

SPECIAL ISSUE

ACQUISITION, RECALL, AND FORGETTING OF VERBAL
INFORMATION IN LONG-TERM MEMORY BY YOUNG,
MIDDLE-AGED, AND ELDERLY INDIVIDUALS

Hasker P. Davis¹, Scott A. Small², Yaakov Stern², Richard Mayeux²,
Simeon N. Feldstein³ and Frederick R. Keller¹

(¹University of Colorado, Colorado Springs; ²Gertrude H. Sergievsky Center, Columbia University College of Physicians and Surgeons, New York; ³Psychology Department, Georgia Institute of Technology, Atlanta)

ABSTRACT

Memory performance by four age groups (30-45 years, 46-60 years, 61-75 years, and 76-90 years) was compared on a multi-trial verbal recall task with 20-minute and 1-day delay free recall and recognition trials. The rate of acquisition across 5 learning trials was similar for all ages except the youngest group whose performance was constrained by a ceiling effect. The level of acquisition achieved was less in the two oldest groups. Words gained across trials and words lost across trials made similar contributions to the shape of the learning curve for the acquisition trials. Subjective organization decreased with age, but remained strongly related to the number of words recalled during acquisition for all age groups. The two oldest age groups demonstrated significant declines in words recalled on the 20-minute and 1-day delay trials. A subset of the oldest group demonstrated more rapid forgetting at the 1-day delay when participants from all age conditions were matched on acquisition. Thus, many aspects of free recall were impaired with age, and variance measurement of recall showed greater inter-individual differences with increasing age. This increase in individual differences could reflect a single form of age-related memory impairment, or it could indicate that memory impairment in the elderly is due to multiple processes. The importance of testing across the life span and using tests that examine a variety of memory components and processes for establishing norms and clarifying age-related deficits are discussed.

Key words: acquisition, recall, and forgetting of verbal information in long-term

MEMORY BY YOUNG, MIDDLE-AGED AND ELDERLY INDIVIDUALS

It is well established that the elderly do not retain recently acquired information to the same extent as the young (Kausler, 1982, 1994). The evidence for this observation comes in large part from studies examining the learning and retention of verbal material. One might assume that after more than a hundred years of studies using verbal material (Ebbinghaus, 1885), that little new knowledge about memory would be gained from another study of verbal recall or recognition. However, this is not the case, and particularly for investigations of age-related changes in learning and memory. Indeed, the development of pharmacological agents for the treatment of memory deficits associated with Alzheimer's disease (e.g., Aricept and Exelon), increases the importance of developing strategies for early detection of memory decline. Further, the

diagnostic categories of Age-Associated Memory Impairment (AAMI), Age-Related Cognitive Decline (ARCD), and Mild Cognitive Impairment (MCI) (Chen et al., 2001; Ferris and Kluger, 1996; Larrabee, 1996; Richards et al., 1999), and their possible treatment with pharmacological agents (e.g., Crook and Youngjohn, 1993), provide an additional rationale for the development of carefully established norms that extend over age ranges that include early onset Alzheimer's disease and the beginning of AAMI, ARCD, or MCI. Indeed, behavioral assessments and establishment of norms is critical to virtually all neuropsychological evaluations. For example, a genetic marker such as the apolipoprotein-4 allele is a valuable indicator of risk for the development of Alzheimer's disease, but behavioral assessment is needed to establish disease onset.

Recently, the cognitive contributions that mediate age-related changes in acquisition and retention of verbal material have been greatly expanded. For example, Dunlosky and Salthouse (1996) decomposed performance on a multi-trial learning task of verbal material into acquisition and forgetting components. These authors suggested that acquisition probably was a greater contributor to the age-related decline in multi-trial learning than forgetting. Carlesimo et al. (1997) examined age effects on forgetting at a 15-minute delay as a function of serial position after multi-trial learning. They suggested that the recency component (words occurring in the last few positions of a word list) was differentially impaired in the elderly.

Another contribution to age-related memory deficits that has remained intractable to clarification, and where findings have been highly inconsistent, is the rate at which information in long-term memory is forgotten. Factors that have played critical roles in rates of forgetting are the time between acquisition and testing, the form of the retention test (i.e., recall or recognition), and the level of acquisition achieved by the participants. The effect of retention interval on rates of forgetting has been reviewed by Kausler (1994) and Salthouse (1991). There are conflicting results, but the general finding is that rates of forgetting for different age groups tend to be similar at short retention intervals. The elderly tend to show greater rates of forgetting at longer intervals, typically 24 hours or greater. However, when recognition tests rather than recall tests are used at longer delays, there are no or only small age differences in the rate of forgetting (e.g., Rybarczyk, et al., 1987). At short retention intervals where recognition tests have been employed, the typical pattern is no age difference (e.g., Lehman and Mellinger, 1986).

Additionally, many studies of age-related memory impairment are restricted by age range. That is, the age groups typically compared are a single group of young adults in their twenties and thirties and a single group of older adults usually in their sixties or older. This restriction of age groups results in a failure to detect changes in mid-life, or graded changes within a group sixty years of age and older.

The purpose of this study is to examine selected aspects of verbal memory that provide information on when people begin to forget verbal material, and how various components of retention change across the life span. We use a wide variety of measures that may be related to different neural mechanisms of memory (e.g., subjective organization may relate to prefrontal cortex function,

forgetting may relate to mid-temporal function). Participants ranging in age from 30 to 90 years of age were administered multi-trial free recall versions of the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964) using a within-group design to examine two different retention intervals. Four age groups of equal range, between 30 and 90, were tested from two geographical locations (Colorado Springs, CO and New York City, NY). The advantage of comparing these multiple age ranges is the detection of early changes in midlife, and late developing changes in the participants over sixty years of age. Drawing participants from different geographical locations increases any generalizations that may be drawn from the results. In addition to assessing forgetting, the tests employ multiple trials that allow for examining the rate and level of acquisition as they vary with age. The contributions of different memory processes and components were examined by assessing free recall and recognition, subjective organization, primacy and recency effects on acquisition and retention trials, and gained access and loss during acquisition (Dunlosky and Salthouse, 1996). Finally, since a prevalent concern with memory research for forgetting of verbal material for different age groups is the level of initial acquisition, we controlled for this inherent difference by matching a subset of participants on level of acquisition.

METHOD

Participants

Participants in Colorado Springs ($n = 97$) were recruited from Psychology classes at the University of Colorado at Colorado Springs or from local senior citizen organizations. Participants in New York City ($n = 39$) were recruited from an ongoing study of aging and cognition. Table I shows the number of subjects from the two samples, age range and mean age, mean educational level, mean full scale IQ, and ratio of males to females for each of the four age groups combined and by geographical location. Participants had no history of neurological disease, head trauma, learning disability, major psychiatric illness, or drug use that might affect cognition. All participants reported they were taking no drugs that they believed would affect their thinking, all were living independently in the community, and all self-reported fair or good health. Participants attended three testing sessions and were given \$10 per session.

Rey Auditory Verbal Learning Test (RAVLT)

Participants were administered two word-frequency matched versions of the Rey Auditory Verbal Learning Test (RAVLT) (Rey, 1964) to assess acquisition and recall of verbal material at two different retention delays. A list of 15 concrete nouns was presented orally five times in succession at the rate of one word every two seconds (intertrial interval = 20 seconds). The order of word presentation was randomized on each trial, and the order of word presentation was the same for all participants. Participants attempted to recall the words

TABLE I
Demographic and Neuropsychological Data Combined and by Geographical Location

	N	Age	Years Education	Shipley IQ	Male/Female Ratio
30-45 years Combined	32	37.2 ± 4.7	15.4 ± 2.0	107.7 ± 12.0	9/23
Colorado	22	38.7 ± 4.3	15.0 ± 1.3	111.4 ± 8.1	6/16
New York	10	33.8 ± 3.9	16.4 ± 2.8	102.7 ± 16.8	3/7
46-60 years Combined	29	54.8 ± 4.1	15.7 ± 2.8	109.7 ± 11.1	11/18
Colorado	17	53.9 ± 3.8	15.5 ± 2.4	113.7 ± 9.2	7/10
New York	12	56.1 ± 4.3	16.0 ± 3.4	107.6 ± 14.2	4/8
61-75 years Combined	40	67.8 ± 5.2	14.4 ± 2.3	116.2 ± 14.9	9/31
Colorado	32	66.8 ± 5.1	14.4 ± 2.3	114.9 ± 14.8	6/26
New York	8	71.9 ± 3.6	14.4 ± 2.3	120.4 ± 11.5	3/5
76-90 years Combined	35	80.2 ± 3.6	14.5 ± 2.4	122.6 ± 14.0	12/23
Colorado	26	80.4 ± 4.0	14.2 ± 2.4	122.6 ± 11.9	9/17
New York	9	79.4 ± 3.7	15.2 ± 2.4	122.2 ± 13.0	3/6

Values are means ± standard deviations.

orally at the end of each trial. Following the five learning trials, a delay trial was given either 20 minutes or 1 day later, and a recognition test was given immediately after the delay trial. Finally, a single study and test trial involving 15 new words was given immediately after the recognition test (Interference Trial). For the recognition trial the 15 target words and 15 frequency-matched distractor words were presented orally one at a time, and participants were asked to indicate whether or not each word was from the study list (Squire and Shimamura, 1986). In summary, participants either had five learning trials and were then tested 20 minutes later, or they had five learning trials and were tested 24 hours later. The two different learning sessions were scheduled at least 7 days apart. Full counterbalancing was achieved on order of testing for the youngest and oldest age groups, but not for the two intermediate age groups due to scheduling problems and coordination between the two testing sites. The major differences in this study were found between the youngest and oldest age groups, and therefore, counterbalancing does not in any way appear to impact the results obtained in this study.

Shipley

The Shipley Institute of Living Scale was also administered to all participants and scored following standard scoring procedures as described by Zachary et al. (1985). The Shipley provides age-estimated WAIS-R IQ scores.

Procedure

Participants were told that the purpose of the study was to assess memory abilities. They were told they would be given an oral test of verbal memory and

a paper and pencil test of thinking ability (Shipley). All testing was carried out in quiet rooms at either the University of Colorado at Colorado Springs or Columbia University. Each testing session required 30 minutes to 1 hour.

The instructions for the RAVLT were as follows:

Trial 1: I'm going to read you a list of words one at a time. Listen carefully, because when I'm done, I'd like you to tell me as many words as you can remember; it doesn't matter in what order.

Trials 2-5: I'm going to read you the list of words again. They will be in a different order. When I'm done, tell me as many as you can remember, including words you have already said. It doesn't matter in what order.

Delayed Recall: Remember that list of words we were reading before? Tell me all the words you can still remember from it; it doesn't matter in what order.

Recognition Trial: Now I'm going to read you a longer list of words in which some of the words will be from the list – the one we went over earlier. Some of the words will not be from that list. I would like you to indicate if the word was from the list we kept going over by saying "yes". If the word was not from the studied list, please say "no". Do you have any questions? Are you ready?

Interference Trial: I'm going to read you a new list of words. Listen carefully, because when I'm done, I'd like you to tell me as many words as you can remember; it doesn't matter in what order.

