Published in final edited form as:
*Exp Aging Res*. 2005 ; 31(3): 291–312. doi:10.1080/03610730590948195.

# A GROWTH CURVE MODEL OF LEARNING ACQUISITION AMONG COGNITIVELY NORMAL OLDER ADULTS

**Richard N. Jones**,
Hebrew Rehabilitation Center for Aged, Research and Training Institute, Boston, Massachusetts, USA

**Adrienne L. Rosenberg**,
Hebrew Rehabilitation Center for Aged, Research and Training Institute, Boston, Massachusetts, USA

**John N. Morris**,
Hebrew Rehabilitation Center for Aged, Research and Training Institute, Boston, Massachusetts, USA

**Jason C. Allaire**,
Department of Psychology, North Carolina State University, Raleigh, North Carolina, USA

**Karin J. M. McCoy**,
Institute on Aging, and Department of Clinical and Health Psychology, University of Florida, Gainesville, Florida, USA

**Michael Marsiske**,
Institute on Aging, and Department of Clinical and Health Psychology, University of Florida, Gainesville, Florida, USA

**Ken P. Kleinman**,
Harvard Medical School and Harvard Pilgrim Health Care, Boston, Massachusetts, USA

**George W. Rebok**, and
Department of Mental Health, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA

**Paul F. Malloy**
Brown University Medical School, Providence, Rhode Island, USA

## Abstract

Address correspondence to Richard N. Jones, ScD, Senior Research Associate, Hebrew Rehabilitation Center for Aged, Research and Training Institute, 1200 Centre Street, Boston, MA 02131, USA. jones@mail.hrca.harvard.edu.

A preliminary version of this study was presented at the 54th Annual Scientific Meeting of the Gerontological Society of America, November 16, 2001, Chicago, Illinois. The ACTIVE study (Advanced Cognitive Training for Independent and Vital Elderly) is a multisite collaborative cognitive intervention trial. The National Institute on Aging Scientific Coordinator at the time of award was Jared Jobe, and current is Daniel Berch. The National Institute on Nursing Research Scientific Coordinator at the time of award was Mary Levek, and is now Karin Helmers. Sharon Tennstedt is the principal investigator at the coordinating center, New England Research Institutes, Watertown, Massachusetts (AG14282). The principal investigators and field sites include Karlene Ball, University of Alabama at Birmingham (AG14289); Michael Marsiske, Institute on Aging, University of Florida, Gainesville (AG14276); John Morris, Hebrew Rehabilitation Center for Aged Research and Training Institute, Boston (NR04507); George Rebok, Johns Hopkins University Bloomberg School of Public Health (AG14260); and Sherry Willis, Penn State University, Gerontology Center (AG14263). David Smith was the principal investigator at Indiana University School of Medicine, Regenstrief Institute, Indianapolis (NR04508), at the time of initial award, currently Fred Unverzagt is the principal investigator.


NIH-PA Author Manuscript
NIH-PA Author Manuscript

The objective of this study was to model recall and learning on the Auditory Verbal Learning Test using latent growth curve techniques. Participants were older adults recruited for the ACTIVE cognitive intervention pilot. A series of nested models revealed that an approximately logarithmic growth curve model provided optimal fit to the data. Although recall and learning factors were statistically uncorrelated, a fitted multivariate model suggested that initial recall was significantly associated with demographic characteristics but unrelated to health factors and cognitive abilities. Individual differences in learning were related to race/ethnicity, speed of processing, verbal knowledge, and global cognitive function level. These results suggest that failing to recognize initial recall and learning as distinct constructs clouds the interpretation of supraspan memory tasks.

---

The Rey Auditory Verbal Learning Test (AVLT) is one of the oldest and most widely used tests of memory functioning (Boake, 2000; Rey, 1964). The AVLT measures many aspects of memory for words, including immediate memory span (initial recall), learning, and retention following an intrusion (Lezak, 1995). The continued relevance of the AVLT is reflected in the availability of norms for various populations (e.g., Ivnik et al., 1990, 1992; Kurylo, Temple, Elliott, & Crawford, 2001; Smith et al., 1992; van-den-Burg & Kingma, 1999) and its utility in making classifications according to the emerging concept of Mild Cognitive Impairment (Petersen et al., 1999). Although there are a number of AVLT summary scores based on various permutations or arithmetic manipulations of AVLT data (e.g., Vakil & Blachstein, 1997), this investigation represents an attempt to build a single statistical model that simultaneously captures distinct aspects of AVLT performance: initial recall and learning.

Of particular importance in this context is the ability of the scoring algorithm used for AVLT performance (and other multitrial learning and memory measures) to permit the separation of initial recall and subsequent learning. Performance on the initial trial of a supraspan list-learning task, like the AVLT, is often thought to be a reflection of verbal attention or immediate word span recall (Delis, Kramer, Kaplan, & Ober, 2000), rather than a measure of verbal learning (e.g., encoding, storage, and retention). Improvement, or performance in general, across multiple recall trials is typically characterized as the learning curve or learning slope (Brandt & Benedict, 2001; Spreen & Strauss, 1998), as encoding and retrieval from long-term memory play an increasingly greater role in performance with each presentation (Lezak, 1995). This learning slope has been calculated a number of ways. For example, on the Hopkins Verbal Learning Test—Revised (HVLT-R; Benedict, Schretlen, Groninger, & Brandt, 1998), learning is defined as the difference between the final immediate recall trial and the initial recall trial, whereas on the California Verbal Learning Test—II (CVLT-II; Delis et al., 2000) learning is based on a least squares regression line that fits the total items correct across the five immediate recall trials. Consequently, the slope of this regression line reflects the amount of new learning per trial, and assumes linearity of the learning trend. These approaches vary in the degree to which they take advantage of the additional information provided in the multiple learning trials. Thus, the level of performance on the initial trial of a list-learning task may be only moderately related to the slope of performance improvement over time, as the two may reflect different cognitive processes.

Previous research suggests that the separation of initial recall and learning may yield important differential outcome predictions. For example, Royall and colleagues have recently demonstrated that among residents of a continuing-care retirement community, learning rate over trials on the CVLT predicts dementia type (Alzheimer's-type versus non-Alzheimer's type), but baseline cross-sectional performance did not (Royall, Palmer, Chiodo, & Polk, 2003). The authors suggest that performance on the CVLT learning trials

captures the quality of hippocampal functioning. Another example is represented by the work of Nettelbeck, Rabbitt, Wilson, & Batt (1996) who found using a serial learning task that initial recall and learning were statistically uncorrelated, and that initial recall was more strongly related to indices of processing speed.

