

# UC Irvine

## UC Irvine Previously Published Works

### Title

Adult Life Span Changes in Immediate Visual Memory and Verbal Intelligence

### Permalink

<https://escholarship.org/uc/item/4pt323q3>

### Journal

Psychology and Aging, 10(1)

### ISSN

0882-7974

### Authors

Giambra, Leonard M

Arenberg, David

Zonderman, Alan B

et al.

### Publication Date

1995-03-01

### DOI

10.1037/0882-7974.10.1.123

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Adult Life Span Changes in Immediate Visual Memory and Verbal Intelligence

Leonard M. Giambra, David Arenberg,  
and Alan B. Zonderman  
National Institute on Aging

Claudia Kawas  
Johns Hopkins Medical Institution

Paul T. Costa, Jr.  
National Institute on Aging

A sample of 558 women and 1,163 men 17 to 102 years old, screened for neurodegenerative and neuropsychiatric disease, was administered tests of immediate visual memory (Benton Visual Retention Test) and crystallized intelligence (Wechsler Adult Intelligence Scale Vocabulary subtest) from 1 to 5 times over 27.7 years. Cross-sectional and longitudinal evidence led to the conclusion that the 65–74-year decade was a watershed for decremental changes in immediate visual memory and verbal intelligence. Age accounted for considerably less variance in vocabulary than in immediate memory. The proportion of individuals whose longitudinal trajectories were contrary to group trends decreased substantially with increased age; observed age changes remained when analyses were restricted to individuals who had perfect or near-perfect mental status scores. Selected neuronal loss and slower reproduction times were considered as possible causes.

Although many individuals change in cognitive functioning with aging, the importance, direction, magnitude, rate, and course of cognitive aging across the many distinct mental processes and functions are not well characterized. The details of normal cognitive aging changes remain among the most important areas of longitudinal investigation. The present study attempted to describe the course of normal aging in the areas of verbal intelligence and immediate visual memory.

## Aging and Immediate Memory for Visual Patterns

Visual patterns are often easily and quickly apprehended because they frequently involve highly distinctive stimuli. Thus, immediate memory for visual patterns might be expected to be well insulated from the detrimental influences of aging. However, this does not appear to be so. The Visual Reproduction subtest of the Wechsler Memory Scale (Wechsler, 1945) re-

quires immediate reproduction from memory of geometric designs after a 10-s exposure. On that subtest, Lezak (1983) reported that 80–92-year-olds had recall at a level 2.6 standard deviations below the level of 20–29-year-olds. Haaland, Linn, Hunt, and Goodwin (1983) reported a mean recall level for those 80 years old and older that was 1.3 standard deviations below the mean of 65–69-year-olds; all participants were healthy, highly educated, motivated, and noninstitutionalized. McCarty, Siegler, and Logue (1982) found significant longitudinal declines on the Visual Reproduction test after 4, 10, and 16 years for individuals initially 60–80 years old.

The Benton Visual Retention Test (BVRT; Benton, 1974), Administration A, requires immediate reproduction of geometric designs after a 10-s viewing. The most extensive investigation of BVRT immediate memory performance and adult aging was conducted by Arenberg (1978, 1982, 1987, 1990; Robertson-Tchabo & Arenberg, 1989) in the Baltimore Longitudinal Study of Aging (BLSA; Shock et al., 1984). Arenberg reported that the mean number of recall errors of 80–89-year-olds was 3.7 20–29-year-old standard deviations above that of the 20–29-year-olds mean; the mean correlation between age and BVRT errors was .48. Longitudinal intraindividual change over 13 years was consistent with cross-sectional differences with regard to trend: For individuals initially 20–39 years old, there was little or no change; for individuals initially in their 40s, 50s, 60s, and 70s, there were increasingly larger changes. Intraindividual change and initial age were correlated .40. Robertson-Tchabo, Arenberg, Tobin, and Plotz (1986) found no evidence of greater errors or greater intraindividual change on the BVRT for non-insulin-dependent diabetic men than for normal controls. A subgroup of BLSA participants who were screened for a history of kidney failure or urinary tract disease, medications affecting renal function, diuretics, and antihypertensive agents had the same pattern of intraindividual change over 13 years as did an

---

Leonard M. Giambra, David Arenberg, Alan B. Zonderman, and Paul T. Costa, Jr., National Institute on Aging, Baltimore, Maryland; Claudia Kawas, Department of Neurology, Johns Hopkins Medical Institution.

The contributions of Marcia H. Swartz, Don Reynolds, Pat Hawthorne, Karen Douglas, Barbara Hiscock, Judith Plotz, Judy Friz, Susan Robinson, Pamela Philips, Melissa Kitner-Triolo, Edwin Rosenberg, and Elizabeth Burke in collecting, scoring, and analyzing data are gratefully acknowledged. We also thank E. Jeffery Metter of the Baltimore Longitudinal Study of Aging (BLSA) for providing diagnostic information on BLSA participants.

Correspondence concerning this article should be addressed to Leonard M. Giambra, Laboratory of Personality and Cognition, Nathan W. Shock Laboratories, Gerontology Research Center, 4940 Eastern Avenue, Baltimore, Maryland 21224. Electronic mail may be sent to lmgmag%vax.dnet@dxl.nih.gov.

unscreened sample (Arenberg, 1987). When corrected for age, 13-year longitudinal change was unrelated to either systolic or diastolic blood pressure for both the healthy subsample and the full sample.

### Aging and Vocabulary

Intelligence has been analyzed into many primary mental abilities. In turn, factor analyses of primary mental abilities have revealed two major dimensions of intelligence, often referred to as fluid and crystallized (Horn, 1978). Crystallized intelligence is the application of intellect to tasks for which efficient completion benefits from prior knowledge acquired through experience, education, and acculturation. Fluid intelligence is the application of intellect to tasks for which efficient completion does not benefit from prior knowledge. Fluid intelligence has been hypothesized to decline with aging in adulthood, presumably as a result of increasing neuronal degeneration (Horn, 1978). Crystallized intelligence level has been hypothesized to remain unchanged or to increase throughout adulthood, including late life. Both hypotheses have received substantial support in cross-sectional studies using fluid and crystallized factor scores (for a review, see Kausler, 1991). Vocabulary is one of two tasks that have the greatest average correlation with the factor of crystallized intelligence (Horn, 1978); consequently, vocabulary measures have been used extensively as a single task representative of crystallized intelligence.

The Verbal Meaning subscale of the Primary Mental Abilities Test (Thurstone & Thurstone, 1949) is a timed test of vocabulary using recognition of a synonym. Age differences and longitudinal changes on this subscale from the Seattle Longitudinal Study have been summarized in Schaie (1983, 1994). Four separate cross-sectional analyses showed that Verbal Meaning scores improve before middle age, followed thereafter by increasingly larger decrements. Longitudinal changes at 7, 14, and 21 years were essentially consistent with the cross-sectional age pattern. Because Verbal Meaning is a timed measure, performance on the scale would be expected to show some deterioration with aging in adulthood (Kausler, 1991). Indeed, Schaie (1989) determined that the correlation of the Verbal Meaning scale with age was reduced to zero when a measure of perceptual speed was partialled out.

Using Educational Testing Service untimed multiple-choice vocabulary tests, Schaie and Willis (1993) reported decrements beginning in the 70s. The Vocabulary subscale of the Stanford-Binet test (Terman & Merrill, 1937) was administered twice (when participants were 25–39 years old and 60–74 years old) to a highly intelligent sample of 14 men and women; a slightly increased score was obtained when these men and women were 60–74 years old (Gilbert, 1973). The Vocabulary (VOC) subtest of the Wechsler Adult Intelligence Scale (Wechsler, 1955) is an untimed measure of vocabulary. Botwinick and Storandt (1974) reported that VOC scores peaked in the 30s decade and monotonically decreased thereafter. However, when education was taken into account, the age differences were no longer significant. Matarazzo (1977) reported essentially the same VOC means from the 20s to the early 60s.

An extensive investigation of VOC performance and adult aging in the BLSA has been carried out by Arenberg (1978,

1990). In three different cross-sectional samples of 20–89-year-old participants, small positive correlations (.14, .07, and .13) were obtained between age and unadjusted VOC scores. Thirteen-year longitudinal changes were examined for more than 400 BLSA men who were initially in their third to eighth decades. The correlation between change and initial age was  $-.43$ , showing that VOC scores increased for the younger participants but declined for the older participants. Wilkie and Eisdorfer (1985) reported a small but significant decrement in VOC scores for 75–94-year-old men and women relative to their scores 15 years earlier.