RESULTS

Demographics – New York and Colorado Springs Samples

We compared each age group from the two locations with respect to distribution of the age of subjects in the two samples, mean age, years of education, full-scale IQ, male/female ratio, and number of words recalled on each test trial. An analysis of variance (ANOVA) yielded an interaction between age and testing site, $F(3, 129) = 6.80$, $MSE = 18.1$, $p < .0005$. For the 30 to 45 year-old group, the Colorado participants ($M = 38.7$, $SE = .91$) were older than the New York participants ($M = 33.8$, $SE = 1.34$), $LSD = 4.93$, $p = .003$. For the 61 to 75 year-old group, the New York participants ($M = 71.9$, $SE = 1.50$) were older than the Colorado participants ($M = 66.8$, $SE = .75$), $LSD = 5.06$, $p = .003$. Given that within-group age differences increase the age variability for those groups, the pooling of data for New York and Colorado participants should improve the generalizability of the findings, providing other systematic differences do not occur between the groups.

The New York and Colorado groups were similar with respect to age distribution of the two samples as measured by the Kolmogorov-Smirnov test ($p > 0.2$). The mean age, full-scale IQ, years of education, and male/female ratio also showed no detectable differences between locations ($p > 0.1$). On the memory tests themselves, the only difference between the two samples was an interaction between age group and test site for the final (Interference) trial in the 1-day delay condition, $F(3, 129) = 4.60$, $MSE = 3.58$, $p = .004$. For the 30 to 45 year-old group, the recall score of the Colorado group on this trial ($M = 7.6$, SE

= .40) was higher than that of the New York group ($M = 6.0$, $SE = .60$), $LSD = 1.59$, $p = .029$. For the 76 to 90 year-old group, the New York group ($M = 5.4$, $SE = .60$) scored higher than the Colorado group ($M = 3.8$, $SE = .37$), $LSD = 1.55$, $p = .029$. Given that no other differences in memory test performance were found between sites, it is likely that the difference detected on the Interference Trial on the 1-day delay condition was random rather than systematic. Further, main effects and interactions were similar in approximately 70% of comparison when the two samples were separately assessed. Most of the remaining cases showed trends in the same direction, although nonsignificant. Only a very few comparisons did not follow this pattern. Accordingly, it seems reasonable to conclude that population sampling, test administration, and test performance were consistent across the two sites and that pooling groups from the two locations in no way complicates the findings. Indeed, increasing power by pooling the two samples increases the likelihood of detecting significant results.

Rate of Acquisition

Performance on the first five recall trials was first examined with number of words recalled as the dependent variable and with independent variables of delay (the 20-minute and 1-day conditions) and trials (Trials 1 through 5), and the between-group variable of age group (4 age groups). An omnibus F test was run ($\alpha = .05$). Main effects were found for trial: $F(3.27, 431.26) = 673.21$, $MSE = 2.10$, $p < .0005$; and for age group, $F(3, 132) = 23.59$, $MSE = 2.64$, $p < .0005$. In addition, a three-way interaction was found (trial number \times age group \times delay condition), $F(10.62, 467.50) = 2.68$, $MSE = 1.65$, $p = .003$, with the Greenhouse-Geisser epsilon used to correct for violation of sphericity. Figures 1a and 1b show the mean words recalled (\pm SEM) on each trial and in each delay condition.

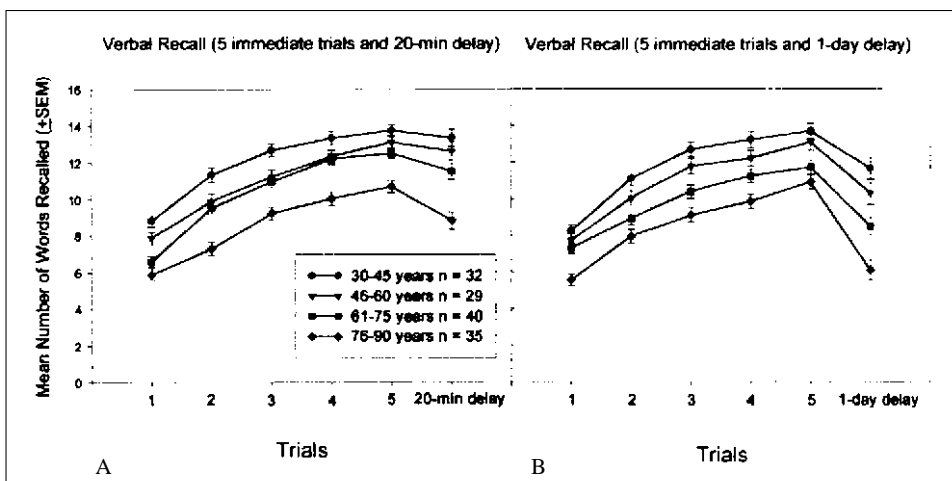


Fig. 1a-b - Mean number of words recalled by participants in four age groups on the Rey Auditory Verbal Learning Test. The left panel shows performance on five consecutive trials and on a 20-minute delay trial. The right panel shows performance on five consecutive trials and on a 1-day delay trial.

An unpredicted finding was the three-way interaction involving delay. In order to clarify the effects of age on rate of acquisition at the two different delays, separate ANOVAs were performed on words recalled with trials and age condition as the independent variables. No significant interaction between trials and age was detected at the 1-day delay. At the 20-minute delay, a significant interaction between trials and age was detected, $F(10.5, 463.4) = 2.3, p = .01$. Pairwise comparisons were performed to detect differences between trials at each age group ($\alpha = .01$). The youngest age group showed no significant improvement from Trial 3 to Trial 4 and from Trial 4 to Trial 5. The participants 61 to 75 years of age showed no significant improvement from Trial 4 to Trial 5. The other two age groups showed significant trial-to-trial improvement for all comparisons. The detection of asymptote performance on the acquisition trials for the 20-minute delay, but not on the 1-day delay, is most likely the basis for the significant three-way interaction reported for the omnibus F test. The lack of asymptote performance at the 1-day delay suggests that no age difference in the rate of acquisition represents the most likely state of affairs.

However, conclusions from the analysis regarding rate of acquisition must necessarily be limited. We found age differences in rate of acquisition, and at one of the delays the youngest age group clearly shows a more rapid rate of acquisition and essentially achieves asymptote by Trial 3. While we do not believe that the difference across delays is systematic, the potential of asymptotic acquisition performance by the youngest group presents a problem for subsequent findings on rates of forgetting. We address this problem in our study by matching subjects on rate and level of acquisition, and we address asymptotic performance by the young in the Discussion. Despite a significant three-way interaction consistent with differential rates of acquisition, no unequivocal evidence for age-related differences in rate of acquisition was obtained. Specifically, the achievement of asymptotic performance by the young is a methodological limitation that precludes strong conclusions about rates of acquisition. The limitations imposed on other aspects of test results will be addressed in the Results and Discussion where appropriate.

Level of Acquisition

Level of acquisition refers to the number of words recalled on the last acquisition trial. To measure level of acquisition, performance by age at Trial 5 was examined using the results from the omnibus F test (noted above) and post-hoc comparisons ($\alpha < .01$). For Trial 5 at the 20-minute delay condition, the total number of words recalled by 30 to 45 year-old participants ($M = 13.72, SE = .34$) was significantly greater than the number of words recalled by the 61 to 75 year-olds ($M = 12.48, SE = .31$) and the 76 to 90 year-old participants ($M = 10.69, SE = .32$), $LSD = 1.24$ and $3.03, p = .007$ and $p < .0005$, respectively. Additionally, both the 46 to 60 year-old ($M = 13.10, SE = .36$) and the 61 to 75 year-old participants recalled significantly more words at Trial 5 than did the 76 to 90 year-old participants, $LSD = 2.42$ and $1.79, ps < .0005$, respectively. Within each age group, the mean number of words recalled was comparable between the 20-minute and the 1-day delays. Comparisons of recall performance

between the four age groups for Trial 5 at the 1-day delay showed the same pattern of significant differences, with the exception that the recall performance of the 61 to 75 year-old participants did not differ significantly from that of the 76 to 90 year-olds. In summary, we find that overall level of acquisition, the number of words recalled on the fifth learning trial, was less in the older age groups. The magnitude of this age-related difference might have been greater if there was not a ceiling effect imposed by the list length on the performance of the youngest age group.

Finally, the lack of support for age differences in rate of acquisition suggests that age differences in level of acquisition might be accounted for by age difference on Trial 1. An ANOVA performed on Trial 1 with age and delay as independent variables detected no significant difference for delay, $F(1, 132) < 1.0$, $p = .69$. Age effects were identical to those reported for Trial 5.

Acquisition Processes

To further evaluate the impact of age on acquisition, we next assessed gained access and lost access using a modification of the method described in Dunlosky and Salthouse (1996). The gained access score was defined as the number of words recalled on trial $n + 1$ that were not recalled on trial n . The new words recalled across trials is one of the components contributing to the learning curve. An ANOVA was run on gained access scores with age condition as the between subjects variable and delay and trials as the within subjects variables. There were two delay conditions as before, however, the trials variable in this case consisted of a gained access score between Trials 1 and 2, Trials 2 and 3, Trials 3 and 4, and Trials 4 and 5. A significant three-way interaction effect, $F(9, 47.76) = 3.15$, $MSE = 1.69$, $p = .001$ was detected, and main effects were found for trials, $F(3, 407.84) = 68.67$, $MSE = 1.98$, $p < .0005$, and for age condition, $F(3, 119) = 10.29$, $MSE = .25$, $p < .0005$.

Post-hoc analyses were conducted to determine patterns of differences in gained access across study trials, delays, and age groups ($\alpha < .01$). The gained access scores for 20-minute trials and 1-day trials are shown in Figures 2a and b. Participants in the 30 to 45 year-old age group showed significantly greater gained access for the verbal material between Trials 1-2 than was found between Trials 2-3, Trials 3-4, and Trials 4-5. Additionally, significantly greater gained access occurred between Trials 2-3 than between Trials 4-5 (for both delay conditions), and Trials 2-3 gained access was significantly larger than Trials 3-4 for the 1-day delay condition. No significant differences were found between Trials 3-4 and Trials 4-5 for the 30 to 45 year-olds. The most likely reason for this lack of gained access for the youngest group is that many of the subjects are limited by there being virtually no new words to recall across trials. Thus, gained access demonstrated another example of the ceiling effect limiting the younger subject's performance.