Previous research has also highlighted other predictors of initial recall and learning (e.g., sociodemographic characteristics such as age, gender, race, and education, and health factors). With respect to age, a number of cross-sectional and longitudinal studies have found evidence of negative age differences and age-related decline performance on measures of episodic memory (see Bäckman, Small, Wahlin, & Larson, 2000). Published age-based norms for the AVLT demonstrate negative age differences in performance across all five trials as well as in the magnitude of learning over trials (Ivnik et al., 1990). Additionally, Lamar, Resnick, and Zonderman (2003) found significant negative age differences specifically in learning over the trials of the CVLT, a measure of word list recall similar to the AVLT. In addition, gender has been associated with the magnitude of learning gains over five trials of the CVLT, with females exhibiting greater gains than males (Lamar et al., 2003). Norman, Evans, Miller, and Heaton (2000) found differences in the CVLT between African American and white older adults in initial recall and overall performance (see also Uchiyama et al., 1995). A number of studies have reported a significant positive relationship between years of education and performance on cognitive measures including word recall tasks (e.g., Anstey & Christensen, 2000) with those individuals with more years of education performing better on measures of cognition than those with less education. Two studies have reported found a significant positive association between years of education and recall performance on the AVLT (Query & Megran, 1983; Uchiyama et al., 1995; see also Norman et al., 2000). Prior studies have not examined the extent to which health is associated with initial recall or growth; however, various indices of health and disease are associated with memory performance (Hultsch, Hammer, & Small, 1993; Verhaeghen, Marcoen, & Goossens, 1992).

Thus, building on prior research, this study seeks to further explore the relationship between initial recall and learning, and their predictors, in a diverse sample of community-dwelling elders. A key innovation of this study is to incorporate contemporary procedures of latent growth curve modeling to enhance the dissociation of initial recall and learning slope. Accurate and reliable measurement of change is essential for longitudinal and clinical outcome research, as well as providing a critical component of clinical assessment (Heaton et al., 2001). The potential value of this analysis is in developing a new conceptualization and method for modeling verbal learning through the use of latent growth curve modeling, which allows for the parameterization of individual differences in recall and rates of learning, here conceptualized as distinct latent variables. In addition, by embedding the growth model in a larger structural equations modeling framework, individual differences in recall and learning can be related to background variables simultaneously, permitting separate description of influences on recall and learning ability.

The goals of this study were twofold. The first goal was to develop a statistical model of AVLT learning using the psychometric approach to latent growth curve modeling. The second goal was to determine the salience of background variables in accounting for individual differences in initial recall and learning. Background variables included sociodemographic factors and cognitive measures thought to be related to memory function.

# METHODS

## Participants and Procedures

The 169 community-dwelling elders who participated in the current study were recruited for the pilot ACTIVE (Advanced Cognitive Training for Vital and Independent Elderly) cognitive intervention trial. The rationale, design, recruitment, and site characteristics of the ACTIVE cognitive intervention trial are described elsewhere (Jobe et al., 2001).

ACTIVE is a multisite study with collaborators in Baltimore, Birmingham, Boston, Detroit, Indianapolis, and State College, Pennsylvania. Potential participants were recruited from congregate senior housing, research subject registries, community centers, and community organizations. Informed consent was obtained from participants with protocols approved by local institutional review boards. Eligibility criteria included age of at least 65 years, and the absence of an active diagnosis of a life-threatening acute illness. Potential participants with a diagnosis of probable Alzheimer's disease or history of stroke within 12 months of screening were excluded. Further, any subject determined to be cognitively impaired after initial screening, defined as a Mini-Mental State Examination score of 22 or less, or for whom the interviewer conducting the screening assessment felt did not at least usually understand or make themselves understood, was excluded. ACTIVE participants were initially contacted by telephone for a screening interview. Individuals who did not trigger exclusionary criteria were invited to in-person testing sessions. These sessions were led by a trained lay interviewer who obtained physiologic, psychological, cognitive and physical functioning measures.

Participant characteristics are summarized in Table 1, which also contains the correlation matrix corresponding to the data analyzed, although the structural equation models estimated used raw data and maximum likelihood estimation via the EM algorithm and Mplus software (Muthén and Muthén, 1998–2003; Los Angeles). Participants ranged in age from 65 to 90 years, with a mean age of 74.3 ($SD$ = 6.1). Slightly more than two-thirds (68%) had completed 12 or more years of education, a level comparable to the U.S. population as suggested by the 1998 Current Population Survey (67%; Wallman, 2000). The level of physical health of participants, as measured by the SF-36 Physical Component Score (PCS), was not different from norms generated from the 1990 National Survey of Functional Health Status for the general U.S. population aged 65 and older (Ware, Kosinski, & Keller, 1994; 42.7 versus 41.3, respectively; $t = 1.4$, $p = .08$). However, ACTIVE pilot participants did score higher on the SF-36 Mental Component Score (MCS) relative to national norms (Ware et al., 1994; 54.3 versus 51.8; $t = 3.8$, $p < .001$), implying better self-reported mental health for ACTIVE participants relative to national norms (the SF-36 PCS and MCS are described below).

## Measures

**Auditory Verbal Learning Test—**One of the cognitive measures included in the ACTIVE battery was the Rey AVLT (Lezak, 1995; Rey, 1964). ACTIVE participants were read a list of 15 words at a 2-interval, and given 3 min to write down as many of those words they could remember. The list was then reread in the same order, and participants allowed to write down however many words they could remember (without referencing their first list). This process was repeated five times. This analysis considers the responses to the five learning trials. Data from the intrusion, delayed recall, and recognition trials were not included.

**Inductive Reasoning—**The word series task from the Schaie-Thurstone Adult Mental Abilities Test (Schaie, 1985) was used as a measure of inductive reasoning. In this task,

participants were asked to discover the pattern to a series of words, and mark the word that would have come next in the series among presented alternatives. The value used in our models was the number of correctly noted patterns. Previous research has demonstrated that fluid abilities, such as inductive reasoning, are significantly related to memory performance (Cockburn & Smith, 1991).

**Verbal Knowledge—**The measure of crystallized verbal knowledge included in the ACTIVE battery was a multiple choice vocabulary test from the Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Derman, 1976). Participants were presented with a stimulus word, and asked to select a synonym from among five choices. The score used in our models was the number of correct choices made, out of 18 possible, within a 4-min time limit.

