Correlates of Memory Decline: A 4-Year Longitudinal Study of Older Adults With Memory Complaints

Joy L. Taylor, Terry P. Miller, and Jared R. Tinklenberg
Department of Psychiatry and Behavioral Sciences, Stanford University
and Psychiatry Service, Department of Veterans Affairs Medical Center, Palo Alto, California

Change in memory performance and its correspondence to change in speed of performance and self-reported memory functioning were investigated longitudinally in 30 older adults with memory complaints. Subjects were assessed by self-report questionnaires and cognitive tests 3 times, at near 2-year intervals. A significant decline in word-recall scores was found, which was accompanied at the group level by significant self-reported decline in everyday memory functioning and nonsignificant decline in Wechsler Adult Intelligence Scale Digit Symbol scores (r = .05). The oldest subjects showed the most substantial declines in memory performance. At the individual level, however, memory change did not significantly correlate with either change in self-reports or change in Digit Symbol scores. Although these results do not support a cognitive slowing model of decline at the intrapersonal level, they do have implications for intervention of age-related memory decline.

Age-related declines in intellectual abilities have been extensively documented across the adult life span by both cross-sectional and longitudinal studies. In cross-sectional studies, age differences in levels of performance have been particularly robust for measures of secondary episodic memory (cf. Craik, 1977; Poon, 1985) and speeded tasks (cf. Cerella, 1985; Salthouse, 1985). Longitudinal findings for speeded tasks essentially agree with cross-sectional findings (Cunningham, 1987; Cunningham, White, & Smook, 1985; Schaie, 1989b). However, as there are few available longitudinal data for measures of episodic memory, the questions of the convergence of cross-sectional with longitudinal methods and the correspondence of memory decline with other intellectual decline remain unanswered.

Models of cognitive aging implicitly predict that longitudinally observed changes in memory will correlate with changes in the speed of performance. Information-processing models of aging, such as general-slowing models (Birren, 1965; Salthouse, 1985) and processing-resources models (Craik & Byrd, 1982; Craik & Simon, 1980; Hasher & Zacks, 1979), assume that aging has a broad impact on encoding, internal manipulation, and response retrieval processes. Horn's (1982) model of adult intellectual development identifies age-related decline of the ability or inclination to concentrate as a basic factor that underlies the decline of speediness, short-term memory, and fluid intelligence with age. The presumed correlation of memory and speed changes is particularly important with respect to Alzheimer's disease-related research: Whereas memory decline is seen as the hallmark of the disease, both memory and speed measures are major discriminants of normal and impairedd performance (Storandt, Botwinick, Danziger, Berg, & Hughes, 1984; Storandt & Hill, 1989; Tierney, Snow, Reid, Zor-zitto, & Fisher, 1987).

Self-perception of memory decline is another potential correlate of memory decline. Cross-sectionally based studies have, however, observed at best a weak relationship between self-reports of everyday memory functioning and actual performance of laboratory-based memory tests (e.g., O'Hara, Hinrichs, Kohout, Wallace, & Lemke, 1986; Scogin, Storandt, & Lott, 1985; see also review by Berry, West, & Dennehey, 1989). However, a number of methodological issues have been raised that cast doubt on such conclusions. Gilewski and Zelinski (1986) and Herrmann (1982) suggested that the relationship might be stronger if the self-report measures were designed more judiciously and if the cognitive tasks were isometric to the self-report items. In addition, it is conceivable that individuals differ in their usage of the self-report response scales (i.e., rater differences in scale calibration). Longitudinal studies could show a closer correspondence than cross-sectional ones if memory performance changed over time and individuals perceived changes and reported them.

Even under optimal measurement conditions, self-report measures are not likely to be strongly correlated with memory performance (Herrmann, 1982; Huhtsch, Hertzog, Dixon, & Davidson, 1988; Sunderland, Watts, Baddeley, & Harris, 1986). Self-report measures may, however, be useful for the attitudinal
information they provide about an individual's stereotypic beliefs of aging (cf. Gilewski & Zelinski, 1986) and memory self-efficacy beliefs (Berry et al., 1989; Hultsch et al., 1988), defined as beliefs about one's own capability to use memory effectively in various situations (cf. Bandura, 1986). Longitudinal data would provide information about the continuity or change of self-efficacy beliefs over time in late adulthood. For example, a cross-sectionally based factor analysis of two self-report metamemory questionnaires (Hertzog, Hultsch, & Dixon, 1989) indicated an age-related increase in the correlation between memory self-efficacy and memory change factors, where change was indexed by self-assessments of current memory functioning relative to a younger age. This leads to the prediction that, longitudinally, individuals reporting memory decline will concomitantly report decreased memory self-efficacy.