For both delay conditions, 46 to 60 year-olds obtained significantly larger gained access scores between Trials 1-2 and 2-3 than between Trials 4-5. For the 1-day delay trials, gained access for Trials 1-2 and 2-3 were significantly larger than gained access for Trials 3-4. For the 61 to 75 year old participants at the 20

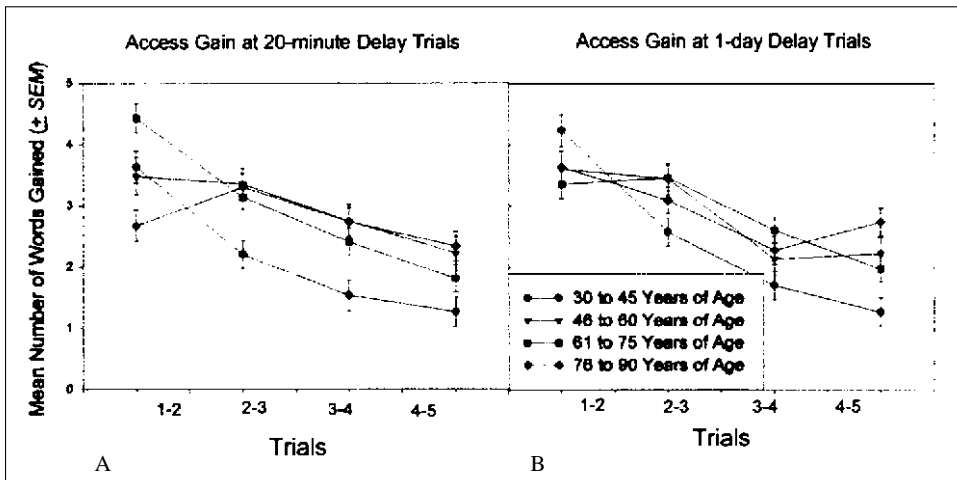


Fig. 2a-b – The panels show the mean number of gained access words at the 20-minute and 1-day delay trials for the different age groups at each intertrial increment. The gained access score was defined as the number of words recalled on trial $n + 1$ that were not recalled on trial n . This component of learning is thought to reflect acquisition.

minute delay condition, gained access was significantly larger for previous than for subsequent intertrials (i.e., trial pairs). That is, Trials 1-2 gained access score was larger than Trials 2-3, 3-4, and 4-5; Trials 2-3 gained access was larger than Trials 3-4 and 4-5 scores. The only intertrial gained access scores that were not significantly different in the 20-minute delay condition were Trials 3-4 and 4-5. For the 1-day delay with 61 to 75 year-old participants the only pairs of intertrial scores where gained access scores were not significantly different from one another were Trials 1-2 and 2-3. For the remaining gained access scores at the 1-day delay, Trials 2-3 were significantly greater than Trials 3-4 and 4-5, and gained access scores were greater between Trials 3-4 than Trials 4-5.

For the 76 to 90 year-olds in the 20-minute delay condition, the only significant difference found in gained access was that Trials 2-3 was significantly larger than Trials 4-5. For the 1-day delay, the Trials 1-2 gained access score was significantly larger than for Trials 3-4 and 4-5, and the Trials 2-3 gained access was significantly larger than that of Trials 3-4. In summary, greater gained access to words from one trial to the next appears to occur in the earlier trials for 30 to 45 year-olds (Trials 1-2, and to a lesser degree Trials 2-3). The other age groups showed the most pronounced gained access to words between Trials 1-2 and 2-3 as well, but gained access between these two intertrials were nearly equivalent. The early increase in gained access by the youngest participants may reflect their achieving a near asymptotic performance sooner than the older groups.

Although intertrial differences across trials appears to show some variation in gained access scores depending on the delay condition, only two significant differences were found when gained access was compared across the delay conditions. For participants in the 61 to 75 year old group, the Trials 1-2 gained access score was significantly larger in the 20-minute delay group than in the 1-

day delay group ($p = .001$). Conversely, the Trials 1-2 gained access score for 76 to 90 year-olds was significantly larger in the 1-day delay group than in the 20-minute delay group ($p = .006$). Thus, there is no systematic difference in gained access on the first five trials for the two delays.

Finally, we examined pairwise comparisons for gained access scores between the different age groups at each delay and intertrial condition. For Trials 1-2, gained access scores were significantly larger (across delay conditions) for the 30 to 45 year-old participants than for the 76 to 90 year-olds, $p = .003$. However, for Trials 2-3, 3-4, and 4-5 all age groups older than the 30 to 45 year-olds obtained significantly greater gained access scores than the 30 to 45 year-old group. No other age group differences were detected. Comparison of this finding with the results from the previous section on recall shows that the Trial 1 mean recall score for the 30 to 45 year-olds leaves less room for gain over later trials, as this age group reaches asymptote earlier than do the older age groups.

The results for gained access in this study appear almost directly opposite to the findings reported by Dunlosky and Salthouse (1996). For example, they show lowest gained access for Trials 1-2, and we show gained access to be greatest for Trial 1-2. These differences are more illusory than real, and most likely reflect differences in the measurement for presentation of gained access, and to a lesser extent the methodology for testing of recall. Dunlosky and Salthouse used a proportion score for gained access that is defined as “the proportion of items recalled on trial $n + 1$ of those that were not recalled on trial n ”. If there are few unrecalled words on the later trials (as is the case for the youngest group), then one or two new words on the subsequent trial may give a high proportional gained access score. One or two words for older subjects with a larger number of unrecalled words from the previous trial will have a lower proportional gained access score. It is for this reason that we choose to present gained access scores in terms of absolute scores, despite an initial analysis that resulted in findings more similar to those reported by Dunlosky and Salthouse. The difficulty with the gained access scores again reflects the problems created by the asymptotic performance of the youngest subjects. Finally, the difference in these studies may also reflect differences in methodology. We presented words at the rate of one word every two seconds and the words are in a different presentation order for each acquisition trial. Dunlosky and Salthouse presented words at the rate of one word every second and the words were presented in the same order on all trials.

Next, we inspected mean differences for lost access scores during the first five acquisition trials (again using a modification of the method from Dunlosky and Salthouse, 1996) across age condition, delay, and trials. Lost access is defined as the total specific words recalled in trial n that were not subsequently recalled in trial $n + 1$. The lost access scores are shown in Figures 3a and 3b. Lost access scores were analyzed using a three-way ANOVA with the aforementioned variables. In this analysis a three-way interaction was not obtained. However, a main effect was seen for age condition, $F(3, 116) = 9.42$, $MSE = .29$, $p < .0005$, and a two-way interaction was found for age condition by trial, $F(9, 357) = 2.25$, $MSE = 1.26$, $p = .018$.

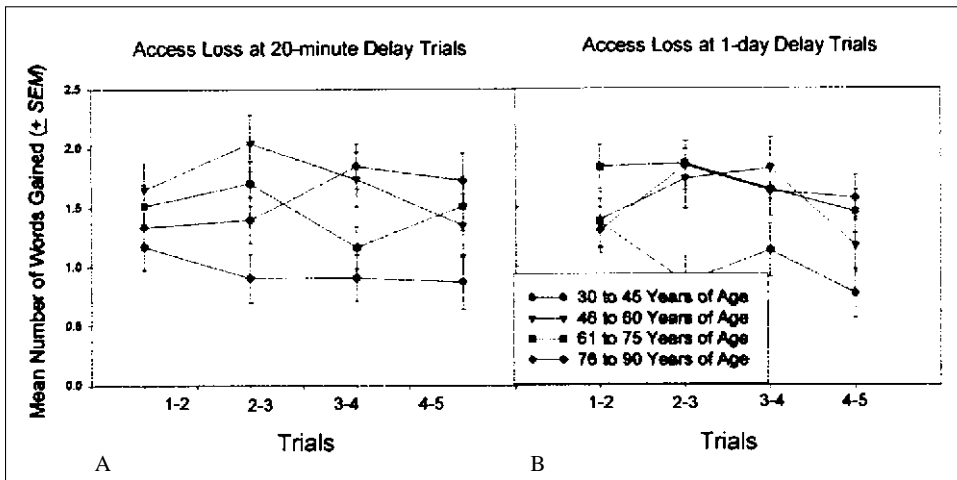


Fig. 3a-b – The panels show the mean number of lost access words at the 20-minute and 1-day delay trials for the different age groups at each intertrial increment. The lost access score was defined as the number of words recalled on trial n that were not recalled on trial $n + 1$. This component of learning is thought to reflect forgetting.

Post-hoc comparisons ($\alpha < .01$) for lost access showed no effect of age on Trials 1-2. On Trials 2-3, the 46 to 60, 61 to 75, and 76 to 90 year-old participants obtained significantly larger lost access scores than the 30 to 45 year-olds, $p < .0005$, $p < .0005$, and $p = .001$, respectively. That is, significantly more forgetting occurred between Trials 2-3 for the three oldest age groups than for the 30 to 45 year-olds. For Trials 3-4, the 46 to 60 and 76 to 90 year-old participants obtained significantly larger lost access scores than the 30 to 45 year-olds, $ps = .001$. Lastly, between Trials 4-5 it was found that the 61 to 75 and 76 to 90 year-olds obtained significantly larger lost access scores than the 30 to 45 year-olds, $p = .005$ and $p = .001$, respectively. Thus, greater lost access to study words occurs, on average, in persons over 45 years of age. Like Dunlosky and Salthouse (1996), our results show an increase in lost access scores with increasing age.

Gained and Lost Access Scores Across Delay Trials

As per the previous analyses on gained and lost access scores, it is shown here that differences exist between Trials 5 and 6 across age groups and delays. A three-way ANOVA was run with gained and lost access scores as the dependent measures, age condition as the between subjects factor, and delay (20-minute and 1-day) and gain-loss comparisons as within subjects factors. No interaction effects were found; however significant main effects were detected for age condition, $F(3, 119) = 8.52$, $MSE = .881$, $p < .0005$; for delay, $F(1, 119) = 56.63$, $MSE = 1.47$, $p < .0005$; and for gained vs. lost access, $F(1, 119) = 187.05$, $MSE = 2.86$, $p < .0005$. Significant two-way interactions were found for gained-lost access by delay, $F(1, 119) = 110.68$, $MSE = 1.53$, $p < .0005$; and gained-lost access by age condition, $F(3, 119) = 7.82$, $MSE = 2.86$, $p < .0005$.

Post-hoc comparisons ($\alpha < .01$) showed that gained access scores were significantly smaller than lost access scores, for both the 20-minute and 1-day delays ($p < .0005$ in all cases). That is, greater forgetting of individual words occurred from Trial 5 to 6 than did recall of words on the delay trial that were not recalled on Trial 5. Comparison of mean scores reveals that for the 20-minute delay, there are approximately twice the number of lost access words as gained access words; but for the 1-day delay, there are over six times as many lost access words as gained access words. This pattern holds within each age group, with larger comparative lost access scores and proportionately smaller gained-lost access ratios at the older age groups. This clearly indicates that post-acquisition forgetting is due to lost access.

The recall of words in the delay trial that were not recalled in Trial 5 (gained access) was significantly larger at the 20-minute delay ($M = 1.01$, $SE = .09$) than at the 1-day delay ($M = .66$, $SE = .08$), $LSD = .357$, $p = .001$. However, this reflects a mean difference of about half a word (largest difference for any individual was 2 words) across delay periods. The magnitude of the difference was not substantial. In contrast, lost access was significantly larger at the 1-day delay ($M = 3.96$, $SE = .20$) than the 20-minute delay scores ($M = 1.94$, $SE = .16$), $LSD = 2.03$, $p < .0005$. In the case of lost access scores, the mean change was two words and for individuals the largest change between 20-minute and 1-day delays was eight words.

No significant differences were found between age groups on gained access scores across the delays, but the following differences were detected for lost access scores. The 76 to 90 year-old participants obtained significantly higher lost access scores than did the 61 to 75 year-olds ($p = .008$), the 46 to 60 year-olds ($p = .003$), and the 30 to 45 year-olds ($p < .0005$). Further, the 61 to 75 year-olds obtained significantly higher lost access scores than the 30 to 45 year-olds ($p = .006$). More individual words were lost between Trial 5 and the delay for older participants than for younger. Examination of mean scores confirms that the increase in loss is linear from one age group to the next.

The overall results indicate that gained access words at the delay that were not recalled on Trail 5 do not substantially vary across age. Greater lost access is noted after a 1-day delay compared to after a 20-minute delay, and steadily increasing lost access to words at the delays occurs with increasing age. Greater forgetting by the elderly participants over the delay trials is due to lost access rather than any age-related change in gained access.