**Psychomotor Speed—**The Digit Symbol Substitution Test (DSST), from the Weschler Adult Intelligence Scale (Wechsler, 1981), was used as a measure of processing speed. The DSST measures how quickly participants place symbols paired with numbers. In our models, we used the number of correct symbols recorded within 90 s as a background variable.

**Mental Status—**Performance on the Mini-Mental State Examination (MMSE) served as an indicator of general mental status (Folstein, Folstein, & McHugh, 1975). All ACTIVE participants scored greater than 22 on the MMSE as a condition of their inclusion in the study. In our models, we used the MMSE score (range 22–30) as background variable.

**Demographics—**Demographic characteristics were collected during the initial telephone interview. Participants were asked their date of birth, educational attainment, income, gender, and ethnic identity. In our final analytic models, age was treated as a continuous variable across the range 65 to 74, and divided by 10 so that regression parameters could be interpreted as differences per 10-year age difference (untransformed data are presented in Table 1).

**Health Status—**Self-reported physical and mental health was represented with scaled component scores derived from responses to the SF-36 (Ware et al., 1994). The SF-36 is a generic health survey with 36 questions that can be used to generate composite general physical and mental health composite summary measures. Participants completed the SF-36 in their homes and returned it to field staff during in-person interview sessions. Scoring the MCS and PCS is a complex procedure based on factor analysis results discussed by Ware and colleagues (1994), and scaled scores are *T* scores relative to the general U.S. population (expected mean 50, *SD* 10). Although untransformed data are presented in Table 1, in our final models PCS and MCS scores were divided by 10 so that regression parameters could be interpreted as differences per 10-point (i.e., 1 standard deviation) differences in expected PCS and MCS scores.

## Analytic Approach

Individual differences in the AVLT learning trajectories were analyzed using latent growth modeling (LGM; Duncan, Duncan, Strycker, Li, & Alpert, 1999; B. Muthén, 1991). Analyses were conducted with the Mplus software package using the limited information maximum likelihood estimator, and the maximum likelihood complete sample approach to missing data (Muthén & Muthén, 1998–2003). The psychometric approach to LGM models, as implemented in the Mplus software, is very flexible in terms of the shape of the developmental trajectory that can be modeled. LGM models are similar to random coefficient or mixed effect models for repeated-measure data, with random effects

reconceptualized as latent variables. LGMs have one latent variable capturing initial status and one or more additional variables capturing change over time.

By conceptualizing random coefficients as latent variables, the treatment of time in LGM is as a model parameter, rather than as fixed data. Determining the most appropriate values for the time steps (represented with $\lambda$) as a function of time (or the repeated learning trials, $t = 1, 2, 3, 4, 5$) is a process of curve fitting (as it is in other longitudinal data analysis methods). In the current study, after examination of the raw data, different functional forms for learning curves were considered, including linear, quadratic, quadratic with orthogonal polynomial time steps, and logarithmic growth trends. We also estimated a model with only the first trial and last trial estimated, and a final model described below as ‘approximately’ logarithmic. Where possible, we attempted to scale the time vector so that the final time step was equal to 1 to facilitate the interpretation of the estimated mean for the learning factor (i.e., gain over five trials). The parameterization of the different models is illustrated and explicated in Figure 1.

Solutions obtained for models of different functional forms were compared by visually inspecting plots of fitted and observed means and by comparing the Akaike Information Criterion (AIC; Akaike, 1974), favoring the model with the smallest AIC. The AIC rewards parsimony and, unlike other measures of fit discussed below, is appropriate for comparing non-nested models (Jöreskog, 1992). Model parsimony was summarized with the parsimony ratio (Raykov, Tomer, & Nesselroade, 1991).

After a suitable growth model for AVLT learning was developed, we evaluated the influence of specific background variables (i.e., sociodemographic, health, cognitive variables) on initial AVLT recall and learning. We examined the relationship of recall and learning with background variables by regressing the latent growth curve factors on background variables. We looked at the association of each background variable and the latent growth curve factors independently and then in a multivariable fashion by building a model using a forward stepwise model selection procedure informed by the model modification index. First, all demographic variables were entered but paths describing associations with latent growth curve factors were fixed to zero. The modification index for the model was then examined for parameters relevant to the regression of the latent factors on sociodemographic variables, and the single path that would result in the greatest improvement in overall model fit was set to be freely estimated, and the model parameters re-estimated. These steps were repeated until no further regressions of latent growth factors on demographic variables would result in significant improvement in model fit. Then, we repeated the model building by adding health and cognitive factors that had a significant independent effect on latent growth curve factors to the model (while retaining fitted effects from sociodemographic variables). All covariates were centered at their respective means in the multivariate analyses, allowing intercepts and mean structures to be interpreted as estimated values for a hypothetical cohort with a distribution of background variables equivalent to the mean of the overall sample. Model fit was assessed with the model chi-square ($\chi^2$), the root mean squared error of approximation (RMSEA; Browne & Cudek, 1993; Muthén & Muthén, 1998–2003), the comparative fit index (CFI; Bentler, 1990; Muthén & Muthén, 1998–2003), and the AIC.

# RESULTS

## Curve Fitting

Analysis began with determining the growth function (i.e., linear, quadratic, logarithmic) that best characterized learning over the course of the five AVLT trials. The results of model fitting, including model chi-square, AIC, parsimony ratio, RMSEA, CFI, and estimated

mean and variance parameters for the growth model are summarized in Table 2. Initially, an intercept-only model was estimated, which fit very poorly and demonstrates the need for model modification. Thirteen degrees of freedom (13 *df*) result from the estimation of seven parameters (five residual variances for the recall trials, and a mean and a variance term for the latent factor) on a mean and covariance matrix with five observed means and 15 unique variance/covariance estimates. Given the parameterization summarized in Figure 1, the estimate for the mean of the initial recall growth model factor corresponds to the average number recalled across all five trials.

The linear growth curve model was estimated by adding a single latent variable with freely estimated mean and variance parameters and covariance with the initial recall factor, and with fixed loadings in the repeated trials as illustrated in Figure 1. The resulting model had three fewer degrees of freedom and although represented a substantial improvement in model fit, still fit poorly. The mean for the initial recall parameter (6.33) corresponds to the intercept of the growth curve given a linear model for AVLT word learning. Because the loading matrix ($\Lambda$, Figure 1) is calibrated so that the final time step ($\lambda_{52}$) has a value of 1, the mean (and standard error, *SE*) of the learning growth curve factor [5.201 (0.24)] can be interpreted as the average additional number of words recalled at the final trial. The variance of the latent growth term was significantly non-zero [estimate and standard error, *SE*, 3.02 (0.77)], revealing significant individual differences in the rate of learning. Significant individual differences were also detected for the initial recall factor [2.22 (0.39)], and the linear model implied covariance of the initial recall and learning factors was statistically significant, and corresponded to a correlation of 0.59.