This article reports 4-year longitudinal data from a small sample of healthy older adults with subjective memory complaints. We address three questions. First, to what extent is change in verbal episodic memory correlated with change in speed of performance? Second, to what extent is change in memory performance related to changes in self-ratings of everyday memory? Third, to what extent is reported memory decline expressed as a change in memory self-efficacy beliefs?

Method

Overview

The subjects were initially participants in a short-term clinical drug trial testing the effects of pramiracetam in healthy nondemented older adults with subjective memory loss. Three times of measurement are reported here; none involve pramiracetam. The first time of measurement was the subject's initial evaluation for the drug study and took place 2 to 4 weeks before starting the 8-week treatment period. Subjects were then recruited to participate in two follow-up evaluations, which took place at approximately 2-year intervals after the first time of measurement (M = 2.3 years, SD = 0.37). (No significant drug effects were found; DeJong, Odenheimer, Hess, & Tinklenberg, 1986. The Hamilton Depression Scale ratings were well below the cutoff score of 24 out of 30 points (29 out of 30 subjects had MMSE scores of 27 or higher; M = 28.4, SD = 1.3). The Hamilton Depression Scale ratings were well below the cutoff score of 18 (M = 3.5, SD = 2.2, range = 0 to 10). Subjects' Year 0 self-ratings of their everyday memory functioning, as indexed by the Memory Functioning Questionnaire (MFQ; Gilewski, Zelinski, & Schae, 1990), were compared to the normative data of Gilewski, Zelinski, Schae, and Thompson (1983): We computed z scores for each subject for the scales of the MFQ in relation to the normative means and standard deviations. Summarized in Table 1 are the mean z scores, standard deviations, and the percentage of subjects who had self-ratings at least 1 standard deviation below the normative mean (one would expect 16% in the normative group). As shown in the table, the mean self-ratings of the present sample were generally below those of Gilewski et al.'s (1983) normative young and older age groups, indicating the present sample initially reported more memory problems than the normative population.

Table 1

Summary Statistics for Standardized Initial Memory Self-Ratings of Subjects With Subjective Memory Decline

<table>
<thead>
<tr>
<th>MFO scales × Age comparison</th>
<th>M</th>
<th>SD</th>
<th>% with z scores of ≤ -1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young adult z score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General rating</td>
<td>-0.80</td>
<td>0.79</td>
<td>31</td>
</tr>
<tr>
<td>FOF</td>
<td>-1.48</td>
<td>1.15</td>
<td>73</td>
</tr>
<tr>
<td>FOF when reading</td>
<td>-0.90</td>
<td>1.08</td>
<td>47</td>
</tr>
<tr>
<td>Seriousness</td>
<td>-0.46</td>
<td>0.93</td>
<td>30</td>
</tr>
<tr>
<td>Mnemonic usage</td>
<td>-1.05</td>
<td>0.72</td>
<td>57</td>
</tr>
<tr>
<td>Age-adjusted z score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General rating</td>
<td>-0.30</td>
<td>0.76</td>
<td>24</td>
</tr>
<tr>
<td>FOF</td>
<td>-0.88</td>
<td>0.95</td>
<td>50</td>
</tr>
<tr>
<td>FOF when reading</td>
<td>-0.23</td>
<td>0.84</td>
<td>23</td>
</tr>
<tr>
<td>Seriousness</td>
<td>-0.71</td>
<td>0.95</td>
<td>43</td>
</tr>
</tbody>
</table>

Note. The young adult z scores are referenced to Gilewski, Zelinski, Schae, and Thompson's (1983) young group, where the mean age was 35 years; the age-adjusted z scores are referenced to the Gilewski et al. (1983) age group that corresponded to the subject's age at entry. An age-adjusted z score was not computed for the Mnemonics Usage scale because age differences have not been observed for this scale. The expected percentage of z scores of ≤ 1 or less was 16%. MFQ = Memory Functioning Questionnaire; FOF = Frequency of Forgetting Scale.