In this article, we extend Arenberg's prior investigations of age differences and age changes on BVRT and VOC performance to (a) a larger cross-sectional age sample, (b) a larger 7-year longitudinal sample including women, (c) a larger 13-year longitudinal sample, and (d) a 19- and 25-year longitudinal sample. It has been shown that disease states such as diabetes and hypertension do not appear to affect BVRT performance. However, diseases associated with mental functioning may be responsible for age differences and changes in BVRT performance as well as the apparent decline of vocabulary in old-old individuals. Therefore, we restricted our sample to those men and women who had no history or diagnosis of neurodegenerative or neuropsychiatric disease. Furthermore, when possible, we restricted the sample to older adults with no evidence of mental impairment as measured by a separate mental status instrument. Finally, we examined the distributions of intraindividual change for evidence of individuals who did not change or who changed in the contrary direction.

We have, on the same individuals, long-term data on intraindividual change for a narrow cognitive function (immediate visual memory for designs) and for a broad cognitive function (vocabulary). Our literature review has pointed to very different trajectories for these two functions. In this article, we show the dissociation between the adult life span trajectories of these two cognitive functions.

## Method

### *Participants*

The BLSA (Shock et al., 1984) began in 1958; it is an open panel in which participants are continually recruited. About 90% of the participants are recruited by word of mouth. The remainder are recruited by invitation of BLSA staff members or through individual initiative after learning of the BLSA in the media. Women were first recruited in 1978. Strenuous effort has been expended in keeping participants from dropping out and, if they have dropped out, in returning them to the panel. About 25% of all participants (excluding instances of deaths) have voluntarily or involuntarily dropped out. Analyses of attrition effects are described in the Results section.

The BVRT (Benton, 1974) and the VOC (Wechsler, 1955) have been administered to all male participants since 1960 and to all female participants since 1978. As of April 1992, 1,384 men and 651 women had had at least one valid measure on the BVRT. In this study, we limited our sample to participants who had not developed, during their lifetime, any of the following neurologic or major psychiatric diseases: psychosis, major depression, organic brain syndrome, dementia, Parkinson's disease, stroke, and epilepsy. After removing participants with major neurological or neuropsychiatric disease, we had 1,163 men and 558 women with a valid BVRT and VOC at the first testing.

The participants in the BLSA are predominantly White, highly edu-

cated men and women. In our sample, 1.7% of men and 2.3% of women had less than 12 years of education, 7.1% of men and 13.7% of women had 12 years of education, 45.5% of men and 44.9% of women had 13–16 years of education, and 53.7% of men and 39.1% of women had 17 or more years of education. There was no correlation between age and years of education for men and women. A high proportion of men and many of the women were employed in, or retired from, administrative or professional positions.

### *Participants in Cross-Sectional Sample*

This sample consisted of 1,721 BLSA participants who took the BVRT and the VOC for the first time, usually on their first BLSA visit. The age distribution (in years), by indicated intervals, was as follows: 17–21, 20 men and 8 women; 22–27, 89 men and 70 women; 28–33, 160 men and 64 women; 34–39, 127 men and 55 women; 40–45, 118 men and 36 women; 46–51, 122 men and 27 women; 52–57, 93 men and 48 women; 58–63, 102 men and 61 women; 64–69, 98 men and 52 women; 70–75, 126 men and 68 women; 76–81, 77 men and 39 women; 82–87, 26 men and 24 women; and 88–102, 5 men and 6 women.

### *Participants in Longitudinal Sample*

Four longitudinal periods were used: 5.95 to 9.7 years ( $M = 6.8$ ,  $SD = 0.8$ ), 11.95 to 15.7 years ( $M = 13.1$ ,  $SD = 0.9$ ), 17.95 to 21.7 years ( $M = 19.0$ ,  $SD = 0.8$ ), and 23.95 to 27.7 years ( $M = 25.1$ ,  $SD = 0.7$ ). These periods are henceforth referred to as the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods, respectively. For most participants, the 6.8-year period occurred at the second testing (599 men and 206 women), the 13.1-year period occurred at the third testing (367 men), the 19.0-year period occurred at the fourth testing (256 men), and the 25.1-year period occurred at the fifth testing (112 men). In some cases, BLSA participants did not always adhere to a 2-year interval between visits; thus, some participants achieved the longer longitudinal periods in fewer than the usual number of testings. However, in most cases, the longer longitudinal period participants were a subsample of participants in the shorter longitudinal intervals.

## *Measures*

### *BVRT*

The BVRT assesses immediate reproductive memory for geometric designs. Three “equivalent” forms were used, each consisting of 10 test cards. A single geometric figure appeared on each of the first 2 cards. The other 8 cards included two major geometric figures and a peripheral figure. Each card was displayed for 10 s and then removed from view. The participant had as much time as needed to reproduce from memory the design on the card. Our performance measure was the total number of errors in reproducing the original designs on all 10 cards. Form C was used at the first and fourth testings. Form E was used at the second and fifth testings. Form D was used at the third testing.

To permit our own evaluation of the equivalence of Forms C, D, and E, we administered two forms to each BLSA participant as follows: Form C followed Form E at the second testing (170 men and 253 women), Form E followed Form D at the third testing (189 men), and Form D followed Form C at the fourth testing (172 men). The correlations were .71 between Forms E and C, .74 between Forms E and D, and .71 between Forms C and D (all  $ps < .01$ ). These correlations were somewhat smaller than the .79 to .85 values reported by Benton (1974). On average, 0.57 more errors occurred on Form E than on Form C,  $t(422) = 4.37$ ,  $p < .001$ ; 0.18 more errors occurred on Form D than on Form E,  $t(188) = 0.95$ ,  $p > .05$ ; and 0.88 more errors occurred on Form D than on Form C,  $t(171) = 4.33$ ,  $p < .001$ . These results suggest that Form D was more difficult than Form E, which was more difficult than Form C. The findings were also consistent with comparisons reported

in the BVRT manual (Benton, 1974). Because the sequence of forms was C, E, D, C, E and because both Forms E and D produced more errors than Form C, longitudinal changes between the first and second, first and third, and first and fifth testings would appear to be biased in the direction of more errors in the second, third, and fifth testings. This represented a bias, small in magnitude, in the direction of longitudinal change, a consequence we consider in the Results section.

### *VOC*

Standard instructions (Wechsler, 1955) were used in administering and scoring the VOC; participants attempted to define, without time limits, each of 40 words. Here we report raw scores (i.e., the scores were unscaled and not adjusted in any way). The BVRT and VOC were almost always given during the same session and were scheduled to occur every third or fourth visit (usually a 6- to 8-year interval).

### *Procedure and Scoring*

Participants' recollections (drawings) for the BVRT were independently evaluated for errors by two psychologists according to the test manual (Benton, 1974), and disagreements were resolved by consensus or by a third independent rater. The VOC was evaluated serially by two psychologists (i.e., the second rater saw the first rater's scoring, according to the test manual; Wechsler, 1955); disagreements were resolved by consensus.

### *Evaluation of Disease and Mental Status*

At each visit, BLSA participants provided a medical history that was reviewed, with the participant, by the examining physician. Histories were often supplemented by summaries of physician and hospital records. The physician conducted a physical examination and, when necessary, ordered special diagnostic laboratory tests. Most participants were given standard laboratory tests including urinalysis, hemograms, creatine clearance, and plasma cholesterol and triglycerides. Also, tests of neuromuscular, pulmonary, and cardiovascular function were carried out (see Shock et al., 1984, for a complete description). Beginning in 1986, older participants were examined by a neurologist and given a neuropsychological battery. Some brains were made available for further pathological analysis after autopsy. The neurologist also conducted follow-up examinations of older participants who had dropped out to determine the presence of neurodegenerative disease. When appropriate, some participants had noninvasive neuroimaging (e.g., computerized tomography).

The Blessed Information-Memory-Concentration test (BIMC; Blessed, Tomlinson, & Roth, 1968) is a brief screening instrument for possible cognitive impairment. Six cognitive domains are measured: orientation to time and place, personal orientation, personal information, knowledge memory, and delayed recall. In this study, the BIMC was individually administered and scored as the number of erroneous responses (maximum of 30). It was introduced to the BLSA participants 70 years of age and older in 1986; in 1990, administration was extended to 60- to 69-year-olds.

## *Results*

### *Overview*

Before presenting the findings in detail, we provide a brief overview of the Results section by indicating its hierarchical organization and a thumbnail sketch of the more interesting outcomes.

### Analyses Involving Participants of Different Ages

*Age differences.* For VOC, age accounted for 1.2% of the variance for women and 2.3% of the variance for men. For BVRT, age accounted for 38.2% of the variance for women and 31.5% of the variance for men.

Little correlation was observed between tenure and initial performance on either test; for men, those who dropped out made 0.3 more BVRT errors than men who remained in the study.

*Gender differences.* Men's VOC scores were 2.7 points higher than those of women.

*Epoch differences.* No consistent epoch effects were observed on either test.