We investigated quadratic models, using simple quadratic functions of time steps as well as orthogonal polynomials (Table 2 and Figure 1). Both fit comparably. However, we found estimates from both sets of quadratic models difficult to interpret (e.g., learning factor means are on an arbitrary scale). Moreover, substantive reasons argue against support for a quadratic model in that it implies a curvilinear learning process showing gain and an implied decay, a pattern not typically observed in the raw data.

Although a logarithmic model (Model 5, Table 2) implied slightly worse fit than the orthogonal polynomial quadratic based on the AIC, RMSEA, and CFI, we found that substantive interpretation of a learning factor describing diminishing gains reaching an asymptote a better conceptual fit to observed data. Furthermore, examination of residuals from the logarithmic model identified the estimation of the number of words recalled at the final trial as an important source of model misfit. Significant improvement in the fit of the logarithmic model was achieved by freely estimating the final time step, and we therefore refer to this model as ‘approximately’ logarithmic. We also show results of model with time steps freely estimated (Table 2, Model 7, and see also Figure 1). Results for this model were similar to the approximately logarithmic model but at the cost of two additional parameters. Unlike other models on this table, Model 7 is nested within Model 6 (Model 7 relaxes three additional parameters and fixes one relative to Model 6). Because the models are nested, the chi-square difference test can be use to assess improvement in model fit. The difference in chi-square estimates of model fit for Model 7 relative to Model 6 is 3.67 on 2 degrees of freedom, and therefore Model 7 does not significantly improve model fit relative to Model 6 ($p = .1560$). Thus, following the formal indicators of model fit and substantive and parsimony considerations, we adopted a functional form for AVLT learning that was approximately logarithmic (the fifth time-step freely estimated).

The parameter estimates for this final logarithmic growth model are provided in Table 3. The mean of the initial recall growth factor was about 5.53 words, whereas the learning growth factor had a mean of 5.21 words, indicating that on average participants recalled

10.74 words at the final trial. Significant individual differences for the initial recall factor and the logarithmic learning factor were detected (ψ, Table 3). The variability in the number of words recalled increased over successive trials as a function of individual differences in the rate of learning and as implied by the estimates of the residual variances for the repeated AVLT learning trials (θ, Table 3).

Finally, we note that initial recall and learning were not significantly correlated (model-implied correlation was .05, $p = .71$). We attribute differences in the apparent significance of the correlation of initial recall and learning between the linear change model and the (approximately) logarithmic model to the fact that the linear model is a misspecified model.

## Effect of Covariates

Results for the LGM model with covariates are summarized in Table 4, which includes the regression parameters (and standard errors) describing the change in latent growth curve mean associated with a unit increase in the background variable. Three sets of model results are presented: the influence of each background variable on initial recall and approximately logarithmic learning factors independently, a fitted multivariable model with sociodemographic variables only, and a fitted multivariable model with sociodemographic, health, and cognitive variables. Many sociodemographic, health, and cognitive variables were related to initial recall and learning factors, and many were related to one growth factor and not another. In addition, the fitted models reveal a number of moderator and suppressor effects that are worthy of note. In this section we discuss the model results by background variable domain.

**Sociodemographics—**There was no apparent relationship between sex and initial recall or learning in the independent effects model, but after controlling for the influence of membership in the Black or African-American race/ethnicity group, women had a clear advantage in terms of initial recall, remembering on average about one word more than men. The suppressor effect on initial recall for women was due to a general underrepresentation of Black or African-American male among participants. At the initial trial, Black or African-American respondents recalled on average about 1.5 fewer words than all other race/ethnicity groups combined (predominantly White, non-Hispanic). However, after controlling for sociodemographic, health, and cognitive factors, we note that cognitive factors suppress a greater learning rate for Black or African-American participants, nearly compensating for the differences in initial recall ($\gamma = 0.83$). This suppressor effect is due to the very strong correlation between the race/ethnicity indicator and the cognitive factors (see Table 1).

In the independent effects models, age was not associated with initial recall, but was strongly related to learning. For every 10-year difference in age, participants recalled about one word fewer at the final recall trial. However, after control for other sociodemographic characteristics, age was associated with poorer initial recall as well. This suppressor effect is also due to the race/ethnicity indicator: Black or African-American respondents tended to be younger in comparison to all other race/ethnicity groups (a difference of 3.7 years [95% confidence interval, 1.9–5.5 years]).

After control for the cognitive factors, age remained associated with differences in initial recall, but the age differences in learning were much attenuated and no longer statistically significant in this sample ($\gamma = -.51$, $p = .07$). The association of age and learning rate is partially suppressed by each of general cognition (MMSE), speed of processing (DSST), as well as race/ethnicity. The effect of Black and African-American race/ethnicity group was found to be suppressed by cognitive factors, and emerged as a significant correlate of learning in the final model.

**Health—**The SF-36 mental health component score (MCS) was significantly associated with neither the initial recall factor nor the approximately logarithmic learning factor. The SF-36 physical component score (PCS) was associated with learning in the independent factor model, but in multivariable models failed to emerge as an important independent correlate of recall and learning.

**Cognitive Variables—**Although inductive reasoning ability (measured by word series), psychomotor speed (digit symbol substitution), and verbal knowledge (vocabulary) were related to initial recall independently, not one was related to initial recall in the final fitted model. However, speed, verbal knowledge, and general mental state (MMSE) were related to learning, and the magnitude of the association in the multivariable models was only slightly attenuated relative to the model assessing independent effects. Moreover, as discussed above, statistical control for these cognitive factors suppressed the effects of age and level of education on learning, and modified the relationship of race/ethnicity on learning.

## DISCUSSION

The purpose of the current investigation was to develop a latent growth curve model of recall and learning as assessed by the AVLT in a sample of cognitively normal, healthy older adults. The model demonstrates significant individual differences in level of initial recall and rate of learning. Individual differences in initial recall were related to age, gender and race/ethnicity, while rate of learning was significantly associated with race/ethnicity, processing speed (DSST), verbal knowledge, and cognitive status (MMSE).