Finally, Pearson product-moment correlations between age, gained access, and lost access scores were examined. The correlations between gained access scores and age for acquisition trials ranged from $-.18$ to $.32$ with a total of four significant correlations. For lost access, the range of correlations was from $-.01$ to $.27$ with a total of four significant correlations. For the acquisition trials the correlations are relatively small and although a number of correlations proved to be significant, age rarely accounted for more than ten percent of the variability in scores. For the delay trials, there was no significant correlation between gained access and age. There were significant correlations between age and access lost at the 20-minute delay ($r = .33$, $p = .01$, 2-tailed) and at the 1-day delay ($r = .42$, $p = .01$, 2-tailed). Thus, the performance on the delay trials appears to be mediated by lost access with little or no contribution from gained access.

Primacy and Recency Contributions to Acquisition

Primacy and recency scores were calculated by adding the number of words recalled in the first three positions (primacy), and the last three serial positions (recency) in each study list. Words recalled from serial positions 4 through 12, inclusive, were totaled to form the recall score for the middle portion of the list. In order to compare the three sets of scores, each was divided by the total number of words possible for recall in that set; i.e., primacy and recency scores were divided by three, and the middle scores were divided by nine. This transformation produces proportion or percent correct scores for primacy, recency, and middle scores, which are shown in Figure 4. As word presentation order was different for each trial, analyses were conducted for Trial 1 only.

Primacy and recency effects were assessed using a three-way factorial design, with delay and serial position (primacy, middle, or recency) as within subjects factors, and age group as the between factor ($\alpha < .05$). Correcting for violation of sphericity (using Greenhouse-Geisser epsilon) significant main effects were found for age, $F(3, 119) = 18.48$, $MSE = .010$, $p < .0005$; and for serial position, $F(1.74, 206.86) = 172.90$, $MSE = .075$, $p < .0005$. No other significant effects were obtained.

Follow-up tests were conducted on the main effect of serial position ($\alpha < .01$). Pairwise comparisons of the main effects of age are less useful in the context of serial position as such comparisons must be collapsed across the delay and serial position variables; the results are the same as those found for the main effect of age on recall and therefore are omitted. Comparisons of proportional serial position scores collapsed across age and delay shows that the mean primacy score ($M = .79$, $SE = .02$) is significantly larger than both the recency ($M = .58$, $SE = .02$) and middle ($M = .35$, $SE = .01$) scores, $LSD = .435$ and $.202$, $ps < .0005$. Recency was likewise found to be significantly larger

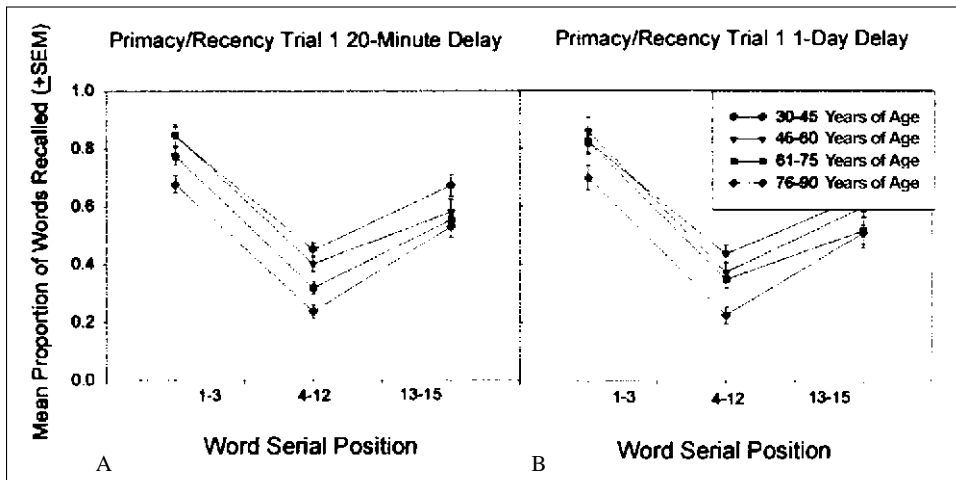


Fig. 4a-b – Primacy, middle, and recency scores (proportion of possible words recalled) at a 20-minute and a 1-day delay.

than the mean middle score, $LSD = .233$, $p < .0005$. Finally, given that no interaction effects were found, serial position effects appear to be equivalent across age groups.

Subjective Organization

For the analyses of subjective organization, two computational methods were used – the unidirectional subjective organization (SO) method, and the bidirectional intertrial repetition, or paired frequency (PF) method – both cited in Sternberg and Tulving (1977). Although the PF method is more widely recommended, the SO method is also in general use. The SO method produces a single score that represents subjective organization across the first five learning trials. The PF method produces four subjective organization scores, one for each pair of trials (i.e., Trials 1-2, Trials 2-3, Trials 3-4, and Trials 4-5). Thus, by using SO and PF methods, it is possible to examine subjective organization collapsed across all trials and its development over trials. Higher scores for both methods indicate consistent word pairings across trials. The SO scores and PF scores are shown in Figures 5 and 6, respectively.

Age correlated moderately with subjective organization. SO correlated negatively with age, $r = -.28$, and was significant ($p = .003$) but only accounts for about eight percent of the variance. Later trial pairs for PF had low to moderate correlations: Trials 3-4 correlated negatively with age, $r = -.43$, as did Trials 4-5, $r = -.38$ ($ps < .0005$), accounting for 16 to 18% of the variance. As will be shown subsequently, PF scores tend to increase across trial pairs, peaking at Trials 3-4, and the youngest age group demonstrates greater subjective organization than the other age groups. Taken together with these correlations, it appears that although subjective organization of verbal material increases with multiple presentations for all age groups, the 30 to 45 year-olds

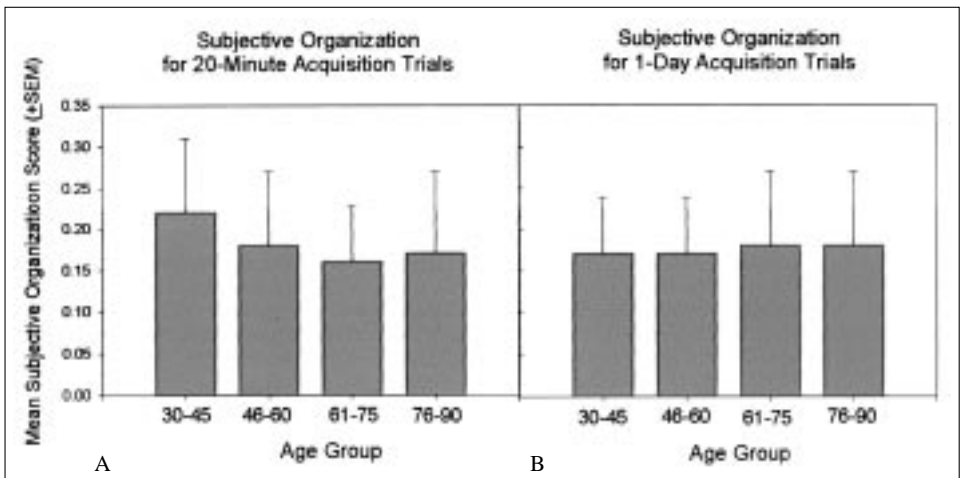


Fig. 5a-b – The panels show the mean unidirectional subjective organization score for the 20-minute delay and 1-day delay acquisition trials. The scores were derived using the formula reported by Tulving (1962).

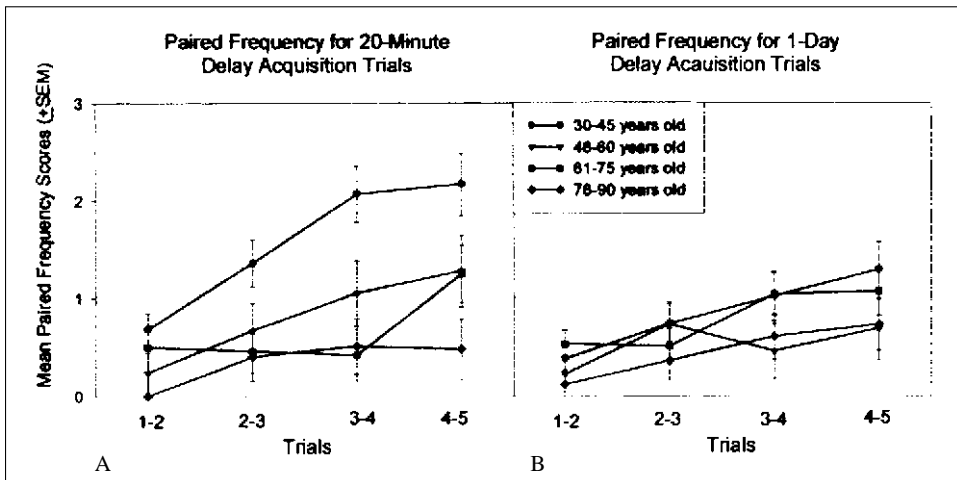


Fig. 6a-b – The panels show the paired frequency scores for the 20-minute delay and 1-day delay acquisition trials. The scores were calculated using the formula reported in Sternberg and Tulving (1977).

make larger increases in subjective organization, and in earlier trials, than is seen with all older age groups.

PF scores correlate positively with number of words recalled, $r_s = .22$ to $.51$, with stronger correlations occurring in the later trials. SO does not correlate as highly with words recalled, with correlations ranging from $.22$ to $.30$.

SO method. A repeated measures ANOVA was run with delay as the within subjects factor, and age group as the between subjects factor, using SO mean scores as the dependent variable. No main effects were found, but an interaction effect was detected, $F(3, 119) = 2.87$, $MSE = .006$, $p = .04$. Although follow-up tests did not reach significance at the $.01$ level, pairwise comparisons revealed that for the 20-minute delay condition, participants in the 30 to 45 year-old age group ($M = .22$, $SE = .016$) obtained higher SO scores than did the participants of the 61 to 75 ($M = .16$, $SE = .015$) and the 76 to 90 year-old ($M = .17$, $SE = .015$) age groups, $p_s = .022$. The magnitude of the effect was also not large, but there is a minor trend seen with the SO scores that indicates more subjective organization occurs in the younger participants.

PF method. As multiple scores are provided across trials by the PF method, greater insight is permitted into the process of subjective organization over time. A three-way ANOVA was run with delay and trial pairs as within subjects factors, and age group as the between subjects factor. A main effect of trials was detected, $F(2.25, 300.28) = 22.57$, $MSE = 1.42$, $p < .0005$, as was a delay by age group interaction, $F(3, 119) = 3.99$, $MSE = 2.97$, $p = .01$.

The correlations between PF subjective organization scores and words recalled at different trials reveals a pattern of increased overlapping variance in later trials, and especially for younger participants. Coupled with the ANOVA results showing trial by trial increases in subjective organization, and the age differences noted, a picture begins to emerge of abilities that relate learning processes across time with aging effects. From the data it appears that younger

participants learn verbal material faster, the learning process is aided substantially by organizational skills or abilities, and the ability to organize verbal material improves with repeated exposure.

Free Recall Delay Trials

A repeated measures ANOVA was performed on the number of words recalled on the two different delay trials with age group as the between independent variable and delay the within variable. The mean number of words recalled on delay trials by each group is shown in Figure 7.