### Analyses Involving Longitudinal Changes in Participants

*Multiple regression, change as dependent measure.* There was little relationship between change on the BVRT and change on the VOC. Second, greater initial scores resulted in greater longitudinal declines in VOC scores and smaller longitudinal increases in BVRT errors. Third, for men administered the VOC, the proportions of variance accounted for by age or age-function variables were 3.1%, 5.9%, 11.1%, and 1.9%, respectively, for the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods. Restricting analyses to participants with BIMC scores of 2 or less (scores of 3 or more indicated evidence of impaired mental status) reduced the variance accounted for in the 6.8- and 19.0-year periods. Finally, the variance accounted for by age or age-function variables on the BVRT was between 11.2% and 22.2% for men and women across the longitudinal periods. BIMC score restrictions had virtually no effect.

### ANOVA-Based Analyses of Longitudinal Change

*ANOVA analyses of longitudinal period differences.* For VOC, significant differences between longitudinal periods occurred for the 64–69 and 76–81 age-at-later-testing groups on the VOC. On the BVRT, significant differences between longitudinal periods occurred for the 64–69 and 70–75 age-at-later-testing groups.

*A finer grained view of longitudinal change.* Regardless of longitudinal period, there was improvement on the VOC in the youngest to the 58–63 age-at-later-testing groups; in the 64–69 age-at-later-testing group and older groups, some decrement occurred that increased with older groups. On the BVRT, the 64–69 age-at-later-testing group and older groups had significant changes for all longitudinal periods; longer periods and older age-at-later-testing groups resulted in progressively larger decrements (approximately).

*Tenure in the BLSA and longitudinal change.* Shorter term participants showed evidence of greater decremental longitudinal changes than longer term participants; group differences were small.

*Individual variation in intraindividual change.* Substantial minorities of all age-at-later-testing groups at all longitudinal periods showed no longitudinal change or changes in a direction opposite to that indicated by group mean changes.



Figure 1. Wechsler Adult Intelligence Scale Vocabulary (unadjusted raw) scores across the adult life span.

### Analyses Involving Participants of Different Ages

#### Age Differences

Figure 1 shows a convex but weak functional relationship between age group and unadjusted (raw) VOC scores. After age 21, there was a small increment in mean scores until middle age, followed by stability until after age 81 when a relatively large decrement in means occurred. This trend was more evident for men than women. The correlations between age and VOC unadjusted (raw) score were .12 for men and .10 for women. A stepwise multiple regression entering successively linear, quadratic, and cubic age components indicated that the linear and quadratic components were significant for men and that the linear component was significant for women. For men, the linear component accounted for 1.3% of the variance,  $F(1, 1162) = 15.68, p < .001$ ; the quadratic component, 0.7%,  $F(1, 1161) = 7.75, p < .01$ ; and the cubic component, 0.3%,  $F(1, 1160) = 3.49, p < .10$ . For women, the linear component accounted for 0.9% of the variance,  $F(1, 557) = 5.13, p < .05$ ; the quadratic component, 0.2%,  $F(1, 556) < 1, p > .05$ ; and the cubic component, 0.1%,  $F(1, 555) < 1, p > .05$ . A one-way ANOVA involving 12 age groups for men and 11 age groups for women (see Figure 1) yielded a significant age group effect for men,  $F(11, 1151) = 3.26, p < .01, MSE = 100.31, \omega^2 = .021$ , but not for women,  $F(10, 534) = 1.47, p > .10$ .

The correlations between age and BVRT errors were .53 for men and .60 for women. Figure 2 depicts the age group means for BVRT total errors for men and women separately; both men and women showed accelerating (i.e., nonlinear) age differences with increasing age. A stepwise multiple regression entering successively linear, quadratic, and cubic age components indicated that all components were significant for men and that the linear and quadratic components were significant for women. For men, the linear component accounted for 28.0% of the vari-

<sup>1</sup> The sample size for the women's 17–21-year age group was too small to be included, and its omission may have been responsible for the lack of a significant effect. When 17–21-year-olds were omitted with the men, however, the age group effect remained significant.

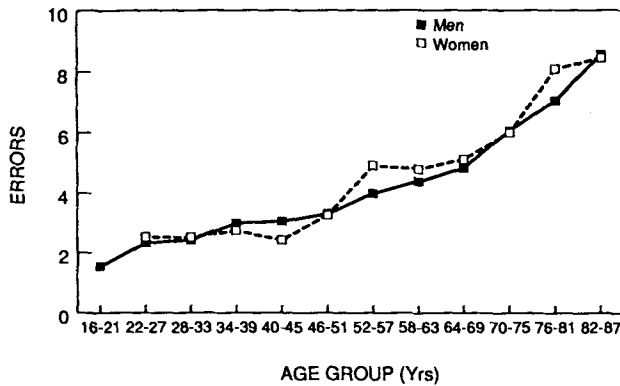


Figure 2. Mean Benton Visual Retention Test total across the adult life span.

ance,  $F(1, 1162) = 452.33, p < .001$ ; the quadratic component, 2.6%,  $F(1, 1161) = 43.55, p < .001$ ; and the cubic component, 0.9%,  $F(1, 1160) = 15.96, p < .001$ . For women, the linear component accounted for 36.1% of the variance,  $F(1, 557) = 314.08, p < .001$ ; the quadratic component, 2.2%,  $F(1, 556) = 19.97, p < .001$ ; and the cubic component, 0.0%,  $F(1, 555) < 1, p > .05$ . Although the nonlinear components contributed substantially less than the linear components, the age function was clearly nonlinear in form. This was confirmed when an exponential function was found to account for essentially the same amount of variance as the linear, quadratic, and cubic components combined. For the men, a least squares fit of an exponential function yielded the function  $BVRT \text{ errors} = 1.89 + 0.1774 \times \exp(0.0432 \times \text{age})$ , which accounted for 31.5% of the variance,  $F(3, 1160) = 177.51, p < .01$ ; all parameter estimates were significant ( $p < .05$ ). For the women, the function  $BVRT \text{ errors} = 1.2048 \times \exp(0.0229 \times \text{age})$  accounted for 38.2% of the variance,  $F(2, 556) = 171.77, p < .01$ , with all parameter estimates significant ( $p < .05$ ). The correlation between VOC scores and BVRT errors was  $-.18$  for both the women and men (both  $ps < .01$ ), indicating that crystallized intelligence tended to covary with immediate memory, but that the relationship was a weak one.

Although selective attrition in the BLSA can play no role in the estimation of any age differences at the first testing, it can have a significant role in longitudinal effects. Selective attrition (for a further explanation, see Kausler, 1991) can act to hide or minimize age changes by a dropping out of the less able participants. One way to determine whether ability is related to attrition is to examine the relationship between tenure (interval from entering to exiting) in the BLSA and performance at first testing (i.e., longer tenure should result in fewer errors on the BVRT and higher scores on the VOC).

Male and female participants who were tested for the first time on the VOC and the BVRT had mean tenures in the BLSA of 11.6 years ( $SD = 9.4$ ) and 6.4 years ( $SD = 5.17$ ), respectively; older participants had shorter mean tenures than younger participants. Tenure was measured from the time of first testing on the VOC and BVRT. Thus, performance at the start of each participant's tenure can be examined relative to tenure length. The most direct measure of the relationship of performance to tenure was a correlation between tenure and BVRT errors or

VOC unadjusted (raw) score. This correlation was obtained for each age group.<sup>2</sup> For the VOC, the within-age-group correlations ranged from .01 to .30 for men, with 2 of the 12 correlations being significant; these correlations ranged from .01 to .41 for women, with only 1 of the 11 being significant ( $p < .05$ ). For the BVRT, the within-age-group correlations ranged from  $-.21$  to .04 for men and from  $-.50$  to .00 for women (only the  $-.21$  and  $-.50$  correlations were significant,  $p < .05$ ). Thus, there was little covariation between initial performance and eventual tenure in the BLSA, regardless of the age group involved.

A more general way to examine the relationship between tenure and performance was through an Age Group  $\times$  Dropout ANOVA. For men, a dropout was defined as not being in the BLSA after 1986 and not having a fifth BVRT or VOC testing. For women, a dropout was defined as not being in the BLSA after 1986 and not having a third BVRT or VOC testing. Table 1 shows the number of dropouts for each age interval. When a minimum cell size of 15 was required, the men had 10 age groups (22–27 through 76–81 years) and the women had 2 age groups (28–33 and 64–69 years) that met this criterion; ANOVAs were restricted to the better-sampled men. For men, the ANOVA yielded a significant dropout effect for BVRT errors,  $F(1, 1092) = 5.10, p < .05, MSE = 5.96, \omega^2 = .003$ . Those BLSA participants who dropped out made more BVRT errors ( $M = 4.29$ ) than those who remained in the program ( $M = 3.93$ ). There also was a significant Age Group  $\times$  Dropout interaction for the VOC,  $F(9, 1092) = 1.93, p < .05, MSE = 95.17, \omega^2 = .007$ . Those VOC participants who dropped out had lower scores (by 2.4 points) in the 22–45, 52–57, and 76–81 age intervals and higher scores (by 2.5 points) in the 46–51 and 58–75 intervals.