Our findings provide evidence for the applicability of a psychometric approach to repeated measures data in clinical neuropsychology, and suggests an alternative conceptualization of outcome data derived from AVLT administration. The fact that learning and recall factors were statistically uncorrelated and differentially related to external variables suggests that fundamentally different neuropsychological processes are captured by the two latent variables. Previous research has demonstrated that the AVLT (Tierney et al., 1994) and similar tests (Hogervorst et al., 2001; Libon et al., 1996) are useful tools as screens for dementia (Gainotti et al., 1998). The screening properties of such tests may be improved by considering a person’s probable location on derived scores corresponding to a latent factor. Insofar as initial recall and learning constructs are uncorrelated, conventional approaches to AVLT data such as summing over five trials (Lezak, 1995) mixes statistically distinct constructs, and may lead to the detection of spurious associations, reduce power to find true associations, and lead to diminished capacity to detect group differences or intervention effects.

Our findings converge with those of previous researchers examining the influence of sociodemographic characteristics on AVLT performance. The most commonly reported of these are age differences in recall and learning (Bolla-Wilson & Bleeker, 1986). For example, in their investigation of list learning performance in 2059 adult males, Uchiyama and colleagues (1995) reported that age, ethnicity, and education were significantly associated with performance on the final trial of the AVLT. Using a general linear multiple regression analysis, it was found that higher scores were associated with younger age and higher education. In another study Mitrushina, Satz, Chervinsky, and D’Elia (1991) found age was related to recall, but not learning, among older adults.

The findings from this study should be interpreted alongside a few limitations. First, data used in these analyses were collected for a pilot program for a cognitive intervention trial, and were not collected to address substantive hypotheses relating to the functional form of

AVLT learning among older adults. Therefore, the available sample limits identification of relationships between background variables and growth factors that describe moderate or larger effects. This analysis is also limited in that the clustered nature of the data (respondents within six field sites) is ignored.

In summary, our results suggest an alternative method for conceptualizing and analyzing performance on a commonly used learning task, the AVLT. Initial recall and learning rates were captured as latent variables in a general structural equations growth curve model. One of the strengths of the latent growth curve approach to longitudinal data analysis is flexibility in curve shape. This permits accurate modeling of individual differences that can subsequently be related to background variables or other factors. Future research should continue the investigation of the most appropriate functional form for the learning curve in larger samples of cognitively normal adults, incorporate the delayed recall and recognition trials into the model, and explore the extent to which the normative trajectory is displayed in samples of persons with neurological disease or injury.

## References


Akaike H. A new look at statistical model identification. IEEE Transactions on Automatic Control 1974;AC-19:716–723.

Anstey K, Christensen H. Education, activity, health, blood pressure and apolipoprotein E as predictors of cognitive change in old age: A review. Gerontology 2000;46:163–177. [PubMed: 10754375]

Bäckman, L.; Small, B.; Wahlin, A.; Larson, M. Cognitive functioning in very old age. In: Craik, F.; Salthouse, T., editors. The handbook of aging and cognition. Mahwah, NJ: Lawrence Erlbaum; 2000. p. 499-558.

Benedict RHB, Schretlen D, Groninger L, Brandt J. Hopkins Verbal Learning Test—Revised: Normative data and analysis of inter-form and test-retest reliability. The Clinical Neuropsychologist 1998;12:43–55.

Bentler PM. Comparative fit indexes in structural models. Psychological Bulletin 1990;107:238–246. [PubMed: 2320703]

Boake C. Edouard Claparede and the auditory verbal learning test. Journal of Clinical and Experimental Neurophysiology 2000;22:286–292.

Bolla-Wilson K, Bleeker M. Influence of verbal intelligence, sex, age and education on the Rey auditory verbal learning test. Developmental Neuropsychology 1986;2:203–211.

Brandt, J.; Benedict, R. Hopkins Verbal Learning Test—Revised: Professional manual. Odessa, FL: Psychological Assessment Resources; 2001.

Browne, M.; Cudek, R. Alternative ways of assessing model fit. In: Bollen, K.; Long, J., editors. Testing structural equation models. Thousand Oaks, CA: Sage; 1993. p. 136-162.

Cockburn J, Smith PT. The relative influence of intelligence and age on everyday memory. Journals of Gerontology: Psychological Sciences 1991;46:31–36.

Delis, D.; Kramer, J.; Kaplan, E.; Ober, B. California Verbal Learning Test. 2. San Antonio, TX: The Psychological Corporation; 2000.

Duncan, TE.; Duncan, SC.; Strycker, LA.; Li, F.; Alpert, A. An introduction to latent variable growth curve modeling: Concepts, issues and applications. Mahwah, NJ: Lawrence Erlbaum Associates; 1999.

Ekstrom, R.; French, J.; Harman, H.; Derman, D. Kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service; 1976. revised ed

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]

Gainotti G, Marra C, Villa G, Parlato V, Chiarotti F. Sensitivity and specificity of some neuropsychological markers of Alzheimer dementia. Alzheimer Disease and Associated Disorders 1998;12:152–162. [PubMed: 9772017]

Heaton RK, Temkin N, Dikmen S, Avitable N, Taylor MJ, Marcotte TD, et al. Detecting change: A comparison of three neuropsychological methods, using normal and clinical samples. Archives of Clinical Neuropsychology 2001;16:75–91. [PubMed: 14590193]

Hogervorst E, Combrinck M, Lapuerta P, Rue J, Swales K, Budge M. The Hopkins Verbal Learning Test and screening for dementia. Dementia and Geriatric Cognitive Disorders 2001;13:13–20. [PubMed: 11731710]

Hultsch DF, Hammer M, Small BJ. Age differences in cognitive performance in later life: Relationships to self-reported health and activity life style. Journals of Gerontology 1993;48:1–11.

Ivnik R, Malec J, Tangalos E, Petersen R, Kokmen E, Kurland L. The auditory-verbal learning test (AVLT): Norms for ages 55 years and older. Psychological Assessment: A Journal of Consulting and Clinical Psychology 1990;2:304–312.

Ivnik R, Malec J, Tangalos E, Petersen R, Kokmen E, Kurland L. Mayo’s older Americans normative studies: Updated AVLT norms for ages 56 to 97. The Clinical Neuropsychologist 1992;6(Suppl): 83–104.

Jobe JB, Smith DM, Ball K, Tennstedt SL, Marsiske M, Willis SL, et al. ACTIVE: A cognitive intervention trial to promote independence in older adults. Controlled Clinical Trials 2001;22:453–479. [PubMed: 11514044]

Jöreskog, K. Testing structural equation models. In: Bollen, K.; Long, J., editors. Testing structural equation models. Newbury Park, CA: Sage; 1992. p. 294-316.