Significant main effects of age, $F(3, 132) = 18.0$, $MSE = 15.4$, $p = .001$, and delay were detected, $F(1, 132) = 143.5$, $MSE = 5.79$, $p = .001$. No interaction between age and delay was detected, $F(3, 132) = 2.1$, $MSE = 2.77$, $p = .10$. A test of age effect (collapsed across delay because no interaction was detected), with an $\alpha < .01$, showed that the 75-90 year-old group recalled significantly fewer words than all younger groups. The 61-75 year-old age group recalled significantly

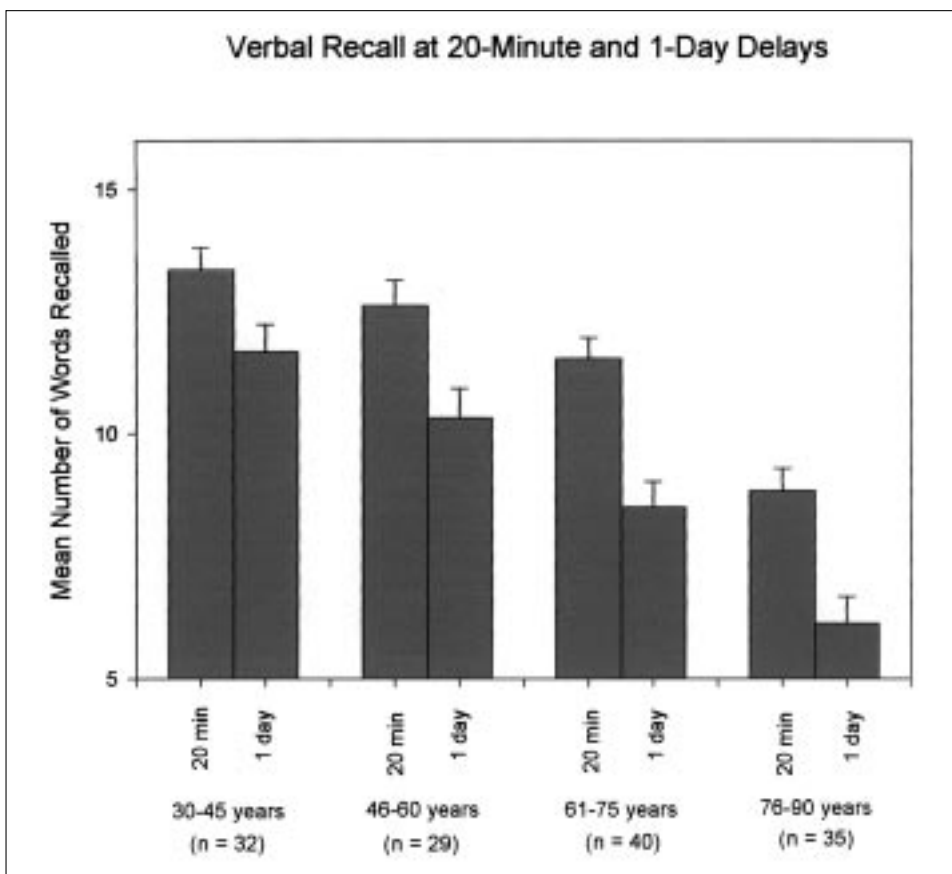


Fig. 7 – The mean number of words recalled by the different age groups at the 20-minute and 1-day delay.

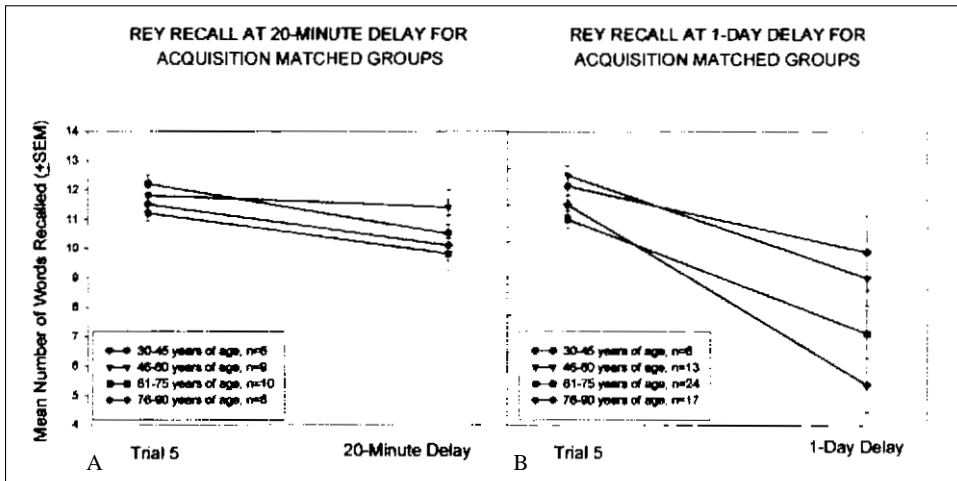


Fig. 8a-8b – The mean number of words recalled over 20-minute delay and 1-day delay by the different age groups when a subset of participants were matched on level and rate of acquisition.

fewer words than the youngest age group. Thus, it appears that even by 20 minutes the elderly are already forgetting significantly more than the young. Because of this apparent early decline in recall, we also examined forgetting by comparing difference scores (Trial 5 minus Trial 6) at the two delays.

Free Recall Forgetting Scores

An ANOVA was performed on forgetting with the difference score (Trial 5 recall minus Trial 6 recall) as the dependent variable with age and delay as the independent variables. Significant effects of age, $F(3, 132) = 10.4$, $MSE = 5.4$, $p = .001$, and delay, $F(1, 132) = 122.7$, $MSE = 2.9$, $p = .001$, were detected, and there was no interaction between age and delay. An examination of the main effect of age with $\alpha < .01$ showed that the oldest group had a larger decrease in words recalled than all younger groups.

Recognition

The mean number of words correctly recognized (from 15 target words and 15 distractor words) at the 20-minute and 1-day delays are shown in Figure 9. A repeated measures ANOVA was performed on the number of words correctly recognized with age the between variable and delay the within variable. Significant main effects of age, $F(3, 132) = 17.1$, $MSE = 6.5$, $p = .001$, and delay were detected, $F(1, 132) = 73.7$, $MSE = 2.3$, $p = .001$. Additionally, a significant interaction between age and delay was obtained, $F(3, 132) = 8.6$, $MSE = 2.3$, $p = .003$. Pairwise post-hoc comparisons of recognition performance by age condition at the different delays showed significantly poorer recognition performance at the 20-minute and 1-day delay by the oldest group relative to all other age groups. The 61 to 75 year-old participants performed significantly

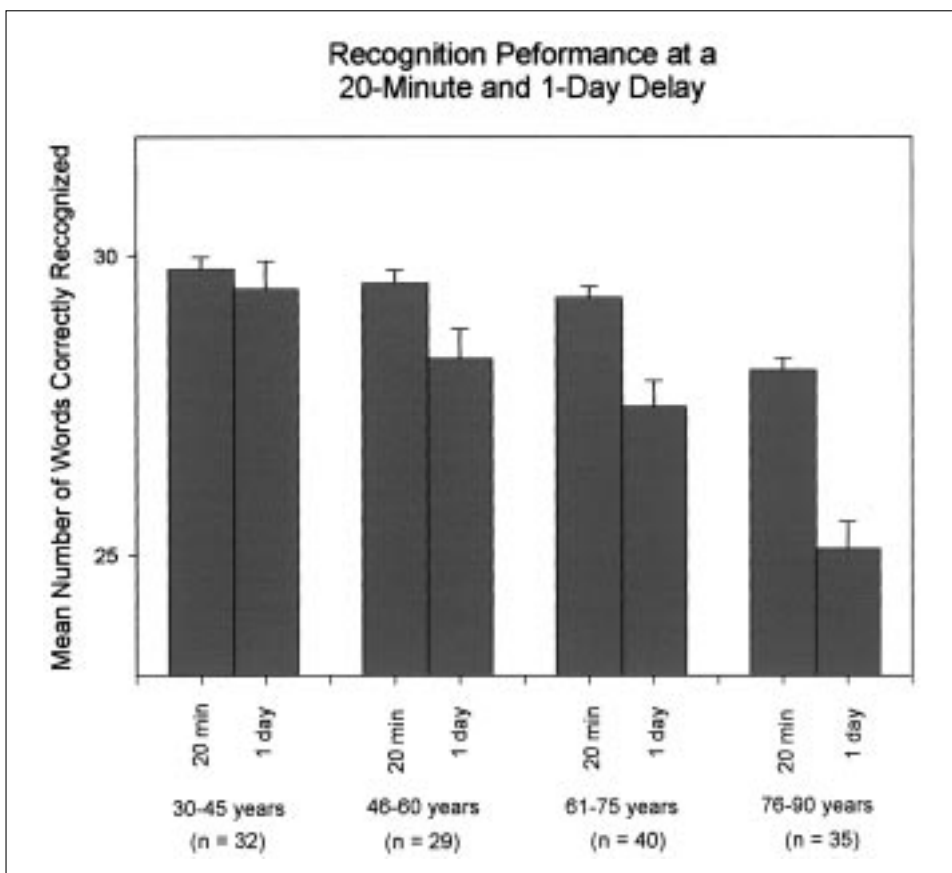


Fig. 9 – Recognition performance is shown for the different age groups at the 20-minute and 1-day delays.

worse than the youngest group at 1 day delay, but not at 20 minute delay. Pairwise comparisons of recognition by delay for the different age conditions showed that the three oldest groups significantly declined from 20-minutes to 1-day (46-60, $p = .002$; 61-75 and 76-90, $ps = .001$). These findings provide support for the view of more rapid forgetting in the elderly.

Free Recall with Acquisition Matched Comparisons

The age by delay interaction ($p = .10$) for recall on the delay trials provides only weak support for more rapid forgetting by the elderly. It is possible that different levels or rates of acquisition across age could lead to the different degrees of forgetting. Specifically, the findings in this study may reflect a classical problem of interpretation that has been reported for amnesic patients and normal elderly. That is, perhaps differential forgetting simply reflects different levels of acquisition. If you reach an acquisition level of 10 words recalled, you may simply forget more than someone who reaches an acquisition

level of 13 words recalled (as occurs in the young group by Trial 3 in the present study with the opportunity for over learning).

A subset of participants was matched on level and rate of acquisition for the 20-minute delay and a second subset was matched for the 1-day delay. Subjects who scored between 5 and 8 words recalled on Trial 1 and less than 14 words on Trials 2-5 were selected for analysis. When ANOVAs were performed on words recalled on acquisition trials at each delay, these parameters resulted in a significant trial effect at both delays [20-min, $F(4, 116) = 108, p < .001$; 1-day, $F(4, 224) = 82.4, p < .001$], no significant effect of age condition [20-min, $F(3, 29) = 1.7, p = .19$; 1-day, $F(3, 56) = 1.7, p = .18$], and no significant age by trial effect [20-min, $F(12, 116) < 1.0, p = .84$; 1-day, $F(12, 224) = 1.3, p = .22$]. Thus, subjects were matched on both the level and rate of acquisition.

Repeated measure ANOVAs were then performed on the number of words recalled at the 20-minute and 1-day delay with trials (the last immediate trial and delay trial) as the within factor and age as the between factor. The mean number of words recalled on the last immediate trial and delay trial for these fully acquisition matched groups at the two delays is shown in Figures 8a and b.

