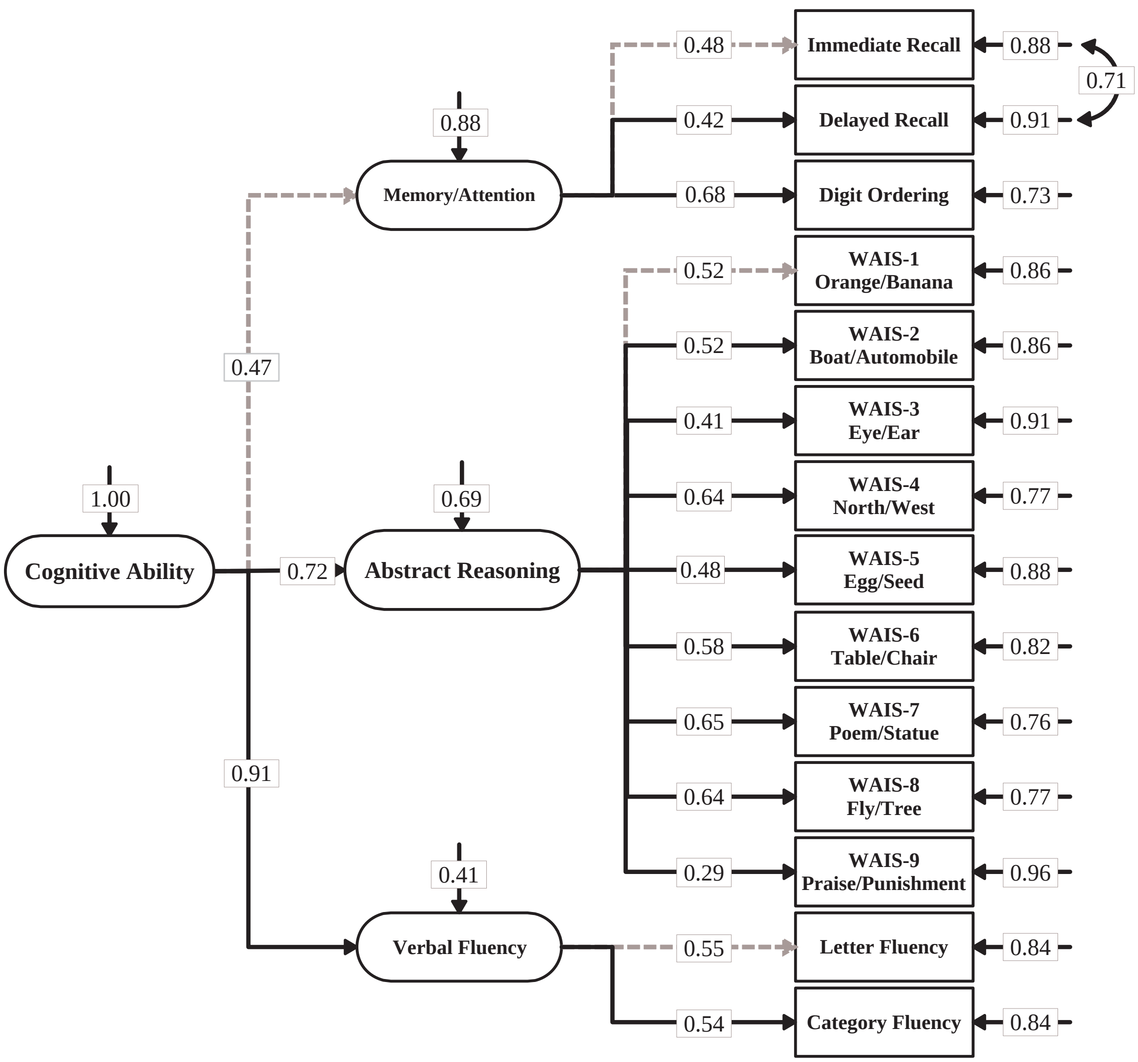


Figure 5: Estimated Standardized Solution, Females: Wisconsin Longitudinal Study, 2003-05



Center for Demography and Ecology
University of Wisconsin
1180 Observatory Drive Rm. 4412
Madison, WI 53706-1393
U.S.A.
608/262-2182
FAX 608/262-8400
comments to: jyonker@ssc.wisc.edu
requests to: cdepubs@ssc.wisc.edu