### Gender Differences

Gender  $\times$  Age Group ANOVAs on first testing BVRT total errors and unadjusted (raw) VOC score yielded a gender main effect on the VOC,  $F(1, 1660) = 21.26, p < .01, MSE = 104.59, \omega^2 = .011$ . Men had a higher mean score (64.2) than women (61.5). When the ANOVAs were restricted to men who entered the BLSA program contemporaneous with the women (see Method section), a significant gender effect was again found for the VOC,  $F(1, 978) = 4.57, p < .05, MSE = 108.21, \omega^2 = .004$  (mean score for men of 63.1). These same men and women also differed on the BVRT; men had fewer errors ( $M = 4.14$ ) than women ( $M = 4.65$ ),  $F(1, 978) = 7.37, p < .01, MSE = 5.78, \omega^2 = .004$ .

### Epoch Differences

Because BLSA participants have been administered the BVRT and VOC since 1960, we examined the sample for period effects. That is, we sought to determine whether there were changes in sample composition or unknown changes in test administration and scoring, or whether there were changes in general exposure to geometric designs or their recall, as a result of

<sup>2</sup> The correlation using all participants was biased because older participants had a shorter tenure and poorer performance than younger ones.

Table 1  
Maximum Sample Size and Number Who Dropped Out for 6-Year Age Intervals

Age at later testing (years)	Longitudinal period (years)													
	First testing				5.95-9.7				11.95-15.7 (men)		17.95-21.7 (men)		23.95-27.7 (men)	
	Women		Men		Women*		Men		Full sample	Dropouts	Full sample	Dropouts	Full sample	Dropouts
16-21	8	1	20	4										
22-27	70	8	89	18										
28-33	64	15	160	38	10	0	23	1						
34-39	55	10	127	33	24	0	102	10	14	0				
40-45	36	5	118	39	27	0	117	14	54	5	10	0		
46-51	27	3	122	57	28	0	156	22	77	4	35	1		
52-57	48	7	93	53	25	1	178	41	102	12	57	3		
58-63	61	12	102	61	30	0	156	40	117	19	77	4	27	0
64-69	52	18	98	48	36	1	130	49	97	22	77	8	29	1
70-75	68	14	126	77	27	1	90	44	58	16	58	4	35	3
76-81	39	7	77	28	16	0	75	42	38	17	28	6	23	0
82-87	24	3	26	16	14	1	39	19	21	14	11	4		
88+	6	2	5	4										

Note. A dropout was defined as a participant who was not involved in the program after 1986 and did not have a specific subsequent testing (third for women and fifth for men). Age at later testing was the participant's age on the later of the two testings included in the longitudinal period (see section "Maximizing age group representation" for a further explanation) and the only age for first test in the cross-sectional sample. The full sample columns represent the number of participants who were tested. The dropouts columns represent the number of participants who eventually dropped out of the study. Thus, the number who stayed was the difference between the full sample and dropouts (e.g., for the 11.95-15.7-year longitudinal period for the 76-81-year age-at-later-testing group, the full sample was 58, the number of dropouts was 16, and the number who remained in the study was 42[58 - 16]).

\* The total sample size is greater than the 206 reported in the Method section as a result of some 57 women who received a third testing.

the epoch in which a participant entered the BLSA study. For men, three epochs (1960-1965, 1966-1977, and 1984-1992) had sufficient sample sizes in four age groups (17-39, 50-59, 60-69, and 70-102) to allow for an examination of period effects. Epoch  $\times$  Age Group ANOVAs on first testing BVRT total errors and unadjusted (raw) VOC score yielded a significant epoch main effect for BVRT total errors,  $F(2, 837) = 6.41$ ,  $p < .01$ ,  $MSE = 7.15$ ,  $\omega^2 = .009$ , and for VOC scores,  $F(2, 840) = 4.25$ ,  $p < .05$ ,  $MSE = 97.27$ ,  $\omega^2 = .007$ ; also, there was a significant Epoch  $\times$  Age Group interaction effect on VOC scores,  $F(6, 840) = 2.17$ ,  $p < .05$ ,  $MSE = 97.27$ ,  $\omega^2 = .004$ . The 1960-1965 epoch produced the highest unadjusted (raw) VOC scores, and the 1966-1977 epoch produced the lowest scores; this pattern occurred for all age groups except the 60-69-year-olds, who had their lowest score in the 1984-1992 epoch. The 1966-1977 epoch produced more BVRT errors than the 1960-1965 and 1984-1992 epochs, which had the same number of errors.

Because women first entered the BLSA in 1978, we examined two epochs (1978-1983 and 1984-1992) that had sufficient representation in five age groups (20-29, 30-39, 50-59, 60-69, and 70-79). Epoch  $\times$  Age Group ANOVAs on BVRT errors and unadjusted VOC score at first testing yielded no significant epoch effects, both  $F_s < 1$ , or Epoch  $\times$  Age Group interactions, both  $F_s(4, 456) < 2.24$ ,  $p_s > .05$ .

#### Analyses Involving Longitudinal Changes in Participants

##### Multiple Regression, Change as Dependent Measure

In multiple regression analyses with change as the dependent measure, participants' performance at the first testing was sub-

tracted from their performance on the  $i$ th testing to obtain a change score for the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods ( $i = 2, 3, 4$ , and 5, respectively). Before presenting the regression outcomes, we examined the zero-order correlations between change and age. For men, the correlations between age and change on the VOC were  $-.20$ ,  $-.34$ ,  $-.39$ , and  $-.18$ , respectively, for the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods; for women, the correlation was  $-.12$  for the 6.8-year longitudinal period. Also, for men, increasing age showed a reduction in a positive longitudinal change or an increased negative longitudinal change in VOC unadjusted (raw) score (i.e., subsequent VOC scores were reduced with increased age). The correlations of age and change in BVRT errors for men were  $.21$ ,  $.26$ ,  $.30$ , and  $.32$ , respectively, for the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods; for women, the correlation was  $.19$  for the 6.8-year period. The longitudinal increase in BVRT errors increased with initial age. Furthermore, this correlation increased with longer longitudinal periods. For men, the correlations of change on the VOC and change in BVRT errors were  $-.11$ ,  $-.17$ ,  $-.14$ , and  $-.07$ , respectively, for the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods; for women, the correlation was  $-.01$  for the 6.8-year period. There was little linear relationship between change on the VOC and change on the BVRT for the longitudinal samples as a whole.

In a stepwise order, the following independent variables were used to predict longitudinal change in performance on the VOC and on the BVRT for each of the longitudinal periods: actual longitudinal period, first test score, first test age, interacting of first test score and first test age, age squared, and age cubed. In a second stepwise multiple regression, only the independent variables that added a significant amount of variance were retained. Table 2 contains a summary of this second multiple re-

Table 2  
*Regression of Longitudinal Change in WAIS Vocabulary Unadjusted Raw Score and Benton Visual Retention Test Total Errors on First Testing and Age: Stepwise Results*

Gender	Longitudinal interval (years)	WAIS Vocabulary raw score				Benton Visual Retention Test errors				Sample size
		Step	Variable	Variance increase	Regression coefficient	Step	Variable	Variance increase	Regression coefficient	
Women	5.95 to 9.7	1	Test 1	.110***	-0.1458***	1	Test 1	.039**	-1.0875***	206
		2		Intercept = 10.70***		2	Test 1 × Age	.112***	0.0129***	
Men	5.95 to 9.7	1	Test 1	.165***	-0.3603***	1	Test 1	.063***	-0.4341***	599
		2	Age	.020***	-0.2723***	2	Age squared	.133***	0.0017***	
		3	Test 1 × Age	.011**	0.0034**	3	Age	.007*	-0.0977*	
	11.95 to 15.7	1	Test 1	.220***	-0.2150***	1	Test 1	.078***	-1.0501***	367
		2	Age	.059***	-0.1044***	2	Age	.140***	0.0434**	
		3		Intercept = 18.66***		3	Test 1 × Age	.022**	0.0122**	
	17.95 to 21.7	1	Test 1	.166***	-0.1971***	1	Test 1	.083***	-1.1673***	256
		2	Age cubed	.111***	-2.62E-05***	2	Age	.156***	-0.2167**	
				Intercept = 15.64***		3	Test 1 × Age	.042***	0.0160**	
						4	Age squared	.024**	0.0026**	
	23.95 to 27.7	1	Test 1	.298***	-0.3324***	1	Test 1	.154***	-0.6079***	112
		2	Age squared	.019	-0.0011	2	Age	.170***	0.1582***	
				Intercept = 23.98***				Intercept = -3.57*		

Note. Longitudinal change was the *i*th testing performance minus the first testing performance (see text for further details). Age is at first testing.