Kurylo M, Temple R, Elliott T, Crawford D. Rey Auditory Verbal Learning Test (AVLT) performance in individuals with recent-onset spinal cord injury. Rehabilitation Psychology 2001;46:247–261.

Lamar M, Resnick SM, Zonderman AB. Longitudinal changes in verbal memory in older adults: Distinguishing the effects of age from repeat testing. Neurology 2003;60:82–86. [PubMed: 12525723]

Lezak, M. Neuropsychological assessment. 3. New York: Oxford University Press; 1995.

Libon D, Mattson R, Glosser G, Kaplan E, et al. A nine-word dementia version of the California Verbal Learning Test. Clinical Neuropsychologist 1996;10:237–244.

Mitrushina M, Satz P, Chervinsky A, D’Elia L. Performance of four age groups of normal elderly on the Rey Auditory Verbal Learning Test. Journal of Clinical Psychology 1991;47:351–357. [PubMed: 2066402]

Muthén, B. Analysis of longitudinal data using latent variable models with varying parameters. In: Collins, LM.; Horn, JL., editors. Best methods for the analysis of change: Recent advances, unanswered questions, future directions. 1. Washington, DC: American Psychological Association; 1991. p. 1-17.

Muthén, LK.; Muthén, BO. Mplus users guide (February 2001, Version 2). Los Angeles, CA: Muthén & Muthén; 1998–2003.

Nettelbeck T, Rabbitt P, Wilson C, Batt R. Uncoupling learning from initial recall: The relationship between speed and memory deficits in old age. British Journal of Psychology 1996;87:593–607. [PubMed: 8962478]

Norman MA, Evans JD, Miller SW, Heaton RK. Demographically corrected norms for the California Verbal Learning Test. Journal of Clinical and Experimental Neuropsychology 2000;22:80–94. [PubMed: 10649547]

Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: Clinical characterization and outcome. Archives of Neurology 1999;56:303–308. [PubMed: 10190820]

Query, WT.; Megran, J. Age-related norms for AVLT in a male patient population; Journal of Clinical Psychology. 1983. p. 136-138.URL-JOURNAL: http://www.intersciencewileycom/jpages/0021-9762/

Raykov T, Tomer A, Nesselroade JR. Reporting structural equation modeling results in Psychology and Aging: Some proposed guidelines. Psychology and Aging 1991;6:499–503. [PubMed: 1777136]

Rey, A. L’examen clinique en psychologie. Paris: Presses Universitaires de France; 1964.

Royall DR, Palmer R, Chiodo LK, Polk M. Decline in learning ability best predicts future dementia type: The Freedom House study. Experimental Aging Research 2003;29:285–406.

Schaie, K. Manual for the Schaie-Thurstone Adult Mental Abilities Test (STAMAT). Palo Alto, CA: Consulting Psychologists Press; 1985.

Smith G, Ivnik R, Malec J, Kokmen E, Tangalos EG, Kurland LT. Mayo's Older Americans Normative Studies (MOANS): Factor structure of a core battery. Psychological Assessment 1992;4:382–390.

Spreen, O.; Strauss, E. A compendium of neuropsychological tests. 2. New York: Oxford University Press; 1998.

Tierney M, Nores A, Snow W, Fisher R, et al. Use of the Rey Auditory Verbal Learning Test in differentiating normal aging from Alzheimer's and Parkinson's dementia. Psychological Assessment 1994;6:129–134.

Uchiyama C, D'Elia L, Dellinger A, Becker J, Selnes O, Wesch J, et al. Alternate forms of the Auditory-Verbal Learning Test: Issues of test comparability, longitudinal reliability, and moderating demographic variables. Archives of Clinical Neuropsychology 1995;10:133–145. [PubMed: 14589735]

Vakil E, Blachstein H. Rey AVLT: Developmental norms for adults and the sensitivity of different memory measures to age. Clinical Neuropsychologist 1997;11:356–369.

van-den-Burg W, Kingma A. Performance of 225 Dutch school children on Rey's Auditory Verbal Learning Test (AVLT): Parallel test-retest reliabilities with an interval of 3 months and normative data. Archives of Clinical Neuropsychology 1999;14:545–559. [PubMed: 14590582]

Verhaeghen P, Marcoen A, Goossens L. Improving memory performance in the aged through mnemonic training: A meta-analytic study. Psychol Aging 1992;7:242–251. [PubMed: 1535198]

Wallman, KK. Older American's 2000: Key indicators of well-being. Washington, D.C: Federal Interagency Forum on Aging Related Statistics; 2000.

Ware, JE.; Kosinski, M.; Keller, SD. SF-36 Physical and Mental Health Summary Scales: A user's manual. Boston: Health Assessment Lab; 1994.

Wechsler, D. Manual: Wechsler Adult Intelligence Scale—Revised. New York: Psychological Corp; 1981.