Subjects

Fifty-three community-dwelling older adults with subjective memory decline participated in the Year 0 evaluation. To be accepted into the study, subjects had to describe their memory changes as having a gradual onset with a duration of at least 6 months. Subjects were selected to be between the ages of 60 and 85 years, with no restriction on gender or educational level as long as they could comprehend the testing materials. They could not show obvious signs of dementia, as indicated by a score of less than 24 on the 30-point Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), a brief screening test of intellectual functioning, or have a current depressive disorder, as indicated by a depression rating greater than 18 on the Hamilton Depression Scale (Hamilton, 1960). Subjects with a history of major psychiatric illness, stroke, or any clinically significant medical condition that would be expected to contribute to cognitive impair-
**Measures**

**Cognitive measures.** Four cognitive tests were administered at all three times of measurement: three tests of the WAIS (Wechsler, 1955), specifically the Arithmetic, Digit Symbol, and Block Design tests, plus a 12-word recall task using the selective-reminding technique (Buschke, 1973; Buschke & Fuld, 1974). The same word list and WAIS materials were used at each time of measurement.

In the recall task, the 12-word list was presented for three study-test learning trials. Words were presented orally at a 2-s rate, and recall was oral. In accordance with the selective-reminding procedure, the entire list was presented on the first trial and was immediately followed by a free-recall test, in which the subject attempted to recall the 12 words in any order she or he preferred. On subsequent trials, only those words not recalled on the immediately preceding trial were presented, but subjects were instructed to try to recall all words from the list. The reminding and recall procedures were explained to subjects before the testing began. Whereas most selective-reminding tasks involve 5 to 15 learning trials, we shortened the task to 3 trials because of the overall length of the assessment procedures. The word list was constructed with the aim that the words be unrelated to one another, as opposed to being easily categorizable. The list was generated by a computer program that randomly selects words from two stimulus pools of 160 high-imagery words and 167 low-imagery words developed by Kraemer, Peabody, Tinklenberg, and Yesavage (1983) from the word norms of Toglia and Battig (1978). The 6 high-imagery words randomly selected had a mean normative imagery rating of 5.86 (range = 5.69 to 6.05, scale = 1 to 7). The 6 low-imagery words had a mean imagery rating of 3.17 (range = 2.67 to 3.38). The serial positions of high- and low-imagery words within the list were balanced by alternating them.

The scores used for data analyses were the raw (unscaled) scores of the WAIS subtests and the number of words recalled on each trial (Trial 1, 2, or 3) of the recall task.

**Self-ratings of memory functioning.** The MFQ (Gilewski et al., 1990) was used to obtain self-evaluations of everyday memory functioning. This questionnaire, developed for use in older adult populations, has 64 items, reducible to four correlated factors. The most salient factor taps frequency of forgetting or memory self-efficacy (cf. Hertzog et al., 1989). The scores used for data analyses were means for the scales, except for the Retrospective Functioning scale. To facilitate a longitudinal analysis of change in self-rated memory decline, we disaggregated the Retrospective Functioning scale and used the first item, “How is your memory compared to the way it was 1 year ago?” with the aim of quantifying subjects' reports of memory change during the time period of the study.

**Clinical measures.** The clinical instruments were the MMSE (Folstein et al., 1975), a brief dementia-screening examination, and the Brief Cognitive Rating Scale (BCRS; Reisberg, Schneck, Ferris, Schwartz, & de Leon, 1983). The BCRS consists of 5 clinician-rated items: Concentration, Recent Memory, Past Memory, Orientation, and Functioning and Self-Care. Each item is rated during a structured interview on a 7-point scale indicating no, subjective, mild, moderate, moderately severe, severe, or very severe cognitive impairment. The MMSE total score and the BCRS mean rating were used for data analyses.

**Procedure**

Subjects were tested individually. At each time of measurement, subjects completed the MFQ before other measures were obtained. At Year 0, subjects were interviewed by a clinician, who administered the word-recall task and full WAIS, except the Information subtest, and then were screened for certain medical abnormalities by physical examination and laboratory tests. The Year 0 evaluation took place during two half-day sessions scheduled 2 weeks apart. At Years 2 and 4, subjects were reinterviewed concerning their memory and medical status, and they were questioned about depressive symptoms. They were then readministered the word-recall task and the three WAIS subtests, which were part of a larger cognitive task battery. The Year 2 and 4 evaluations lasted 3 to 3.5 hr.

**Results**

Table 2 shows the means and standard deviations for the cognitive, self-reported memory functioning, and clinical measures at the three times of measurement for the residual sample of 30 subjects. For all measures, except the BCRS, higher scores indicate better performance or more positive ratings.

**Plan of Analysis**

**Data reduction.** Because the sample size \( n = 30 \) was modest relative to the number of dependent variables (14), we sought to identify a more limited set of variables that would span the domains tested and be relatively independent. Variable selection was aided by principal-components analysis (PCA) with varimax rotation. The data used for the PCA were the Year 0 scores for the variables listed in Table 2 for all subjects tested at that time \( (n = 43) \). The input correlation matrix, based on Spearman's rank correlation coefficient computation, is shown in Table 3. The PCA suggested a four-factor solution, which accounted for 60% of the variance.