\* Shorthand for -0.0000262.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

gression, including final regression coefficients and the constant (i.e.,  $y$ -intercept).

It can be seen from Table 2 that, for all longitudinal periods and for both the VOC and the BVRT, score at first testing predicted a significant amount of the variance of longitudinal change (from 3.9% to 29.8%). Furthermore, the partial regression coefficients were all negative, indicating that the higher the initial VOC or BVRT score, the greater the decline in VOC scores and the smaller the increase in BVRT errors. For men only, a significant amount of VOC change variance was accounted for by age in the 6.8-year (2.0%) and 13.1-year (5.9%) longitudinal periods and by age cubed in the 19.0-year (11.1%) period. In all cases, the partial regression coefficients were negative, indicating that with greater age there was a greater likelihood of VOC scores falling. The Age × First Test score interaction accounted for a small but significant amount of variance (1.1%) for men for the 6.8-year longitudinal period; greater initial VOC scores in combination with greater age resulted in an increased second VOC score. When the samples were restricted to participants who had never had a BIMC score of 3 or more (i.e., had never shown any evidence of impaired mental status), the effect of the previously described age variables was (a) no longer significant for the 6.8-year longitudinal period, (b) unchanged for the 13.1- and 25.1-year periods, and (c) reduced in significance and percentage of variance accounted for (by two thirds) in the 19.0-year period.

In terms of BVRT errors, a significant amount of variance (14.0% to 18.0%) was accounted for by age or age squared for

men in all four longitudinal periods after the influence of initial score had been removed. For the 6.8- and 19.0-year longitudinal periods, the age and age squared terms were opposite in sign, indicating minimum changes (negative in value) at 29.1 years of age for the 6.8-year period and at 37.8 years of age for the 19.0-year period (see, however, the significant interaction described next). Age × First Test Score interactions were significant for women in the 6.8-year longitudinal period and for men in the 13.1- and 19.0-year periods. In all cases, the partial regression coefficients were positive, indicating that greater age combined with greater initial errors to disproportionately increase positive change scores (i.e., subsequent testings for older participants led to greatly increased BVRT errors). When the samples were restricted to participants who had never had a BIMC score of 3 or more, i.e., never showed any evidence of impaired mental status, the age effects were essentially unchanged for all longitudinal periods.

Earlier we determined that there was a procedural bias toward more BVRT errors with increased age resulting from Form E or D being used in the second, third, and fifth testings (i.e., the 6.8-, 13.1-, and 25.1-year longitudinal periods). This bias did not affect the percentage of variance and partial regression coefficients in the multiple regression analyses described earlier because it would have affected all participants. However, the bias did increase the magnitude of the regression intercept. For example, because 199 of 206 women had Form E followed by Form C at their second testing, in comparison with 149 of 599 men, one can directly compare regression outcomes for C



Table 3

*An Illustration of How to Maximize Sample Size for a Longitudinal Period for a Particular Age at Later Testing: Age at Testing of Four Hypothetical Longitudinal Participants*

Participant	Longitudinal period (years)		Age at later testing (years)	Testing				
	Maximum	Used		Last - 4	Last - 3	Last - 2	Last - 1	Last
P1	6	6	56				50	56
P2	12	6	56			44	50	56
P3	18	6	56		38	44	50	56
P4	24	6	56	32	38	44	50	56
P2	12	6	50			44	50	56
P3	18	6	50		38	44	50	56
P4	24	6	50	32	38	44	50	56
P3	18	6	44		38	44	50	56
P4	24	6	44	32	38	44	50	56
P4	24	6	38	32	38	44	50	56
P2	12	12	56			44	50	56
P3	18	12	56		38	44	50	56
P4	24	12	56	32	38	44	50	56
P3	18	12	50		38	44	50	56
P4	24	12	50	32	38	44	50	56
P4	24	12	44	32	38	44	50	56
P3	18	18	56		38	44	50	56
P4	24	18	56	32	38	44	50	56
P4	24	18	50	32	38	44	50	56

to C and C to E changes. For the C to C changes, we found that (a) the first test accounted for 3.7% of variance and had a coefficient of  $-1.0933$ , (b) First Test Score  $\times$  Age accounted for 13.4% of variance and had a coefficient of  $0.0133$ , and (c) the intercept was  $1.64$  (see Table 2 for comparison values).

#### *ANOVA-Based Examinations of Longitudinal Change*

*Maximizing age group representation.* Some longitudinal participants were tested as many as five times. Table 3 shows four hypothetical longitudinal participants who were last tested at age 56. Participant 1 (P1) was tested twice; Participant 2 (P2) was tested three times; Participant 3 (P3) was tested four times; and Participant 4 (P4) was tested five times. Each of these participants had 6 years between testings and a different initial testing age. P1 had one 6-year longitudinal period (50–56 years); P2 had two 6-year longitudinal periods (44–50 and 50–56 years); P3 had three 6-year longitudinal periods (38–44, 44–50, and 50–56 years); and P4 had four 6-year longitudinal periods (32–38, 38–44, 44–50, and 50–56 years). Therefore, P1 contributed information on longitudinal change for one 6-year interval, P2 contributed information on two 6-year intervals, P3 contributed information on three 6-year intervals, and P4 contributed information on four 6-year intervals. Thus, if the maximum information on 6-year longitudinal change between 50 and 56 years of age were to be obtained, all four participants would contribute; for 44–50 years, three participants (P2, P3, and P4) would contribute; for 38–44 years, two participants (P3 and P4) would contribute; and, for 32–38 years, one participant (P4) would contribute. The cost of this gain in informa-

tion at specific ages is that longitudinal change estimates for different ages come from interdependent samples instead of independent samples, as was the case in the earlier regression analyses for the 6.8-year longitudinal period.<sup>3</sup> A similar maximization of sample size at a particular age interval can occur for the 12-year and 18-year longitudinal periods depicted in Table 3. With the 12-year longitudinal period, for example, P2, P3, and P4 could contribute to the 44–56-year interval; P3 and P4 could contribute to the 38–50-year interval; and P4 could contribute to the 32–44-year interval. Similar sample size maximization can occur for the 18.0-year longitudinal period. Table 1 shows the sample sizes for the four longitudinal periods used in this study at specific age-at-later-testing intervals. The VOC and BVRT means and standard deviations for each of these groups are shown in Appendix A.

In addition to maximizing sample sizes to examine changes at any particular age, this method of aggregating longitudinal participants also permits a longitudinal period one-way ANOVA for each age at later testing. This method statistically compares longitudinal changes associated with different longitudinal periods; as a result of the use of age at later testing, the different longitudinal period samples were independent of each other (i.e., they included different participants). An ANOVA was carried out for each age-at-later-testing interval with two or

<sup>3</sup> By using 6-year age intervals with the 5.95–9.7-year longitudinal period, we eliminated virtually all age overlap between the earlier and later testings. For the longer longitudinal periods, age overlap between the earlier and later testings could not occur.

Table 4  
Results of One-Way Longitudinal Period Analyses of Variance on Longitudinal Change Scores: WAIS Vocabulary Unadjusted Raw Scores and BVRT Total Errors for Men at Specific 6-Year Age-at-Later-Testing Intervals

Age at later testing (years)	Longitudinal periods (years)	df	F	WAIS Vocabulary significant contrast (longitudinal period)	F	BVRT significant contrast (longitudinal period)
40-45	6.8, 13.1	1, 169	<1		1.45	
46-51	6.8, 13.1, 19.0	2, 265	<1		<1	
52-57	6.8, 13.1, 19.0	2, 334	<1		<1	
58-63	6.8, 13.1, 19.0, 25.1	3, 373	<1		<1	
64-69	6.8, 13.1, 19.0, 25.1	3, 329	2.52	6.8-25.1* 13.1-25.1*	1.79	6.8-25.1*
70-75	6.8, 13.1, 19.0, 25.1	3, 237	<1		1.93	6.8-25.1* 19.0-25.1*
76-81	6.8, 13.1, 19.0, 25.1	3, 160	2.82*	6.8-19.0* 6.8-25.1*	<1	
82-87	6.8, 13.1	1, 58	1.97		<1	

Note. For longitudinal periods, 6.8 = 5.95-9.7, 13.1 = 11.95-15.7, 19.0 = 17.95-21.7, and 25.1 = 23.95-27.7. The change score was defined as later testing score minus earlier testing score. WAIS = Wechsler Adult Intelligence Scale; BVRT = Benton Visual Retention Test.  
\*  $p < .05$ .

more longitudinal periods consisting of samples of at least 15 men. For example, for the 46-51-year age-at-later-testing ANOVA, three longitudinal intervals (6.8, 13.1, and 19.0 years) were represented by (maximized) independent samples of sufficient size (see Table 4).