For the 20-minute delay, there was a significant effect of delay trial, $F(1, 29) = 10.3, MSE = 22.7, p = .003$. There was no significant effect of age or interaction between age and delay. At the 1-day delay, there were significant effects of delay, $F(1, 56) = 81.9, MSE = 22.8, p = .001$, and a significant age and delay trial interaction, $F(3, 56) = 3.6, MSE = 10.1, p = .02$. There was no significant effect of age. Pairwise post-hoc comparisons for the different age groups were carried out across the delay trial. The oldest group recalled significantly fewer words than the 30-45 year-old group ($p = .01$), the 46-60 year-old group ($p = .01$), and the 61-75 year-old group ($p = .05$). These findings are consistent with the view that the elderly forget more rapidly than the young, and that more rapid forgetting cannot be accounted for by level or rate of acquisition. However, it must be kept in mind that there was no significant interaction on recall or forgetting when all subjects were included in the analyses. There was greater forgetting on the recognition trials by the elderly at 1 day than at 20 minutes. Thus, the most parsimonious view of these results, as they relate to rates of forgetting, is that they provide qualified support for more rapid forgetting of verbal material in the elderly.

Forgetting Scores with Acquisition Matched Comparisons

These same subsets of participants describe above were selected for further analysis. Separate oneway ANOVAs were performed on the difference scores (Trial 5 minus Trial 6) for words recalled at the 20-minute delay and the 1-day delay trials with age condition as the independent variable. No significant effect of age was detected at the 20-minute delay, $F(3, 29) < 1.0, p = .59$. A significant effect of age was detected at the 1-day delay, $F(3, 56) = 3.6, MSE = 5.6, p = .019$. A post-hoc analysis with the LSD test, $\alpha < .01$, indicated significantly greater forgetting at the 1-day delay by the oldest age group as compared to all other age groups. This finding provides further support for more rapid forgetting by the 76 to 90 year-old participants.

Recognition with Acquisition Matched Comparisons

The same subset of participants was selected for further recognition score analysis. Separate one way ANOVAs were performed on Trial 7 recognition scores at the 20-minute delay and the 1-day delay trials with age as the independent variable. No significant effects of age were detected at the 20-minute delay, $F(3, 29) = 2.0$, $MSE = 1.5$, $p = .14$, or at the 1-day delay, $F(3, 56) = 2.6$, $MSE = 19.3$, $p = .06$. The absolute decline in recognition scores was greater at the 1-day delay, but there was much less variance in the scores of the three younger subjects at the 20-minute delay. The trend for an effect of age at the 1-day delay on recognition scores provides further, but qualified support for more rapid forgetting by the elderly.

Proactive Interference

Following the recognition test, participant recall was tested on a single trial for a new list of words. Proactive interference was assessed by performing a repeated measures ANOVA on the number of words recalled on Trial 1 minus the number of words recalled on Trial 8 with age and delay as independent variables. There were no significant main effects of age, $F(3, 132) < 1.0$, $MSE = 3.0$, $p = .51$, or delay by age, $F(3, 132) = 1.6$, $MSE = 8.4$, $p = .18$. Thus, no effect of proactive interference on different age groups was detected in this study.

Variance and Distribution Shape

It is clear from observation of the distributions of recall scores that not only do certain age groups differ on average scores, but they also differ on variability in performance and distribution properties of scores (see Figure 10). Although it is common to simply compare mean differences to delineate functional performance, a discussion of fine gradations in performance across the life span should necessarily include comparisons of variability and distribution shape. The tendencies of individuals to score above or below certain performance markers, and to score consistently or variably compared with others within an age group, are issues well worth exploring. Accordingly, variance and distribution shape were compared across the age conditions. Comparisons were made at both delays for Trials 1, 5, and 6; Trial 1 provides examination of initial learning, Trial 5 is the final study trial, and Trial 6 is the recall task after a 20-minute or a 1-day delay period.

First, performance was examined by age group at Trials 1 and 5 with delay as the within subjects factor. Using the Wilcoxon signed-rank test for related samples, significant differences between delay conditions were found at Trial 1 and Trial 5, for the 61 to 75 year-old age group. As no other groups showed significant differences at these trials across the delays, it is concluded that there are no systematic age differences from one delay condition to the next. Accordingly, for the purpose of illustration in Figure 10, performance on Trials 1 and 5 are collapsed across delays for each age group. Thus, Figure 10 shows the distribution of recall scores for each age group at Trial 1, Trial 5, Trial 6 after a 20-minute delay, and Trial 6 after a 1-day delay.

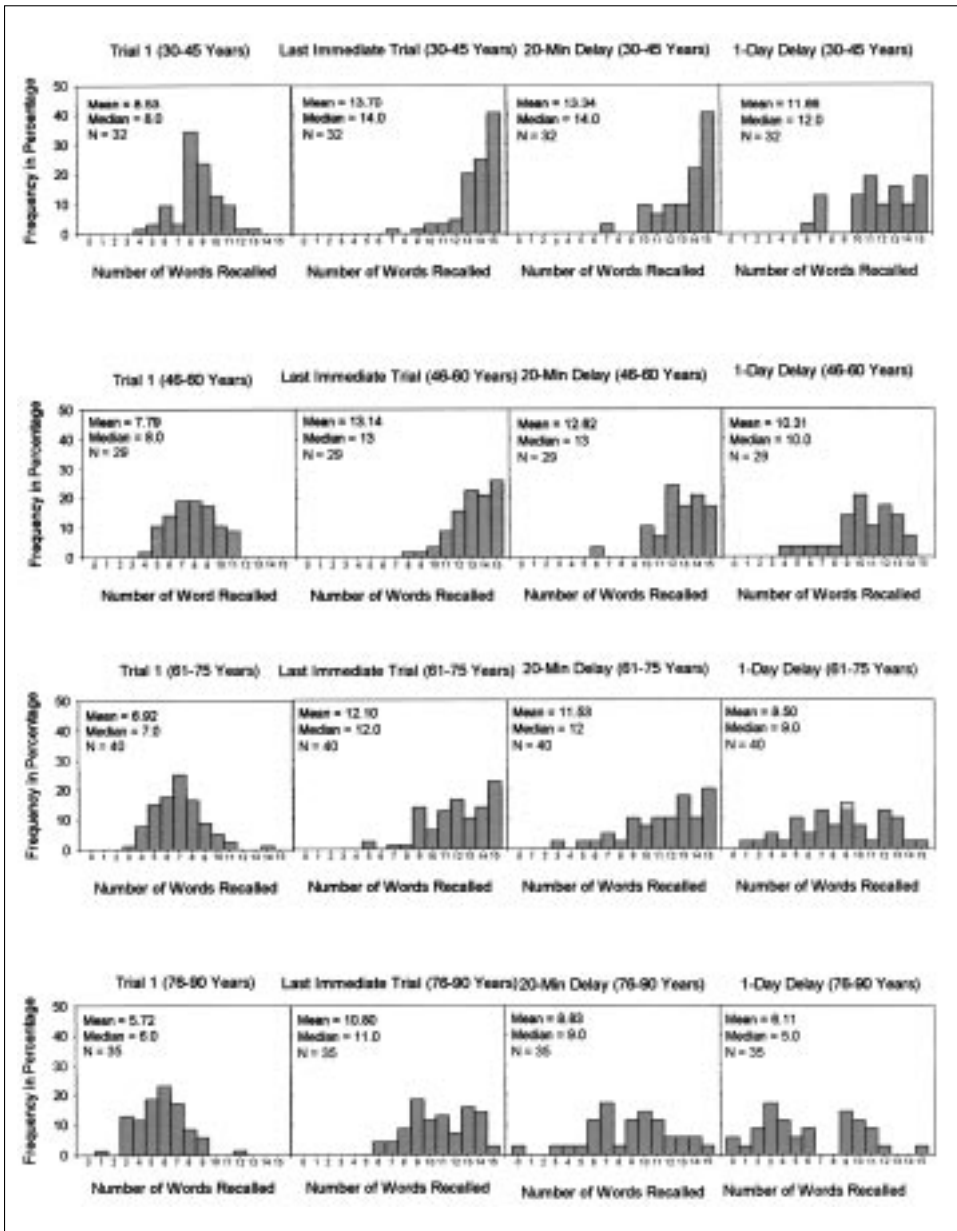


Fig. 10 – Distribution of number of words recalled given as frequency by participants in the four different age groups on the first acquisition trial (Trial 1), the last acquisition trial (Trial 5), the 20-minute delay trial (Trial 6 at 20 minutes), and the 1-day delay trial (Trial 6 at 1 day).

To compare the recall variability of each age group at each delay for Trials 1 and 5, Levene’s test of homogeneity of variance for two independent samples was used, $\alpha < .01$. On Trial 5 for the 20-min delay, the 61 to 75 year-old participants were found to have significantly larger recall variance ($SD = 2.23$)

than both the 30 to 45 (SD = 1.37) and the 46 to 60 year-olds (SD = 1.50), $t(70) = 12.90$ and $t(67) = 9.79$, $ps = .001$ and $.003$, respectively. For the same trial and delay, the variance for the 76 to 90 year-olds (SD = 2.26) was significantly larger than the variance for the 30 to 45 year-olds (SD = 1.37), $t(65) = 8.91$, $p = .004$. Likewise, analysis of Trial 5 for the 1-day delay condition showed that the 76 to 90 year-old group recall variance (SD = 2.61) was significantly higher than the variance for both the 30 to 45 (SD = 1.89) and the 46 to 60 year-olds (SD = 1.85), $t(65) = 9.30$ and $t(62) = 6.99$, $ps = .003$ and $.01$, respectively.

Age differences in variance for the recall trials (Trial 6 for the 20-minute and 1-day delays) are as follows: Recall variance for the 20-minute delay trial was significantly higher for the 76 to 90 year-olds (SD = 3.32) than for the 46 to 60 year-olds (SD = 2.01), $t(62) = 7.22$, $p = .009$; and the difference between the 76 to 90 year-olds and the 30 to 45 year-olds (SD = 2.04) was not significant, $t(65) = 6.56$, $p = .013$. Comparisons for the 1-day delay trial showed that the recall variance for the 76 to 90 year-olds (SD = 3.90) was significantly higher than the recall variance for the 30 to 45 year-olds (SD = 2.70) and the 46 to 60 year-olds (SD = 2.55), $t(65) = 8.56$ and $t(62) = 10.79$, $ps = .005$ and $.002$, respectively.

Planned comparisons for differences in recall distribution shapes were measured by the nonparametric Kolmogorov-Smirnov (K-S) test for two independent samples ($\alpha < .01$, see Siegel and Castellan, 1988). To assess differences in shape all variables were first converted to z scores, to standardize means and variance. Since the K-S test assesses differences in central tendency, variability, and distribution shape, standardizing the variables limits the K-S test to assessing differences in shape only.

For the 20-minute delay, Trial 1 recall, the distribution for the 30 to 45 year-olds is roughly normally distributed and the distributions for the 61 to 75 and 76 to 90 year-olds are significantly more positively and negatively skewed, respectively, $Ds = .675$, $ps < .0005$. The comparisons of shape for the 1-day delay at Trial 1 show that the recall distribution for the 76 to 90 year-olds is slightly more positively skewed and considerably more leptokurtic than the roughly normally distributed distributions for both the 30 to 45 and the 46 to 60 year-olds, $Ds = .667$ and $.472$, $p < .0005$ and $p = .002$, respectively.