Loading strongly on the first factor were three of the MFQ scales, Frequency of Forgetting (FOF), FOF when reading, and Seriousness of Forgetting when it occurs, and the BCRS mean rating. Because the MFQ–FOF scale typically exhibits good psychometric properties (Gilewski et al., 1983) and has been identified as a marker of memory self-efficacy (Hertzog et al., 1989), the MFQ–FOF scale was selected to represent Factor 1. The Retrospective Functioning item and Mnemonics scale were somewhat independent of the FOF scales, which is consistent with previous analyses of the MFQ (Gilewski et al., 1983; Hertzog et al., 1989).

Loading most strongly on the second factor were the WAIS Digit Symbol and Block Design tests and word recall Trial 1 variables. Because of our interest in speed of performance and because the Digit Symbol test generally has good test-retest reliability (Wechsler, 1955), we selected the Digit Symbol score to represent Factor 2.

The word recall Trial 2 and 3 variables loaded strongly on Factor 3. Although the differential loadings for Trials 1, 2, and 3 were not predicted ad hoc, the pattern was congruent with multiple-store and multiple-process models of verbal learning and memory. We consider here the distinction between primary and secondary episodic memory. Crowder (1976, pp. 146–151) concluded in his review of the immediate free-recall literature that the estimated capacity of primary memory is two or three items, with the capacity estimate holding constant across subject characteristics and experimental manipulations. In examining the present results, where the number of words recalled increased monotonically across successive trials (mean recall of 4.60, 7.27, and 8.07 words, on Trials 1, 2, and 3, respectively), it
Table 2
Means and Standard Deviations for Cognitive, Self-Reported Memory, and Clinical Measures at Three Times of Measurement

<table>
<thead>
<tr>
<th>Measure</th>
<th>Year 0</th>
<th>Year 2</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Word recall (range = 0–12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trial 1</td>
<td>4.60</td>
<td>1.30</td>
<td>4.80</td>
</tr>
<tr>
<td>Trial 2</td>
<td>7.27</td>
<td>1.60</td>
<td>7.23</td>
</tr>
<tr>
<td>Trial 3</td>
<td>8.07</td>
<td>1.48</td>
<td>8.17</td>
</tr>
<tr>
<td>WAIS raw subscores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arithmetic (range = 0–18)</td>
<td>13.40</td>
<td>2.79</td>
<td>13.20</td>
</tr>
<tr>
<td>Digit Symbol (range = 0–90)</td>
<td>48.30</td>
<td>10.89</td>
<td>48.53</td>
</tr>
<tr>
<td>Block design (range = 0–48)</td>
<td>35.70</td>
<td>6.68</td>
<td>35.03</td>
</tr>
<tr>
<td>MFQ (range = 1–7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General rating</td>
<td>3.84</td>
<td>0.94</td>
<td>3.82</td>
</tr>
<tr>
<td>Retrospective functioning*</td>
<td>3.52</td>
<td>0.75</td>
<td>3.53</td>
</tr>
<tr>
<td>FOF</td>
<td>3.83</td>
<td>0.86</td>
<td>4.14</td>
</tr>
<tr>
<td>FOF when reading</td>
<td>4.72</td>
<td>0.97</td>
<td>4.86</td>
</tr>
<tr>
<td>seriousness</td>
<td>3.98</td>
<td>1.14</td>
<td>4.04</td>
</tr>
<tr>
<td>Mnemonics usage</td>
<td>2.62</td>
<td>0.87</td>
<td>2.78</td>
</tr>
<tr>
<td>Clinical ratings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCRS mean rating (range = 1–7)</td>
<td>1.73</td>
<td>0.28</td>
<td>1.72</td>
</tr>
<tr>
<td>MMSE total score (range = 0–30)</td>
<td>28.43</td>
<td>1.28</td>
<td>28.99</td>
</tr>
</tbody>
</table>

Note. For all measures except BCRS, higher scores indicate better performance or more positive ratings. WAIS = Wechsler Adult Intelligence Scale; MFQ = Memory Functioning Questionnaire; FOF = Frequency of Forgetting Scale; BCRS = Brief Cognitive Rating Scale; MMSE = Mini-Mental State Examination.

* The rating is for MFQ Item 2a. For this item, subjects were asked, "How is your memory compared to the way it was one year ago?"

It is logical to infer that recall performance on successive learning trials reflected increasingly larger contributions from secondary memory. Tentatively, we conclude that Trial 2 and 3 scores were markers of secondary memory ability. These two scores were averaged to represent Factor 3. As the MFQ Retrospective Functioning item had the highest loading on the fourth factor, it was selected to represent Factor 4.