Finally, by maximizing sample sizes, this approach also permitted examination of the effect of remaining in or dropping out of the BLSA study for the 6.8- and 13.1-year longitudinal periods. For participants in a particular age-at-later-testing interval, longitudinal changes for those who remained in the study were compared with changes for those who dropped out of the study in an independent groups ANOVA for each of the two longitudinal periods. Specifically, for participants in a particular age-at-later-testing interval, longitudinal changes for those who never would have more than two or three testings were compared with changes for those who eventually would have four or five testings in an independent groups ANOVA (see Table 1 for the number of participants who dropped out within a particular age-at-later-testing interval and longitudinal period).

*Analyses of longitudinal period differences.* For men, a one-way longitudinal period ANOVA was carried out on longitudinal change (later minus earlier score) at each 6-year age-at-later-testing interval for all 6-year intervals with at least two longitudinal periods represented. Also, a priori contrasts between each pair of longitudinal periods were carried out. Table 4 indicates the outcomes of the ANOVAs and contrasts. In terms of changes on the VOC, longitudinal periods differed only for the 76-81-year-olds at later testing,  $F(3, 160) = 2.82, p < .05, \omega^2 = .032$ ; significant contrasts occurred between the 6.8- and 25.1-year longitudinal periods and between the 13.1- and 25.1-year periods. The mean changes were -1.0, -2.4, -3.0, and -3.7, respectively, for the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods (see Figure 3). Significant contrasts also occurred between the 6.8- and 25.1-year periods and the 13.1- and 25.1-year periods for the 64-69-year-olds at later testing. For BVRT errors, no overall effect was found at any age at later testing; however, significant contrasts occurred between the 6.8- and 25.1-year periods for the 64-69-year-olds and 70-75-year-olds

at later testing and between the 19.0- and 25.1-year periods for the 70-75-year-olds at later testing (see Figure 4 for the mean longitudinal changes).

*Finer grained view of longitudinal changes.* To determine whether the mean longitudinal changes were significantly different from zero, we conducted a one-way testing (earlier or later) repeated measures ANOVA using the maximized samples for each 6-year age-at-later-testing interval for the 6.8-year longitudinal period for men and women separately and for the 13.1-, 19.0-, and 25.1-year periods for men only. Table 5 con-

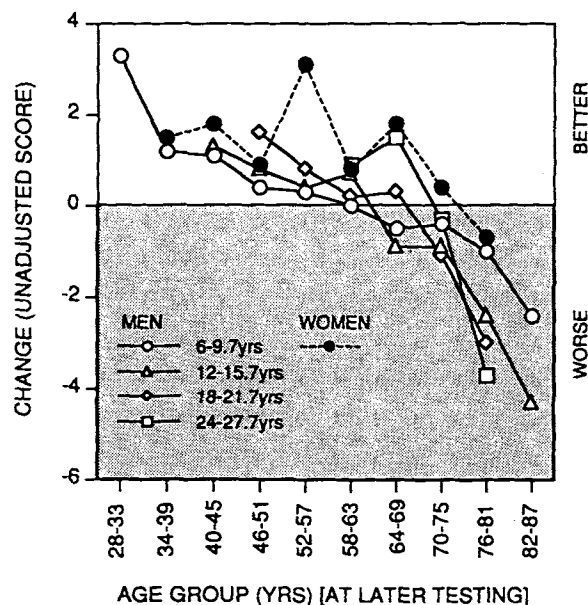


Figure 3. Mean changes (later minus earlier) in Wechsler Adult Intelligence Scale Vocabulary (unadjusted raw) scores for the 5.95-9.7, 11.95-15.7, 17.95-21.7, and 23.95-27.7 longitudinal periods for 6-year age-at-later-testing intervals. Age at earlier testing can be derived by subtracting the mean longitudinal period from age at later testing.

tains the results of these ANOVAs; entries are the percentages of variance accounted for by the longitudinal period (as estimated by  $\omega^2$ ) when the longitudinal change was significantly different from zero.

On the VOC, the 46–51, 70–75, and 76–81 age-at-later-testing intervals showed nonsignificant 6.8-year longitudinal changes and the 40–45 interval showed significant 6.8-year longitudinal changes for both men and women. In the remaining intervals, either the women or the men, but not both, had a significant change. For men 28–33, 34–39, and 40–45 years old, VOC scores increased significantly from when these men were 6.8 years younger; for 82–87-year-old men, VOC scores declined from when the men were 6.8 years younger; for 82–87-year-old men, VOC scores declined from when the men were 6.8 years younger. Two intervals—52–57 years,  $F(1, 201) = 8.55, p < .01, \omega^2 = .036$ , and 64–69 years,  $F(1, 164) = 6.89, p < .01, \omega^2 = .034$ —exhibited the only significant gender effects of the eight age groups; Figure 3 shows the extent of these gender differences.

For 40–45-year-old men, VOC scores increased significantly from when these men were 13.1 years younger; for men 64–69, 76–81, and 82–87 years old, VOC scores declined significantly from when the men were 13.1 years younger. For 76–81-year-old men, VOC scores also declined significantly from when the men were 19.0 and 25.1 years younger. With the exception of the 64–69-year-olds, we found that a significant longitudinal change for a shorter longitudinal period was always accompanied by significant longitudinal changes for the longer longitudinal periods. When these analyses were restricted to participants with BIMC scores of 2 or less, the outcomes were unchanged.

For the BVRT, no significant longitudinal changes occurred until the 64–69-year-old age-at-later-testing interval. Thereafter, all longitudinal changes were significant for all periods and for both men and women. Furthermore, longer periods tended to

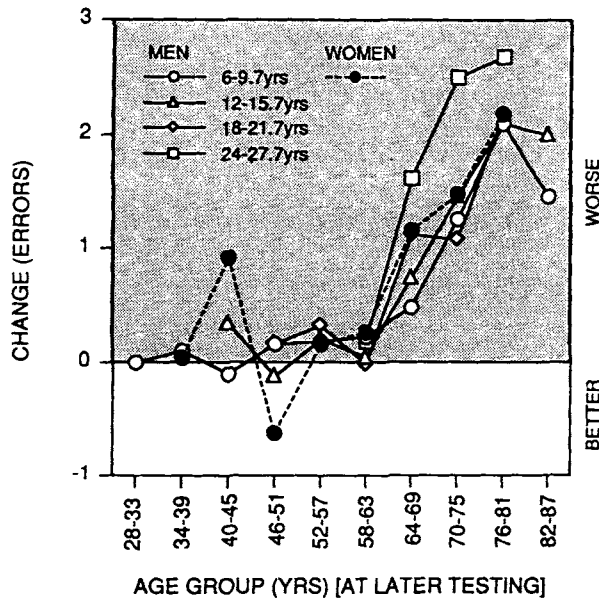


Figure 4. Mean changes (later minus earlier) in Benton Visual Retention Test total errors for the 5.95–9.7, 11.95–15.7, 17.95–21.7, and 23.95–27.7 longitudinal periods for 6-year age-at-later-testing intervals. Age at earlier testing can be derived by subtracting the mean longitudinal period from age at later testing.

Table 5  
Longitudinal Change on WAIS Vocabulary and Benton Visual Retention Test for Specific Maximum Sample Size Age-at-Later-Testing Intervals (ANOVA Results When Significant,  $p < .05$ , the Percentages of Variance Accounted for by Longitudinal Periods are Entered)

Age at later testing (years)	WAIS Vocabulary				Benton Visual Retention Test				
	5.95–9.7-year period		11.95–15.7-year period (men)		17.95–21.7-year period (men)		23.95–27.7-year period (men)		
	Women	Men	Women	Men	Women	Men	Women	Men	
28–33	—	22.5	—	—	—	—	—	—	—
34–39	ns	5.9	—	—	—	—	—	—	—
40–45	20.9	3.7	6.3	—	—	—	—	—	—
46–51	ns	ns	ns	ns	ns	ns	ns	ns	ns
52–57	33.4	ns	ns	ns	ns	ns	ns	ns	ns
58–63	12.0	ns	ns	ns	ns	ns	ns	ns	ns
64–69	12.2	ns	3.2	ns	5.9	14.8	12.7	27.6	62.1
70–75	ns	ns	ns	ns	13.0	12.7	12.7	27.6	62.1
76–81	ns	16.5	17.9	32.4	26.3	46.4	46.4	38.7	—
82–87	—	—	35.9	—	18.3	—	—	—	—

Note. Age at later testing was the participant's age on the later of the two testings included in the longitudinal interval. Dashes indicate a sample size of less than 15 (see Table 1). WAIS = Wechsler Adult Intelligence Scale. ns = Not significant,  $p > .05$ .

account for greater percentages of variance in longitudinal changes and to exhibit greater increases in errors (see Figure 4). When the just-described analyses were restricted to participants with BIMC scores of 2 or less, the outcomes were unchanged with one exception; the 82–87-year-olds now showed essentially no change in number of BVRT errors from 6.8 years earlier. The only significant gender difference occurred at the 40–45-year-old age-at-later-testing interval,  $F(1, 142) = 3.97, p < .05, \omega^2 = .020$ ; the women had a mean increase in errors of 0.92 from 6.8 years earlier, and the men had a mean decrease in errors of 0.10.