In the 20-minute delay condition for Trial 5, the recall scores for both the 30 to 45 and the 46 to 60 year-olds are significantly more negatively skewed than the distribution of scores for the 76 to 90 year-olds, $Ds = .563$ and $.519$, respectively, $ps < .0005$. Trial 5 recall scores for the 1-day delay condition show an extreme negative skew and leptokurtosis for the 30 to 45 year-olds, and is significantly different from both the lesser negatively skewed distribution for the 61 to 75 year-olds, and the relatively normal recall distribution for the 76 to 90 year-olds, $Ds = .450$ and $.504$, $p = .001$ and $p < .0005$, respectively.

For the Trial 6, 20-minute delay trial, distributions of recall scores were significantly more negatively skewed for the 30 to 45, 46 to 60, and 61 to 75 year-olds, as compared with the platykurtic distributed recall scores of the 76 to 90 year-olds, $Ds = .613$, $.593$, and $.375$, $p < .0005$, $p < .0005$, and $p = .01$, respectively. Trial 6 recall scores for the 30 to 45 year-olds, 1-day delay, are slightly negatively skewed and significantly different from the platykurtic

distribution of scores in the 61 to 75 year-old age group, $D = .469$, $p = .01$. Similarly, recall distributions for both the 30 to 45 and the 46 to 60 year-olds at the 1-day delay. Trial 6 are both slightly negatively skewed, and are significantly different from the slight positively skewed or platykurtic distribution of scores obtained by the 76 to 90 year-old participants, $D_s = .587$ and $.497$, $p < .0005$ and $p = .001$, respectively.

The pattern of differences found by examining variability indicates that recall scores obtained by younger participants have smaller variance than do the scores of older participants. Shape differences were seen to differ by trial. At Trial 1, younger participants tended to obtain scores in a roughly normally distributed pattern, while the older age groups tended towards skewed distributions. In Trials 5 and 6, the pattern was somewhat reversed, with younger participants showing a ceiling effect and strong negatively skewed distributions, whereas the distribution of recall scores for the older participants was less skewed and more platykurtic. In addition to the mean differences discussed in earlier sections, the results of the variance and distribution shape comparisons suggest that recall becomes more variable, or perhaps less consistent, with greater age. Further consideration for these apparent differences will be given in the Discussion section.

DISCUSSION

A major purpose of this study was to evaluate the utility of a test of verbal memory and provide empirical data on when memory processes first begin to change with age, and the course of changes in later life. The test was designed to examine a variety of theoretical constructs associated with memory while allowing comparisons with results from previous studies. There are several strong points to the tests used in this study, but there are also, in retrospect, certain needed modifications. The shortcomings of our tests and needed modifications will be discussed first, and then followed by a discussion of confirmations of previous findings, divergences from previous findings, and successes of this study.

The youngest age group (30 to 45 year-olds) did not significantly increase the number of words recalled from Trial 3 onward, indicating that the younger participants may have reached asymptote sooner than participants in the three older age groups. The most likely explanation for this is that the youngest age group experienced a ceiling effect (i.e., 65 % of the youngest subjects recalled 14 or 15 words by the end of the acquisition trials). One implication of this finding is that the age difference in rate of acquisition reported in the Results section may be an artifact of the ceiling effect for the youngest age group. Another implication is the possible effects of over learning by the youngest group on retention and rates of forgetting. That is, the repeated recollection of words on Trials 4 and 5 by the youngest participants may reduce their forgetting, and thereby makes suspect any age-related differences in forgetting. While this last concern applies to many studies, it is not applicable to the present study because participants from the different age groups were matched on level and rate of acquisition.

The current findings indicate the most frequently used free recall tests in aging are susceptible to interpretation problems because, like the tests used in this study, they use approximately 15 words and young participants reach asymptote during multi-trial acquisition. Specifically, the California Verbal Learning Test (CVLT) (Delis et al., 1988; Kramer, and Delis, 1991) uses 16 words, and most of the versions of the RAVLT (Peaker, and Stewart, 1989; Wiens et al., 1988) use 15 words. This potential ceiling effect in acquisition can be eliminated by the use of a longer word list. For example, we would recommend the study list contain 24 words to eliminate the possibility of ceiling effects during acquisition, and any subsequent effects on forgetting due to over learning.

Despite this caveat, for the most part, the results for acquisition are typical and in agreement with previous studies. As indicated above, there was an age-related effect on rate of acquisition with the youngest group reaching asymptote earlier than older participants, but this finding is equivocal and may be due to a ceiling effect. Indeed, previous examination of age-related changes in the rate of acquisition for multi-trial word recall can also be described as equivocal. There are a number of studies where the young demonstrate a significantly faster rate of learning across trials (Carlesimo et al., 1998; Macht and Buschke, 1983; Query and Megran, 1983; Witte et al., 1993), and other studies report equivalent rates of learning for the young and elderly participants (Hultsch, 1974; Keitz and Gounard; 1976). For level of acquisition, a reliable and incremental age-related decrease was obtained and is consistent with previous reports (Bolla-Wilson and Bleecker, 1986; Dunlosky and Salthouse, 1996; Geffen et al., 1990; Mitrushina et al., 1991; Wiens et al., 1988). The age-related differences in level of acquisition may be underestimated in the present and other multi-trial free recall studies, again because of possible ceiling effects on the performance of younger participants. The difference in level of acquisition on a multi-trial test translates into more trials for the elderly to achieve the same level of performance as young participants. This fact is dramatically demonstrated in Figure 1. Specifically, the elderly require five trials to reach the same level of performance achieved by the youngest participants in two trials. Finally, we show that age differences in level of acquisition can be accounted for by performance on the initial learning trial.

Our results differed from those of Dunlosky and Salthouse (1996) in the contribution of "gained access" and "lost access" to the five acquisition trials. While their findings showed that both acquisition (gained access) and forgetting (lost access) contribute to learning performance, evidence from several different analyses in their study converge to indicate that acquisition played a more important role than did forgetting. The results of the current study are consistent with the view that both acquisition and forgetting contribute to learning performance. Methodological limitations in both studies make it difficult to accurately assess the contributions to acquisition. The asymptotic performance in the Dunlosky and Salthouse study and the current study can result in a low score or a score of zero by a participant on gained access. These rather minimal differences in results can be resolved by multi-trial studies using longer word list that avoid asymptotic performance by younger participants.

As was the case in the Dunlosky and Salthouse (1996), the present study shows minimal effect of gained access on delay trials. Age-related decreases for delay recall in both studies can be attributed primarily to forgetting (lost access).

When recall of primacy, recency, and middle components of the word list were examined, the results were generally found to be consistent with previous findings (e.g., Carlesimo et al., 1997; Dunlosky and Salthouse, 1996). The proportion of primacy and recency words recalled was greater than the proportion of middle words recalled. An additional consistency is the age-related decline in primacy. There are differences among these three studies that may be attributable to methodological differences. However, if the view that primacy words represent words transferred into long-term memory, then the results of the current study, as well as results from these previous studies, are consistent with an age-related decreased efficiency in transferring information into long-term memory.

It is commonly believed that aging leads to increasing susceptibility to interference. In the current study there is a decrease in recall with increasing age on the interference trial (Trial 8), but no age effects were obtained on our measure of proactive interference, providing no evidence for increased susceptibility to interference with age.

Subjective organization is thought to play a major role in word list recall, and is postulated to exert an increasingly powerful effect on recall of words across trials (Tulving, 1962, 1964). The significant increase in PF scores across trials and the significant positive correlations between PF scores and number of words recalled supports this view of subjective organization. There were also significant negative correlations between age condition and both types of subjective organization scores (SO and PF scores). Specifically, there was a significant reduction in subjective organization for all age groups as compared to the youngest subjects. Thus, the three oldest groups demonstrated less subjective organization. However, it is important not to over estimate the contribution of subjective organization to remembering. The effect size for the SO measure was relatively small, never accounting for more than 10% of the variance. The PF scores on later trials accounted for less than 20% of the variance.

Our initial examination of recall showed that participants 61 years of age and older recalled fewer words than the youngest group of participants. Similarly, the oldest participant's results for forgetting as measured by differences scores (Trial 5 minus Trial 6) was greater than for the youngest subjects. There was no apparent difference in the rate of forgetting when all participants were included in the analysis. When participants were matched on level of acquisition, the 76 to 90 year-old participants showed more rapid forgetting at the 1-day delay. It is under these circumstances of matched acquisition that the best conditions for comparing rates of forgetting for different age groups are obtained. Thus, we consider our findings under these conditions as strong support for more rapid forgetting in the elderly. Our results are also consistent with reviews on rates of forgetting that suggest no age differences at short delays (e.g., 20 minutes), but age differences at long delays (e.g., 1 day or longer) (Kausler, 1994; Salthouse, 1991). The more rapid forgetting by the oldest group when matched on acquisition with the younger groups indicates more rapid forgetting was not due

to encoding. Indeed, this more rapid forgetting is qualitatively similar to the forgetting demonstrated by amnesic patients with hippocampal damage (Squire, and Kandel, 1999), albeit the time course for forgetting is greater and the deficit is quantitatively smaller. This similarity is consistent with the idea that our behavioral data mirrors neural changes associated with aging (i.e., changes in midtemporal regions). The elderly subjects not matched on acquisition demonstrated poor acquisition, and this suggests an additional deficit superimposed on the accelerated forgetting.

The consideration of a variety of processes and variables in this study using measures of central tendency provides compelling evidence that age alters various aspects of memory. It is assumed that as variance increases with age as it did in this study, that this reflects a differential effect on memory performance in the elderly. Despite the infrequent measure of variance in the aging literature, strong support for examining variance is provided in a series of cross sectional aging studies (Morse, 1993), and in the Victoria Longitudinal Study of aging (Hultsch et al., 1998). The results from the present study also support the view that memory is subject to differential aging. The possible reasons for the increased differences in performance by the elderly include a greater time span for the expression of genetic predisposition in ability, a greater time span for experiential factors (health, educational activity, etc.) to effect performance, or possibly increased differences in performance reflect different individuals expressing different neural mechanisms associated with aging processes.

The current findings show the possibility that declines in cognitive functioning are detectable by middle age. Decreased performance by the 46-60 years old group and older groups was detected for lost access from Trials 2-3 onward (p. 16), and PF scores (see Figure 6). The detection of impairment on two measures by the 46 to 60 years old group raises the possibility of memory changes in middle age. Most studies on aging and cognition do not include a middle age group, but when such a group has been included, deficits have been reported for subjects in their 40s and 50s on the nondeclarative task of conditioned eye blink response (Durkin et al., 1993; Woodruff-Pak, 1997) and heart rate conditioning (Durkin et al., 1993), the declarative task of word recall (Davis and Keller, 1998), and on tests of frontal lobe functioning (Daigneault and Braun, 1993). This detectable change in cognitive function by middle age has also been observed in animal studies that have included what is considered a middle-aged group. For example, 11 and 17 month-old rats were impaired on place discrimination in the Morris water maze (Frick et al., 1995), and monkeys in their late teens demonstrate impaired spatial memory (Bachevalier et al., 1991). The brain changes that might underlie these deficits have not been clearly delineated. However, there are a number of brain changes detectable in middle age that might be related to changes in cognition. For example, a 6% decrease per decade in resting brain glucose utilization was reported for human subjects between 20 and 67 years of age (Petit-Taboue et al., 1998), a reduction in c-fos mRNA in rat brain following open field exposure was reported in limited areas at 14 months of age with more wide-spread reduction at 21 months (Nagahara and Handa, 1997), a decrease in protein kinase C has been reported in rats to occur as early as 8 months (Battaini et al., 1995), and for five age groups of rats

from young to old a significant correlation was observed between an age-related decline in water maze performance and forebrain cholinergic neuronal size (Fischer et al., 1991). Thus, it is suggested that as age groupings become less restricted in both human and animal studies, there will be an increase in reports of cognitive changes in middle-aged organisms.