To summarize, four relatively independent measures (hereinafter referred to as primary dependent variables) were selected.

Table 3
Intercorrelations Among the Cognitive Performance, Clinical Ratings, and Memory Functioning Self-Report Measures at Year 0

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Trial 1</td>
<td></td>
<td>.49</td>
<td>.20</td>
<td>.31</td>
<td>.31</td>
<td>.21</td>
<td>.45</td>
<td>-.14</td>
<td>-.10</td>
<td>-.01</td>
<td>.04</td>
<td>-.20</td>
<td>-.19</td>
<td>.07</td>
</tr>
<tr>
<td>2. Trial 2</td>
<td></td>
<td>.57</td>
<td>.10</td>
<td>.09</td>
<td>-.00</td>
<td>.25</td>
<td>.10</td>
<td>-.25</td>
<td>-.05</td>
<td>-.03</td>
<td>-.10</td>
<td>-.05</td>
<td>-.17</td>
<td></td>
</tr>
<tr>
<td>3. Trial 3</td>
<td></td>
<td>.13</td>
<td>.12</td>
<td>.09</td>
<td>.30</td>
<td>.09</td>
<td>-.17</td>
<td>.00</td>
<td>-.02</td>
<td>.05</td>
<td>-.02</td>
<td>.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS subscale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Arithmetic</td>
<td></td>
<td>-.16</td>
<td>.32</td>
<td>.40</td>
<td>-.13</td>
<td>.22</td>
<td>.35</td>
<td>.26</td>
<td>.36</td>
<td>.16</td>
<td>.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Digit Symbol</td>
<td></td>
<td>-.32</td>
<td>.25</td>
<td>-.07</td>
<td>-.01</td>
<td>.05</td>
<td>.18</td>
<td>-.15</td>
<td>.02</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Block design</td>
<td></td>
<td>-.39</td>
<td>-.11</td>
<td>.26</td>
<td>.12</td>
<td>.22</td>
<td>.04</td>
<td>.03</td>
<td>.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. MMSE total</td>
<td></td>
<td></td>
<td>-.00</td>
<td>.01</td>
<td>-.02</td>
<td>-.00</td>
<td>-.11</td>
<td>-.01</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. BCRS mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Global</td>
<td></td>
<td>.30</td>
<td>.23</td>
<td>.40</td>
<td>.40</td>
<td>.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. FOF general</td>
<td></td>
<td>-.47</td>
<td>.49</td>
<td>.32</td>
<td>.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. FOF reading</td>
<td></td>
<td>.39</td>
<td>.10</td>
<td>.02</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Seriousness</td>
<td></td>
<td>-1.9</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Mnemonics</td>
<td></td>
<td>-.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Retrospective rating (1 year ago)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. WAIS = Wechsler Adult Intelligence Scale; MMSE = Mini-Mental State Examination; BCRS = Brief Cognitive Rating Scale; MFQ = Memory Functioning Questionnaire; FOF = Frequency of Forgetting Scale.
for hypothesis testing. They were (a) word recall, the average number of words recalled on Trials 2 and 3, as a measure of secondary episodic verbal memory, (b) WAIS Digit Symbol raw score, as a measure of speed of performance, (c) MFQ-FOF as a measure of memory self-efficacy, (d) MFQ Retrospective Functioning ("How is your memory compared with the way it was one year ago?") as a measure of self-rated memory change.

Measurement of intraindividual change. A major objective of the statistical analyses was to relate intraindividual change over time in word recall to change in Digit Symbol performance and memory self-ratings. For each subject we derived the average rate of change (ARC) over the time period for each measure and then analyzed the ARC scores with correlational procedures, as described more fully in the next section. To derive the ARC values, a subject's scores on a given variable at the three times of measurement were linearly regressed on the time (in years) elapsed from the first measurement. Because the MFQ Retrospective Functioning item ("How is your memory compared with the way it was one year ago?") is itself a measure of change (on a 7-point Likert scale, 1 indicates much worse, 4 indicates same, and 7 indicates much better), the Year 2 and 4 ratings were averaged to measure self-reported retrospective memory change during the period of the study.

We emphasize that the index of change used in this study represents the average rate of change (ARC) over the time period. As mathematically demonstrated by Kraemer and Thiemann (1989), the slope validly estimates the ARC over a time period for nonlinear as well as linear growth curves. Although the few extensive longitudinal studies indicate that some cognitive abilities decline in an accelerating manner for the age range studied here (Arenberg, 1990; Field, Schaie, & Leino, 1988; Schaie, 1989a, 1989b), we did not model the functional form of the change within individuals because there were too few observations on each subject to model deviations from linearity.