Recall that the order of BVRT forms was C, E, D, C, E; that both Forms E and D were more difficult than Form C; and that Form D was more difficult than Form E. For the age-at-later-testing maximized samples, this created a bias toward a longitudinal increase in errors. However, for every longitudinal period, such bias was diluted by oppositely biased comparisons or by same-form comparisons. For example, C to E, C to D, and E to D changes were biased toward increased error, whereas E to C, D to C, and D to E changes were biased toward decreased error. In the 19.0-year longitudinal period, 93.3% of the comparisons were of the same form at both testings (C to C or E to E). Recall further that different subsamples of participants were administered two forms at the second testing (E and then C), the third testing (D and then E), and the fourth testing (C and then D). Through the use of either the first or second administered form at these testings, some of the previously described ANOVAs could be restricted to same-form longitudinal change comparisons. When a minimum age-at-later-testing sample size of 15 was required, same-form comparisons were possible only within the 6.8- and 19.0-year longitudinal periods.

Virtually all of the women had Form C administered second at the second testing. In the 34–39, 40–45, 52–57, and 58–63 age groups, the longitudinal performance change was improvement (i.e., fewer errors were made 6.8 years later) when Form C was used at both testings in contrast to a longitudinal decline when Forms C and E were used. However, longitudinal change was not significant in any of these age groups. In the 46–51-year-old group, form had little impact on the 6.8-year longitudinal change in errors. In the 64–69 and 70–75 age groups, the longitudinal change was significant when Form C was used at both testings. In the first group, the decline was reduced relative to the comparisons of Forms C and E; in the second group, the decline was increased relative to these comparisons.

For the men in the 6.8-year longitudinal period, 23.9% could be matched to identical forms for the earlier and later testings. For the age-at-later-testing groups, this ranged from a low of 8% for the 70–75-year-olds to a high of 51% for the 34–39-year-olds. Sample sizes of fewer than 15 occurred in the 28–33 and 70–75 age groups. In five groups, the same-form comparisons resulted in a greater decrease in errors or a smaller increase in errors than with different-form comparisons; in two groups, an augmented increase in errors occurred; and, in one group, errors increased with same-form comparisons when different forms errors decreased. In the 40–45 and 46–51 age groups, a decrease in errors on the later testing was obtained with identical forms and was significantly different from zero; in both cases, the comparisons with different forms were not significant. In the 19.0-year longitudinal period, 93.3% of the participants could be matched to identical forms for the earlier and later testing with no essential change in the outcomes previously reported.

Change on the VOC and change on the BVRT were correlated within each age group for each longitudinal period. For the women, no correlation was significant at the .05 level. For the men, only 3 of 18 correlations were significant; 2 were negative ( $-.23$  and  $-.19$ ), and 1 was positive ( $.26$ ). It would appear that there was little, if any, covariation between changes on the VOC and the BVRT regardless of age at later testing and length of the longitudinal period.

*Tenure in the BLSA and longitudinal change.* We looked again for evidence that BLSA participants who remained in the program differed in ability from those who dropped out so that we could assess any bias in our measures of longitudinal change. Here we examined tenure effects for the two shorter longitudinal periods as indicative of possible bias for the two longer periods. For men within the 6.8- and 13.1-year longitudinal periods, each age-at-later-testing group was partitioned into two subgroups: those who had been tested two or three times and those who had been tested four or five times (see earlier explanation under “Maximizing age group representation”). This partition permitted an evaluation of longitudinal change as a function of extent of participation in the BLSA. For example, we determined whether longer term participants had smaller longitudinal changes than shorter term participants. For each age-at-later-testing group at the 6.8- and 13.1-year longitudinal periods, a Number of Testings (2 or 3 vs. 4 or 5)  $\times$  Repeat (earlier vs. later) ANOVA was carried out when a cell size was at least 15. The interaction term was of primary interest because longitudinal change was defined as later minus earlier performance.

On the VOC, significant main effects were found for number of testings in the 46–51-year group for the 6.8-year longitudinal period,  $F(1, 115) = 4.54, p < .05, \omega^2 = .025$ , and in the 52–57-year group for the 13.1-year longitudinal period,  $F(1, 100) = 5.29, p < .05, \omega^2 = .040$ . In both instances, the longer term participants had higher VOC scores than the shorter term participants. The longitudinal change was significantly affected by the number of testings in only one instance; VOC scores for the shorter term participants ( $-1.2$ ) declined more than those for the longer term participants ( $-0.2$ ) in the 70–75-year group for the 6.8-year longitudinal period,  $F(1, 88) = 4.81, p < .05, \omega^2 = .040$ . Across all age groups for the 6.8-year longitudinal period, the mean longitudinal changes (later minus earlier) were 0.0 for the shorter term participants and 0.5 for the longer term participants. For the 13.1-year longitudinal period, the means were  $-0.8$  and  $0.0$ , respectively, for the shorter and longer term participants. Thus, although there was some evidence that shorter term BLSA participants were more likely to show declines or smaller improvements than were longer term participants, the magnitude of the difference was quite small and seldom significant.

On the BVRT, a significant main effect for number of testings occurred only with the 76–81-year group for the 6.8-year longitudinal period,  $F(1, 73) = 6.28, p < .05, \omega^2 = .049$ . The shorter term participants made fewer errors than the longer term participants. The longitudinal change was significantly affected by number of testings (i.e., the interaction term) in the 46–51-year group for the 6.8-year longitudinal period,  $F(1, 154) = 8.00, p < .01, \omega^2 = .041$ , and in the 64–69-year group for the 13.1-year period,  $F(1, 95) = 5.14, p < .05, \omega^2 = .037$ . In both instances, the shorter term participants had a greater longitudinal increase in errors than the longer term participants (by 1.14 for the 46–

51-year-olds and by 1.57 for the 64–69-year-olds). For the 6.8-year longitudinal period across all age groups, the mean increases in errors were 0.81 for the shorter term participants and 0.39 for the longer term participants; for the 13.1-year period, the means were 0.92 and 0.68, respectively, for the shorter and longer term participants. Although there was some evidence that shorter term BLSA participants had greater longitudinal increases in BVRT errors than longer term participants, the magnitude of the difference was quite small and seldom significant.

*Individual variation in intraindividual change.* The preceding description of longitudinal change has been in terms of group central tendencies. In this section, we examine individual variation in these age changes. Appendix B gives the distribution of change scores, defined as later minus earlier scores, on the VOC for each of the age-at-later-testing groups in each of the longitudinal intervals; Appendix C provides the same distribution for the BVRT. These distributions permitted an assessment of individual variation in intraindividual change. Of particular interest were those individuals who exhibited change contrary to that indicated by significant central tendencies.

On the VOC, the significant mean longitudinal changes were positive—indicating improvement—for all age-at-later-testing groups younger than 64–69 years of age (see Table 5). For women, the 64–69-year group also had a positive longitudinal change. Twenty-nine percent, 32%, 23%, and 39% of women, respectively, in the 40–45, 52–57, 58–63, and 64–69 age groups had either no change or a negative change (indicating decrement). For men in the 6.8-year longitudinal period, 21%, 43%, and 47%, respectively, in the 28–33, 34–39, and 40–45 age groups had either no change or a negative change. For men in the 13.1-year longitudinal period, 42% of the 40–45-year-old group had either no change or a negative change. The significant mean longitudinal changes were negative for all age-at-later-testing groups 76 years old and older (see Table 5 and Figure 3). For the 76–81-year-old age-at-later-testing groups, 36%, 22%, and 25% of participants, respectively, in the 13.1-, 19.0-, and 25.1-year longitudinal periods had either no change or a positive change (indicating improvement) in the 82–87-year-old age-at-later-testing groups, the proportions who had either no change or a positive change were 39% and 24%, respectively, for the 6.8- and 13.1-year longitudinal periods. Thus, substantial minorities within all statistically significant change groups showed longitudinal changes counter to those indicated by the central tendencies. (In the young and middle-aged groups, there were individuals who did not improve from a testing at a younger age. In the young-old groups, there were individuals who did not decline from a testing at a younger age.) Appendix B provides the details of these contrary changes.