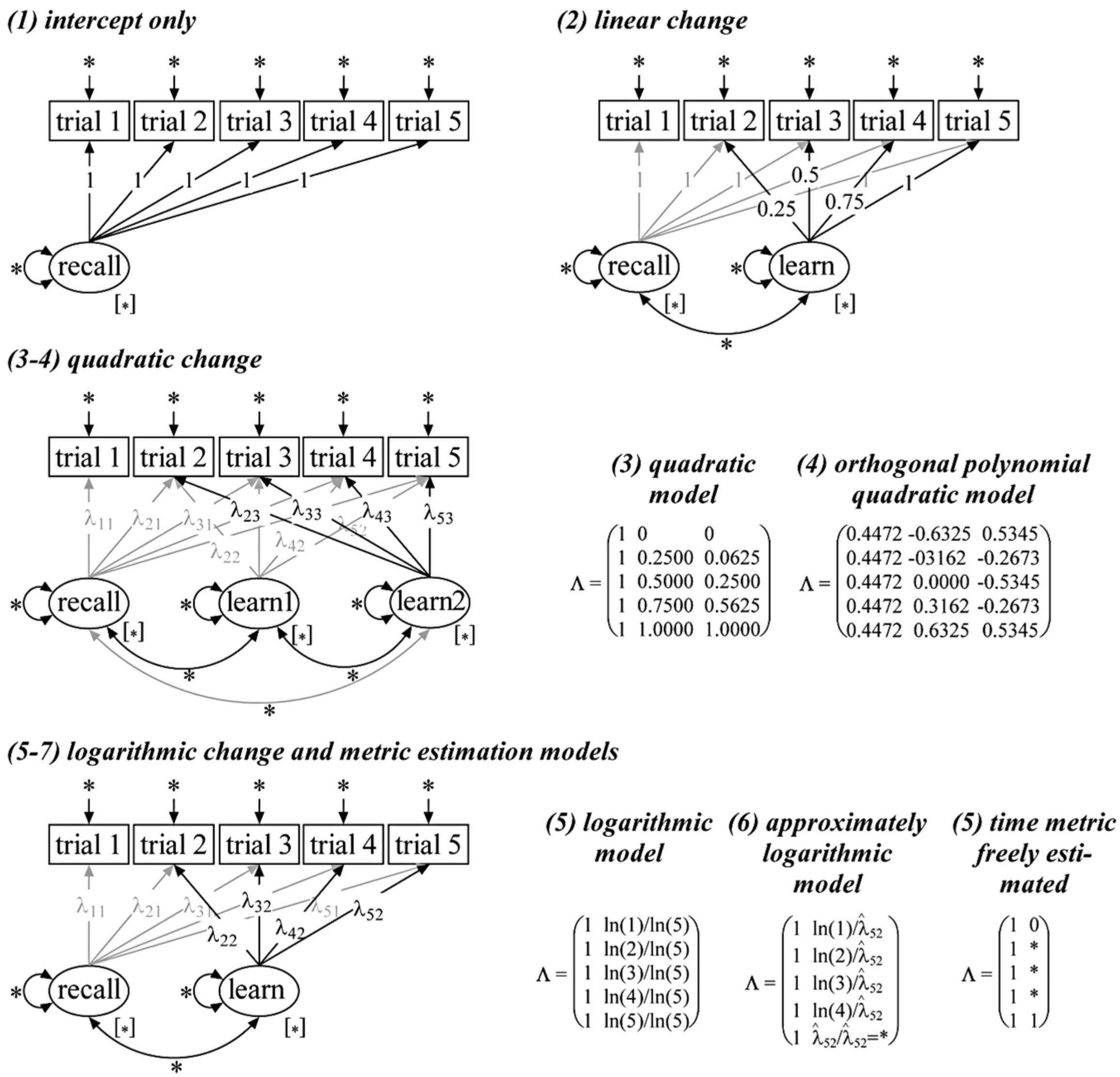


**Figure 1.**

Path diagrams and parameterization details for different learning curve–fitting models. Observed variables (words recalled at successive AVLT trials) shown in rectangles, latent variables capturing random growth model parameters shown as ovals. Free parameters indicated with asterisks (*), fixed parameters shown to their numerical value. Parameters not illustrated are fixed to 0. Mean structures for growth model parameters shown in square brackets. See text for additional detail on model building.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

**Table 1**

Descriptive statistics and correlation matrix: ACTIVE pilot sample ($N$ = 169)

| | | | Correlation matrix | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Characteristic** | **Mean** | **Var** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** | **(10)** | **(11)** | **(12)** | **(13)** | **(14)** |
| (1) Women (indicator, vs. men) | 0.83 | 0.14 | 1.00 | | | | | | | | | | | | | |
| (2) Age (years) | 74.26 | 37.39 | 0.11 | 1.00 | | | | | | | | | | | | |
| (3) Black, African-American (indicator, vs. all others) | 0.56 | 0.25 | 0.19 | −0.31 | 1.00 | | | | | | | | | | | |
| (4) Fewer than 12 years of education (indicator, vs. 12 or more) | 0.32 | 0.22 | 0.04 | 0.10 | 0.15 | 1.00 | | | | | | | | | | |
| (5) SF-36 Physical Component Score (observed range 18–58) | 42.30 | 94.57 | 0.01 | −0.19 | −0.13 | −0.28 | 1.00 | | | | | | | | | |
| (6) SF-36 Mental Component Score (observed range 27–67) | 54.04 | 64.33 | −0.04 | −0.18 | −0.03 | −0.20 | 0.18 | 1.00 | | | | | | | | |
| (7) Word series number correct (range 0–21) | 6.48 | 20.58 | −0.10 | −0.18 | −0.35 | −0.40 | 0.24 | 0.25 | 1.00 | | | | | | | |
| (8) DSST number correct (range 0–74) | 34.68 | 222.31 | 0.01 | −0.28 | −0.32 | −0.37 | 0.31 | 0.24 | 0.62 | 1.00 | | | | | | |
| (9) Vocabulary number correct (range 0–14) | 6.10 | 12.06 | −0.06 | 0.04 | −0.31 | −0.48 | 0.19 | 0.16 | 0.59 | 0.49 | 1.00 | | | | | |
| (10) MMSE score (observed range 23–30) | 26.98 | 4.70 | −0.13 | −0.17 | −0.26 | −0.40 | 0.23 | 0.16 | 0.52 | 0.51 | 0.45 | 1.00 | | | | |
| (11) AVLT trial 1 number of words recalled (observed range 1–10) | 5.52 | 2.88 | 0.08 | −0.02 | −0.32 | −0.12 | 0.14 | 0.09 | 0.20 | 0.26 | 0.22 | 0.13 | 1.00 | | | |
| (12) AVLT trial 2 number of words recalled (observed range 2–14) | 7.95 | 5.06 | 0.12 | −0.12 | −0.27 | −0.23 | 0.24 | 0.15 | 0.41 | 0.47 | 0.41 | 0.31 | 0.68 | 1.00 | | |
| (13) AVLT trial 3 number of words recalled (observed range 3–15) | 9.47 | 6.18 | 0.10 | −0.25 | −0.23 | −0.23 | 0.24 | 0.12 | 0.39 | 0.47 | 0.33 | 0.42 | 0.65 | 0.80 | 1.00 | |
| (14) AVLT trial 4 number of words recalled (observed range 2–15) | 10.26 | 7.31 | 0.11 | −0.22 | −0.17 | −0.27 | 0.25 | 0.12 | 0.44 | 0.50 | 0.38 | 0.44 | 0.55 | 0.75 | 0.86 | 1.00 |
| (15) AVLT trial 5 number of words recalled (observed range 0–15) | 10.74 | 8.76 | 0.11 | −0.26 | −0.19 | −0.21 | 0.20 | 0.16 | 0.49 | 0.53 | 0.45 | 0.45 | 0.51 | 0.72 | 0.79 | 0.83 |

*Abbreviations:* Var, variance; DSST, Digit Symbol Substitution Test; MMSE, Mini-Mental State Examination.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