Statistical tests. (a) Two-sample t tests were used to test for the presence of subject selection biases related to attrition. (b) One-sample t tests were applied to the four ARC variables to test presence of longitudinal change. (c) Univariate correlation coefficients were computed among the ARC variables to examine the following expected relationships: (a) that change in verbal memory ability, as measured by the ARC for Word Recall (Trial 2 and 3 scores) would be positively correlated with change in speed of performance, as measured by the ARC for WAIS Digit Symbol scores; (b) that change in verbal memory would be positively correlated with change in memory self-efficacy or degree of self-reported memory change, as measured by the ARC for the MFQ-FOF scale and the MFQ Retrospective Functioning rating; and (c) that change in memory self-efficacy would be positively correlated with degree of self-reported memory change.

Effects of Attrition

Ten subjects dropped out of the longitudinal study. There were no statistically significant differences between the 10 dropped and the 30 remaining subjects for the primary dependent variables at Year 0 (all ps > .20 by t tests), nor were there differences in age, level of education, or depression ratings. Thus, there were no indications of selective attrition effects in the study.

Group Changes

The average rates of change per year for the primary dependent variables are summarized in Table 4. For all measures, negative values indicate decline over time or increasingly negative self-ratings. Listed in the table are the means, standard deviations, minimum and maximum values observed, and effect sizes—the mean divided by the standard deviation. Cohen (1977) has suggested that effect sizes expressed in such units be considered small if the absolute values are less than 0.2, moderate if they reach 0.5, and large if they are greater than 1.0.

Cognitive ability changes. For the word-recall task, there was an average rate of decline of 0.19 words recalled per year from an initial mean of 7.7 words. This decline was statistically significant, t(29) = -3.43 (p < .01, two-tailed test), and represented an average loss of 10% from the mean initial level of performance over the 4-year period of the study. The Digit Symbol results indicate an average rate of decline of 0.6 items per year from an initial mean score of 48 items completed and an average loss of 5%. The decline in Digit Symbol scores was not statistically significant, t(29) = -2.00, p = .06.

Changes in self-reported memory functioning. With respect to the MFQ Retrospective Functioning self-rating ("How is your memory compared with the way it was one year ago?"), subjects rated their memory functioning as declining during the period of the study. The mean response for Years 2 and 4 was 3.53, which was significantly lower than 4, the same response for this rating, t(29) = -3.56, p < .01. The effect size was of the same order of magnitude as that observed for decline in word-recall performance. In contrast with subjects' retrospective reports of memory decline, there was not a significant parallel increase in FOF t was less than 1 for the FOF average change measure. Interestingly, the FOF-ARC mean was in the direction of a small decrease over time in self-estimated FOF.

Correlates of Memory Change

Contrary to our predictions, decline in word recall was not significantly correlated with decline in WAIS Digit Symbol performance (r = .04). Furthermore, decline in word recall was not significantly correlated with either self-estimated increase in FOF (r = -.09) or self-reports of memory decline during the period of the study (r = -.13). On the other hand, the two self-report measures of change in memory functioning were moderately correlated with each other as expected (r = .49, p < .01) for the MFQ-FOF ARC score with the MFQ Retrospective Functioning rating.

Because we did not detect significant correlates of the decline in word recall, we turned to indices of how much systematic variance was apparent in the ARC measures. The cross-time correlation or stability coefficients provide some information regarding the extent of individual differences in amount of change. As reviewed by Hertzog and Nesselroade (1987), stability coefficients can be high when there is consistent change across individuals, when only a few individuals in a sample
show change, or when the amount of intraindividual change is small relative to the magnitude of interindividual differences. Table 5 shows the three sets of stability coefficients, that is, for Years 0–2, Years 2–4, and Years 0–4. The correlations were fairly high for the Digit Symbol test where they were .86, .86, and .78 for Years 0–2, Years 2–4, and Years 0–4, respectively. Also listed in Table 5 are the correlations of the ARC measures with their respective Time 0 variables. These coefficients provide additional information about variation in intraindividual change, as $r^2$ indicates the extent of variation in ARC scores predictable from initial scores. In the case of the Digit Symbol and the MFQ–FOF measures, these correlations were moderate and negative, suggesting that some of the variance in these two ARC measures may be accounted for by regression to the mean. Because the initial scores on the primary dependent variables were relatively independent and the previously mentioned pattern of correlations was found, there was an upper limit for the magnitude of correlations involving change in Digit Symbol scores and MFQ–FOF self-ratings. On the other hand, the word-recall change measure appeared to be less restricted in this way as the stability coefficients were .68, .74, and .68, and the correlation of change with initial scores was low ($-.10$).