The findings reported here did not always result in strong and clear conclusions. However, a set of reasonable conclusions emerges from our findings and can be summarized as the following:

1. Ceiling effects occurred for young participants during acquisition trials in the multi-trial free recall test, and potentially confounded measures such as forgetting, level of acquisition, susceptible to interference, etc. When multi-trial versions of the RAVLT or the CVLT have been used to compare or establish norms for different age groups the same confounds exist. Thus, a multi-trial free recall test that avoids this ceiling effect during acquisition needs to be developed and normed across the life span.

2. Both forgetting and the acquisition of words determine the form of the learning curve across trials.

3. Subjective organization makes a significant, albeit small to moderate, contribution to the acquisition of words across trials and decreases with age.

4. The elderly demonstrate more rapid forgetting when matched for level of acquisition with younger subjects.

5. Possible age-related deficits in verbal memory were detectable in middle age (i.e., fifties). We predict that more changes will be detected as studies expand the age range studied and include middle-aged subjects.

6. Greater performance differences in the elderly were indicated by greater variance on measures of recall, and difference in distribution shapes (e.g., a bimodal distribution for the oldest participants at the 1-day delay). The increased performance differences are consistent with greater variance in a single process of aging, the occurrence of multiple aging processes, or a combination of the two. For example, some elderly demonstrate normal acquisition with more rapid forgetting and others demonstrate additional impairment in acquisition. Greater variance in the elderly may reflect differential neural changes in the aging process.

7. The effect of age on recall from long-term memory is clearly and dramatically revealed in the 1-Day Delay column of Figure 10. Less than 33% of the 61-75 year old subjects recalled more words than the mean number recalled by the youngest age group. Additionally, approximately 50% of the subjects in the oldest age group recalled fewer words than the worst subject in the youngest age group.

8. The average adult from the oldest age group may be impaired relative to young adults, but their greater variability reveals a subset of approximately 10% that recall more words than the mean recall achieved by the youngest group. Similarly, over 25% of the 61-75 year old participants recalled more words than the mean of the youngest group.

Acknowledgments. This research was generously supported by funds from Memory Pharmaceuticals Corporation. We thank Larry R. Squire for experimental design suggestions and comments on an early version of this paper.

REFERENCES

- BACHEVALIER J, LANDIS LS, WALKER LC, BRICKSON M, MISHKIN M, PRICE DL and CORK LC. Aged monkeys exhibit behavioral deficits indicative of widespread cerebral dysfunction. *Neurobiology of Aging*, 12: 99-111, 1991.
- BATTAINI F, ELKABES S, BERGAMASCHI S, LADISA V, LUCCHI L, DEGRAAN PNE, SCHURMAN T, WETSEL WC, TRABUCCHI M and GOVONI S. Protein kinase C activity, translocation and conventional isoforms in aging rat brain. *Neurobiology of Aging*, 16: 137-148, 1995.
- BOLLA-WILSON K and BLEECKER ML. Influence of verbal intelligence, sex, age and education on the Rey Auditory Verbal Learning Test. *Developmental Neuropsychology*, 2: 203-211, 1986.
- CARLESIMO GA, MAURI M, GRACEFFA AMS, FADDA L, LOASSES A, LORUSSO S and CALTAGIRONE C. Memory performances in young, elderly, and very old healthy individuals versus patients with Alzheimer's disease: Evidence for discontinuity between normal and pathological aging. *Journal of Clinical and Experimental Neuropsychology*, 20: 14-29, 1998.
- CARLESIMO GA, SABBADINI M, FADDA L and CALTAGIRONE C. Word-list forgetting in young and elderly subjects: Evidence for age-related decline in transferring information from transitory to permanent memory condition. *Cortex*, 33: 155-166, 1997.
- CHEN P, RATCLIFF G, BELLE SH, CAULEY JA, DEKOSKY ST and GANGULI M. Patterns of cognitive decline in presymptomatic Alzheimer Disease. *Archives of General Psychiatry*, 58: 853-858, 2001.
- CROOK TH and YOUNGJOHN JR. Development of treatments for memory disorders: The necessary meeting of basic and everyday memory research. *Applied Cognitive Psychology*, 7: 619-630, 1993.
- DAVIS HP and KELLER FR. *Colorado Assessment Tests (version 1.0)*. Colorado Springs: Colorado Assessment Tests Co., 1998.
- DELIS DC, FREELAND J, KRAMER JH and KAPLIN E. Integrating clinical assessment with cognitive neuroscience: Construct validation of the California Verbal Learning Test. *Journal of Consulting and Clinical Psychology*, 56: 123-130, 1988.
- DAIGNEAULT S and BRAUN CMJ. Working memory and the self-ordered pointing task: Further evidence of early prefrontal decline in normal aging. *Journal of Clinical and Experimental Neuropsychology*, 15: 881-895, 1993.
- DUNLOSKY J and SALTHOUSE TA. A decomposition of age-related differences in multi-trial free recall. *Aging, Neuropsychology, and Cognition*, 3: 2-14, 1996.
- DURKIN M, PRESCOTT L, FURCHTOTT E, CANTOR J and POWELL DA. Concomitant eye-blink and heart rate classical conditioning in young, middle-aged, and elderly human subjects. *Psychology and Aging*, 8: 571-581, 1993.
- EBBINGHAUS H. *Über das Gedächtnis: Untersuchungen zur experimentellen psychologie*. Leipzig: Duncker and Humboldt, 1885.
- FERRIS SH and KLUGER A. Commentary on age-associated memory impairment, age-related cognitive decline and mild cognitive impairment. *Aging, Neuropsychology, and Cognition*, 3: 148-153, 1996.
- FISCHER W, CHEN KS, GAGE FH and BJORKLUND A. Progressive decline in spatial learning and integrity of forebrain cholinergic neurons in rats during aging. *Neurobiology of Aging*, 13: 9-23, 1991.
- FRICK KM, BAXTER MG, MARKOWSKA AL, OLTON, DS and PRICE DL. Age-related spatial reference and working memory deficits assessed in the water maze. *Neurobiology of Aging*, 16: 149-160, 1995.
- GEFFEN G, MOAR KJ, O'HANLON AP, CLARK CR and GEFFEN LB. Performance measures of 16- to 86-year-old males and females on the auditory verbal learning test. *The Clinical Neuropsychologist*, 4: 45-63, 1990.
- HULTSCH DF. Learning to learn in adulthood. *Journal of Gerontology*, 29: 302-308, 1974.
- HULTSCH DF, HERTZOG C, DIXON RA and SMALM BJ. *Memory Change in the Aged*. Cambridge: Cambridge University Press, 1998.
- KAUSLER DH. *Experimental Psychology and Human Aging*. New York: John Wiley and Sons, 1982.
- KAUSLER DH. *Learning and Memory in Normal Aging*. New York: Academic Press, 1994.
- KEITZ SM and GOUNARD BR. Age differences in adults' free recall of pictorial and word stimuli. *Educational Gerontology*, 1: 237-241, 1976.
- KRAMER JH and DELIS DC. Interference effects on the California Verbal Learning Test: A construct validation study. *Psychological Assessment*, 3: 299-302, 1991.
- LARRABEE GJ. Age-associated memory impairment: Definition and psychometric characteristics. *Aging, Neuropsychology, and Cognition*, 3: 118-131, 1996.
- LEHMAN EB and MELLINGER JC. Forgetting rates in modality memory for young, mid-life, and older women. *Psychology and Aging*, 1: 178-179, 1986.
- MACHT ML and BUSCHKE H. Age differences in cognitive effort in recall. *Journal of Gerontology*, 38: 695-700, 1983.
- MITRUSHINA M, SATZ P, CHERVINSKY A and D'ELIA L. Performance of four age groups of normal elderly on the Rey auditory-verbal learning test. *Journal of Clinical Psychology*, 47: 351-357, 1991.
- MORSE CK. Does variability increase with age? An archival study of cognitive measures. *Psychology and Aging*, 8: 156-164, 1993.

- NAGAHARA AH and HANDA RJ. Age-related changes in c-fos mRNA induction after open-field exposure in the rat brain. *Neurobiology of Aging*, 18: 45-55, 1997.
- PEAKER A and STEWART LE. Rey's auditory verbal learning test – A review. In JR Crawford and DM Parker (Eds), *Developments in Clinical and Experimental Neuropsychology*. New York: Plenum Press, 1989, pp. 219-236.
- PETIT-TABOUE MC, LANDEAU B, DESSON JF, DESGRANGES B and BARON JC. Effects of healthy aging on the regional cerebral metabolic rate of glucose assessed with statistical parametric mapping. *Neuroimage*, 7: 176-184, 1998.
- QUERY WT and MEGRAN J. Age-related norms for AVLT in a male patient population. *Journal of Clinical Psychology*, 39: 136-138, 1983.
- REY A. *L'Examen clinique en psychologie*. Paris: Press Universitaires de France, 1964.
- RICHARDS M, TOUCHON J, LEDESERT B and RICHIE K. Cognitive decline in ageing: Are AAMI and AACD distinct entities? *International Journal of Geriatric Psychiatry*, 14: 534-540, 1999.
- RYBARCZYK BD, HART RP and HARKINS SW. Age and forgetting rate with pictorial stimuli. *Psychology and Aging*, 2: 404-406, 1987.
- SALTHOUSE TA. *Theoretical Perspectives on Cognitive Aging*. Hillsdale: Lawrence Erlbaum, 1991.
- SIEGEL S and CASTELLAN NJ, Jr. *Nonparametric Statistics for the Behavioral Sciences*, (2nd Edition). New York: McGraw-Hill, 1988.
- SQUIRE LR and KANDEL ER. *Memory: From Mind to Molecules*. New York: Freeman, 1999.
- SQUIRE LR and SHIMAMURA AP. Characterizing amnesic patients for neurobehavioral study. *Behavioral Neuroscience*, 100: 866-877, 1986.
- STERNBERG RJ and TULVING E. The measurement of subjective organization in free recall. *Psychological Bulletin*, 84: 539-556, 1977.
- TULVING E. Subjective organization in free recall of unrelated words. *Psychological Review*, 69: 344-354, 1962.
- TULVING E. Intratrial and intertrial retention: Notes toward a theory of free recall verbal learning. *Psychological Review*, 71: 219-237, 1964.
- WIENS AN, MCMINN MR and CROSSEN JR. Rey auditory-verbal learning test: Development of norms for healthy young adults. *The Clinical Neuropsychologist*, 2: 67-87, 1988.
- WITTE KL, FREUND JS and BROWN-WHISTLER S. Adult age differences in free recall and category clustering. *Experimental Aging Research*, 19: 15-28, 1993.
- WOODRUFF-PAK DS. *The Neuropsychology of Aging*. Oxford: Blackwell, 1997.
- ZACHARY RA, PAULSON MJ and GORSUCH RL. Estimating WAIS-IQ from the Shipley Institute of Living Scale using continuously adjusted age norms. *Journal of Clinical Psychology*, 41: 820-831, 1985.