**Table 2**

Summary of curve fit model building for AVLT initial recall and learning. ACTIVE pilot sample ($N = 169$)

| | | | | | | | | Growth parameter estimates | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Intercept | | Slope | | Quadratic | |
| **Model** | $\chi^2$ | *df* | *p* | **AIC** | **Parsimony ratio** | **RMSEA** | **CFI** | **Mean** | **Var** | **Mean** | **Var** | **Mean** | **Var** |
| (1) Intercept only | 754.894 | 13 | <.001 | 3890.329 | 0.867 | 0.581 | 0.000 | 9.498 | 5.521 | na | na | na | na |
| (2) Linear change | 204.955 | 10 | <.001 | 3346.390 | 0.667 | 0.340 | 0.733 | 6.329 | 2.218 | 5.207 | 3.018 | na | na |
| (3) Quadratic change | 9.185 | 6 | 0.163 | 3286.168 | 0.400 | 0.056 | 0.996 | 5.548 | 2.344 | 10.523 | 18.177 | −5.433 | 7.015 |
| (4) Quadratic change—orthogonal | 9.409 | 6 | 0.152 | 3158.844 | 0.400 | 0.058 | 0.995 | 19.612 | 21.817 | 4.022 | 2.538 | −1.269 | 0.382 |
| (5) Logarithmic change | 21.686 | 10 | 0.168 | 3163.121 | 0.667 | 0.083 | 0.984 | 5.558 | 2.455 | 5.469 | 4.607 | na | na |
| (6) Approximately logarithmic change | 9.593 | 9 | 0.384 | 3153.027 | 0.600 | 0.020 | 0.999 | 5.531 | 2.525 | 5.207 | 4.321 | na | na |
| (7) Time metric freely estimated | 5.919 | 7 | 0.549 | 3153.353 | 0.467 | 0.000 | 1.000 | 5.519 | 2.544 | 5.199 | 4.336 | na | na |

*Abbreviations:* AVLT, Rey Auditory Verbal Learning Test; *df*, degrees of freedom; AIC, Akaike Information Criteria; RMSEA, root mean square error of approximation; CFI, Comparative Fit Index; Var, variance. Parsimony ratio defined as $df/[.5k(k+1)]$ where $k$ is the number of observed variables.

## Table 3

Fitted approximately logarithmic AVLT performance growth model: ACTIVE pilot (*n* = 169)

| Model part | Parameter | Estimate (SE)[a] |
|---|---|---|
| Time scores | | |
| First trial | $\lambda_{12}$ | *0.000*[b] |
| Second trial | $\lambda_{22}$ | *0.464* |
| Third trial | $\lambda_{32}$ | *0.736* |
| Fourth trial | $\lambda_{42}$ | *0.928* |
| Fifth trial | $\lambda_{52}$ | 0.996 (0.023) |
| Growth factor means | | |
| Initial recall (words) | $\alpha_1$ | 5.531 (0.131) |
| Learning (words at trial five) | $\alpha_2$ | 5.207 (0.181) |
| Growth factor variance/covariance | | |
| Initial recall | $\psi_{11}$ | 2.525 (0.364) |
| Learning | $\psi_{22}$ | 4.321 (0.674) |
| Covariance of initial recall, learning | $\psi_{21}$ | 0.147 (0.39)[c] |
| Observed variable (AVLT) residual variance | | |
| First trial | $\theta_{11}$ | 0.387 (0.240) |
| Second trial | $\theta_{22}$ | 1.322 (0.170) |
| Third trial | $\theta_{33}$ | 0.857 (0.132) |
| Fourth trial | $\theta_{44}$ | 0.933 (0.170) |
| Fifth trial | $\theta_{55}$ | 1.996 (0.278) |

[a]Standard error.

[b]Entries in italic typeface are not estimated, but fixed equal to the shown value.

[c]Parameter estimate not significantly different from zero.

NIH-PA Author Manuscript
NIH-PA Author Manuscript



## Table 4

Independent and fitted relationship of background variables and initial recall and learning factors for AVLT performance: ACTIVE pilot sample ($N$ = 169)

| | Independent effects | | | | Fitted multivariate model sociodemographic characteristics only | | | | Fitted multivariate model sociodemographic, health, and cognition characteristics | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Initial recall | | Learning | | Initial recall | | Learning | | Initial recall | | Learning | |
| **Characteristic** | **Est.** | **(*SE*)** | **Est.** | **(*SE*)** | **Est.** | **(*SE*)** | **Est.** | **(*SE*)** | **Est.** | **(*SE*)** | **Est.** | **(*SE*)** |
| Women (vs. men) | 0.390 | (0.345) | 0.458 | (0.469) | 0.927 | (0.329)* | 0.000 | — | 0.969 | (0.325)* | 0.000 | — |
| Age (per 10 years) | −0.068 | (0.215) | −1.141 | (0.278)* | −0.477 | (0.211)* | −1.062 | (0.274)* | −0.473 | (0.211)* | −0.507 | (0.277) |
| Black, African-American (vs. Whites and others) | −1.107 | (0.249)* | 0.047 | (0.357) | −1.446 | (0.260)* | 0.000 | — | −1.428 | (0.263)* | 0.832 | (0.342)* |
| Completed fewer than 12 years of education (vs. more) | −0.458 | (0.278) | −1.103 | (0.371)* | 0.000 | — | −1.006 | (0.353)* | 0.000 | — | 0.318 | (0.355) |
| SF-36 physical component score (per 10 points) | 0.257 | (0.140) | 0.493 | (0.195)* | — | — | — | — | 0.000 | — | 0.000 | — |
| SF-36 mental component score (per 10 points) | 0.182 | (0.172) | 0.306 | (0.244) | — | — | — | — | — | — | — | — |
| Word series number correct | 0.077 | (0.028)* | 0.213 | (0.036)* | — | — | — | — | 0.000 | — | 0.000 | — |
| DSST number correct | 0.030 | (0.008)* | 0.069 | (0.011)* | — | — | — | — | 0.000 | — | 0.040 | (0.013)* |
| Vocabulary number correct | 0.110 | (0.037)* | 0.223 | (0.048)* | — | — | — | — | 0.000 | — | 0.122 | (0.053)* |
| MMSE score | 0.102 | (0.060) | 0.497 | (0.073)* | — | — | — | — | 0.000 | — | 0.328 | (0.080)* |
| Latent variable $r^2$ | — | | — | | 0.190 | | 0.165 | | 0.188 | | 0.445 | |

*Abbreviations:* AVLT, Rey Auditory Verbal Learning Test; DSST, Digit Symbol Subtitution Test; MMSE, Mini-Mental State Examination; Est., parameter estimate; *SE*, standard error.

* $p$ <.05;

— in estimate column indicates parameter not included in model; entries of 0 in estimate column and — in *SE* column indicate parameter fixed to 0.