We explored age at entry into the study as a predictor of decline in word-recall scores and found that age moderately predicted ARC in word-recall scores ($r = -.41$). Age did not predict much change in WAIS Digit Symbol scores ($r = -.10$). Scattered plots revealed that all 8 subjects who entered the study while in their 70s had negative ARC scores for word recall, evincing decline in memory test scores over time. Of the 22 individuals who entered the study in their 60s, 13 (59%) showed memory decrement by this measure. (Age was a weak predictor of initial word-recall performance ($r = -.20$), whereas it was a strong predictor of initial Digit Symbol performance ($r = -.63$)).

As mentioned earlier, the two self-report measures of change in memory functioning were moderately correlated with each other ($r = .49$). Scattered plots revealed that, of the 16 subjects who rated their memory functioning as declining during the period of the study, as assessed by the MFQ Retrospective Memory Functioning item, slightly more than one half (9 subjects, or 56%) did tend to report increased estimates for FOF during this period (i.e., their ARC values were negatively signed). Of the subjects reporting no change in memory functioning, many (9 of 12 subjects, or 75%) tended to lower their estimates of FOF.

Discussion

In this longitudinal study of healthy older adults with subjective memory complaints, we observed declines in performance of a verbal secondary episodic memory test that were roughly paralleled at the group level by self-reported decline in everyday memory functioning. No significant decline was observed in speed of performance as measured by the WAIS Digit Symbol test.

Cognitive Ability Changes

There was an average loss of 10% from initial levels of performance in the word-recall task over the 4-year period of the study, with the average rate of decline being faster for the oldest subjects. As the words used in the recall task were not readily categorizable in terms of semantic categories, the verbal memory task used here may have been particularly sensitive to cognitive aging viewed as decline in self-initiated encoding and retrieval processes (Craik, 1983; Hasher & Zacks, 1979). One would expect that longitudinally measured age changes would be smaller in memory tasks where the testing procedures or
materials encourage the appropriate encoding and retrieval operations (cf. Craik, Byrd, & Swanson, 1987). Although longitudinal designs have obvious advantages because of their potential for gaining insights into developmental processes, the measurement of change in longitudinal studies is obscured by the effects of attrition and carry-over effects, particularly practice. No effects of attrition were found in the present study, but the use of the same testing materials throughout the study because of unavailability of validated equivalent forms may have allowed for practice effects. Practice effects would result in an underestimation of age decline for the age range studied here. In contrasting the observed decline in word recall to the nonsignificant change in Digit Symbol scores, we emphasize that these results may reflect a larger effect of aging on word recall than on Digit Symbol performance or, alternatively, a smaller practice effect on the word-recall task. Importantly, decline was observed in a test of verbal episodic secondary memory.

The subjects of the present study were generally highly educated, intelligent, and in good health relative to their age group. Thus, our findings may not generalize to samples including a broader segment of the elderly population. One may speculate that subjects in the present study started with higher memory ability as young adults and are declining more slowly relative to the older adult population. This extrapolation is based on cross-sectional studies of memory performance of older adults, which indicate that adults having high verbal ability and levels of social activity perform better on tests of episodic memory (e.g., Arbuckle, Gold, & Andres, 1986; Craik et al., 1987) and on longitudinal studies, which indicate that individuals who perform at high levels on tests of verbal, spatial, and inductive-reasoning tests in midlife show relatively smaller decrements in late life than less able individuals (Schaei, 1989a).

Memory Loss and Cognitive Slowing

Correlational analyses yielded little evidence that decline in memory performance within an individual was associated with decline in speed of performance. In terms of models of age changes in intellectual performance, these results do not support a single-effect model where general slowing of cognitive processing accounts for decline in intellectual performance. However, such a model cannot be rejected, because the data from this study are limited by the number and characteristics of subjects sampled as well as the period of follow-up. A more definitive longitudinal study of change in memory functioning would require larger and sequential samples of individuals having greater baseline variability in performance or self-assessments. Coupling such a sampling strategy with longer time periods between measurements should increase the amount of systematic variance available in the measures of change.

Finally, the Digit Symbol test may have a relatively large motor component, which may overshadow the measurement of attentional capacity constructs, such as effortless encoding, that are considered to be integral to episodic memorization (cf. Hasher & Zacks, 1979). Multiple or more finely tuned measures may be required to examine the correspondence between memory changes and attentional capacity changes definitively.

Changes in Self-Report

On the average, subjects did rate their memory functioning as declining during the period of the study, as indexed by sequential self-judgments of current memory functioning relative to that 1 year earlier. Memory self-efficacy, as indexed by self-estimates of frequency of memory problems in various daily situations, was not observed to change at the group level. At the individual level, average rate of self-reported memory decline was moderately correlated with longitudinal change in memory self-efficacy. Thus, subjects displayed a moderate degree of consistency in their metamemory beliefs of self-functioning. We observed, however, that the correspondence between self-perceived rate of decline and change in the estimated frequency of memory problems was not direct. For example, most of our subjects who reported no decline in memory during the study tended, over time, to lower their estimates of how often they experienced memory difficulties, as opposed to making unchanging frequency estimates.

Memory Loss and Complaints

Although subjects in the study displayed a moderate extent of internal consistency in their metamemory judgments over time, self-report of memory change was not correlated with change in performance. It is worth noting that the correlation of memory self-evaluation with performance may have been attenuated, because sampling of individuals with memory complaints is conducive to suppression of variance. In integrating these preliminary longitudinal findings with the cross-sectional ones, we agree that memory self-assessments should not be used as substitute measures of memory performance. They are useful for the attitudinal information they provide.

Implications

The most important finding was that the decline observed in word recall was at least as great as the decline observed in Digit Symbol performance. We believe that this psychometrically observed decline may impact on performance in cognitively demanding real-life situations and conclude that metamemory skill training courses and other activities designed to optimize everyday functioning should be evaluated for their effectiveness. It is worth mentioning, however, that individuals vary in how much benefit they get from mnemonic training techniques. Older adults in the upper age range have been found to show fewer gains than those in the lower range from some techniques (Yesavage, Sheikh, Friedman, & Tanke, 1990). Our study suggests that these adults in the upper age range tend to show the fastest rates of memory decline, making this an important area for further research in possible intervention.

We feel it is unlikely that these individuals are showing indications of early dementia, on the basis of the lack of concomitant decline in speed of performance in these subjects as well as a comparison of their initial and follow-up psychometric test scores to those of patients with probable Alzheimer's disease (Taylor et al., 1988). The question of whether older adults with self-perceived memory difficulties show age changes in mem-
ory abilities that are different from their noncomplaining peers remains unanswered.

References


Received March 14, 1991
Revision received June 25, 1991
Accepted July 2, 1991

---

**Instructions to Authors**

Authors should prepare manuscripts according to the *Publication Manual of the American Psychological Association* (3rd ed.). All manuscripts must include an abstract containing a maximum of 960 characters and spaces (which is approximately 120 words) typed on a separate sheet of paper. Typing instructions (all copy must be double-spaced) and instructions on preparing tables, figures, references, metrics, and abstracts appear in the *Manual*. Also, all manuscripts are subject to editing for sexist language.

Masked reviews are optional, and authors who wish masked reviews must specifically request them when submitting their manuscripts. Each copy of a manuscript to be mask reviewed should include a separate title page with authors' names and affiliations, and these should not appear anywhere else on the manuscript. Footnotes that identify the authors should be typed on a separate page. Authors should make every effort to see that the manuscript itself contains no clues to their identities.

Authors will be required to state in writing that they have complied with APA ethical standards in the treatment of their sample, human or animal, or to describe the details of treatment. (A copy of the APA Ethical Principles may be obtained from the APA Ethics Office, 750 First Street, NE, Washington, DC 20002-4242.)

APA policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. APA policy also prohibits duplicate publication, that is, publication of a manuscript that has already been published in whole or in substantial part elsewhere. Authors have an obligation to consult journal editors if there is any chance or question that the paper might not be suitable for publication in an APA journal. Authors who do not inform the Editor of the prior publication history of previously published material will be engaging in unethical behavior. Also, authors of manuscripts submitted to APA journals are expected to have available their raw data throughout the editorial review process and for at least 5 years after the date of publication.

Authors should submit manuscripts in quadruplicate. All copies should be clear, readable, and on paper of good quality. A dot matrix or unusual typeface is acceptable only if it is clear and legible. Dittoed or mimeographed copies are not acceptable and will not be considered. Authors should keep a copy of the manuscript to guard against loss. In addition to addresses and phone numbers, authors should supply electronic mail addresses and fax numbers, if available, for potential use by the editorial office and later by the production office. Mail manuscripts to the Editor, T. A. Salthouse, School of Psychology, Georgia Institute of Technology, Atlanta, Georgia 30332-0170.