On the BVRT, the significant mean changes were all positive—indicating decrement—and occurred for all age-at-later-testing groups 64–69 years and older. In the 64–69-year-old groups, 37% of the women had either no change or improvement 6.8 years later; the values for men were 52%, 52%, 46%, and 37%, respectively, for the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods. In the 70–75-year-old groups, 34% of the women had either no change or an improvement 6.8 years later; the values were 44%, 34%, 40%, and 15%, respectively, for men in the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods. In the 76–81-year-old male groups, 34%, 35%, 22%, and 27%, re-

spectively, had either no change or improvement for the 6.8-, 13.1-, 19.0-, and 25.1-year longitudinal periods. In the 82–87-year-old male groups, 44% and 38%, respectively, had either no change or improvement for the 6.8- and 13.1-year longitudinal periods. Thus, with one exception, at least 22% of every age-at-later-testing group had included participants who did not decline from a testing from as many as 25.1 years earlier. Clearly, for substantial minorities of young-old individuals, immediate memory for designs was at least as good as in earlier middle age, and, for substantial minorities of old-old individuals, immediate memory for designs was at least as good as in their middle and late middle age. Appendix C provides the details of the structure of these minority changes. Furthermore, these percentages were, in all likelihood, underestimates for the 6.8-, 13.1-, and 25.1-year longitudinal periods, because there was some bias toward more difficult forms of the BVRT being given later than earlier (see earlier description of this bias).

## Discussion

Cross-sectionally, immediate visual memory began to show decrement in late middle age; from late middle to old age, the decrement accelerated. Longitudinally, intraindividual change did not reach a magnitude sufficient to be significant until after 64 years. The 65 to 70 age period is particularly important because cognitive decrements begin to appear (see also Schaie, 1994). In that period, significance change occurred from as little as 6.8 years earlier. Longer longitudinal periods tended to result in greater decremental changes in immediate visual memory, although only some of these differences were significant. Except for the 82–87-year-old group, the older a subject at later testing, the greater the magnitude of the decrement; this finding provides further evidence of a positively accelerating age function. Vocabulary showed increments to middle age, and a plateau occurred to early old age; very modest decrements were found to occur in old-old individuals. Furthermore, there was little indication that change in immediate visual memory and change in vocabulary covaried, regardless of the longitudinal interval and age at later testing. This outcome suggests that, in individuals with no evidence of neurodegenerative or neuropsychiatric disease, there was no generalized lowering in efficiency of the brain. Rather, efficiency decrements, if they occurred, were isolated or confined to specific brain regions.

If vocabulary is a valid marker for crystallized intelligence, then the evidence places crystallized intelligence alongside fluid intelligence as negatively affected by aging. Of course, both factors of intelligence do not change at the same rate, at the same ages, or to the same degree. Crystallized intelligence is weakly negatively affected somewhat later in life, the mid-70s, but more strongly affected in the late 70s and early 80s. Mean changes were 4 or fewer unadjusted points (equivalent to entirely forgetting two words or less adequately defining four words), and about 1 in 4 individuals showed no decrement in vocabulary from when they were in their early 50s. The health and mental status screening of our participants reduced substantially the likelihood that these late-life decrements were the result of a few poorly performing individuals with diseases. If disease is responsible for this late-life decline in crystallized intelligence, then it is occult and clearly subclinical. Before accepting that these late-life declines are the result of occult disease or a site-

specific area of neuronal degeneration, one should investigate other potential causes. For example, perhaps these declining individuals were not sufficiently motivated, or perhaps their retrieval from long-term memory was incomplete at that time or they were suffering from the effects of sleep disturbance common late in life. However, until these other potential causes are investigated, it seems that the current view of crystallized intelligence as entirely preserved throughout the life span must be revised. This conclusion is consistent with the results of 30 years of longitudinal study of primary mental abilities by Schaie (1994), who concluded that “—reliable average decrement is indeed found for all abilities by age 67” (p. 308).

Age accounted for 2 to greater than 10 times the amount of variance in immediate visual memory than in vocabulary, indicating that age has much more profound effects on immediate visual memory. There is evidence that short-term memory is tied to hippocampal functioning (see Mishkin, Malamut, & Bachevalier, 1984, for a review) and that loss of neurons in the hippocampus occurs with aging (see Coleman & Flood, 1987, for a review). However, our connection of neuronal loss in the hippocampus to the observed decrement in immediate visual memory must remain conjectural until appropriate prospective studies establish that link.

Slower reproduction times in older people represent another potential cause of an age-based decrement in immediate visual memory. It has been observed that memory span for lists of verbal material seems to be limited, on average, to that which can be spoken in 2 s (Cowan, 1993; Schweickert, 1993). It is inferred that items in immediate “verbal” memory decay to nonretrievability in about 2 s. This outcome could be summarized as follows: The longer a person takes to produce items from immediate memory, the fewer the number of items recalled. The Benton Visual Retention Test requires that each geometric design be reproduced on paper. Older adults are slower writers and figure tracers than younger adults (Dixon, Kurzman, & Friesen, 1993; Welford, 1977). If one can infer slower drawing from slower tracing, then poorer immediate memory with increased age could be simply a production artifact. This deduction, of course, is based on the truth of the unspoken premise that visual reproductive memory has decay characteristics quantitatively similar to those of memory for words. Further evidence supporting an age decrement in immediate visual memory to longer reproduction times comes from the digit-span task, which measures immediate memory and requires verbal reproduction. There is evidence for reduced speech rates (two thirds; Mysak & Hanley, 1958) and reduced subvocalization rates (three fourths; Salthouse, 1980) in old people relative to the young people. Robertson-Tchabo and Arenberg (1989) reported small decreases in digit span between the third and ninth decades (the old digit span was 90% that of the young digit span). These results suggest that age decrements in immediate memory for designs would be eliminated if some correction or control were made for slower reproduction times in the elderly. Conversely, Baddeley (1986) has provided evidence that age differences in working memory are primarily the result of impairments in the central executive component, as opposed to the articulatory loop and to a lesser extent the visual-spatial scratch pad, of his working memory model. Furthermore, Schaie (1994) has provided rather compelling supporting evidence that only maturation is responsible for the

mid-60s to early 70s being a watershed for cognitive processes in normal aging. By looking at a sample well screened for mental status as well as neurodegenerative and neuropsychiatric disease, we have provided considerable additional evidence for considering the mid-60s to early 70s as a watershed.

## References

- Arenberg, D. (1978). Differences and changes with age in the Benton Visual Retention Test. *Journal of Gerontology*, *33*, 534–540.
- Arenberg, D. (1982). Estimates of age changes on the Benton Visual Retention Test. *Journal of Gerontology*, *37*, 87–90.
- Arenberg, D. (1987). A note on differences and changes in memory with age. In M. W. Riley, J. D. Matarazzo, & A. Baum (Eds.), *Perspective on behavioral medicine: The aging dimension* (pp. 39–47). Hillsdale, NJ: Erlbaum.
- Arenberg, D. (1990). Longitudinal changes in cognitive performance. In R. J. Wurtman, S. Corkin, J. H. Growdon, & E. Ritter-Walker (Eds.), *Advances in neurology*, Vol. 51: *Alzheimer's disease* (pp. 207–209). New York: Raven Press.
- Baddeley, A. (1986). *Working memory*. New York: Oxford University Press.
- Benton, A. L. (1974). *The Revised Visual Retention Test: Clinical and experimental applications* (4th ed.). New York: Psychological Corporation.
- Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in cerebral gray matter of elderly subjects. *British Journal of Psychiatry*, *114*, 797–811.
- Botwinick, J., & Storandt, M. (1974). *Memory, related functions and age*. Springfield, IL: Charles C Thomas.
- Coleman, P. D., & Flood, D. G. (1987). Neuron numbers and dendritic extent in normal aging and Alzheimer's disease. *Neurobiology of Aging*, *8*, 521–545.
- Cowan, N. (1993). Activation, attention, and short-term memory. *Memory & Cognition*, *21*, 162–167.
- Dixon, R. A., Kurzman, D., & Friesen, I. C. (1993). Handwriting performance in younger and older adults: Age, familiarity, and practice effects. *Psychology and Aging*, *8*, 360–370.
- Gilbert, J. G. (1973). Thirty-five-year follow-up study of intellectual functioning. *Journal of Gerontology*, *28*, 68–72.
- Haaland, K. Y., Linn, R. T., Hunt, W. C., & Goodwin, J. S. (1983). A normative study of Russell's variant of the Wechsler Memory Scale in a healthy elderly population. *Journal of Consulting and Clinical Psychology*, *6*, 878–881.
- Horn, J. L. (1978). Human ability systems. In P. B. Baltes (Ed.), *Lifespan development and behavior* (pp. 212–257). New York: Academic Press.
- Kausler, D. H. (1991). *Experimental psychology, cognition, and human aging*. New York: Springer-Verlag.
- Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.
- Matarazzo, J. D. (1977). *Wechsler's measurement and appraisal of adult intelligence* (5th enlarged ed.). Baltimore: Williams & Wilkins.
- McCarty, S. M., Siegler, I. C., & Logue, P. E. (1982). Cross-sectional and longitudinal patterns of memory scores. *Journal of Gerontology*, *37*, 169–175.
- Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. L. McGaugh, & N. W. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 65–77). New York: Guilford Press.
- Mysak, E. D., & Hanley, T. D. (1958). Aging process in speech: Pitch and duration characteristics. *Journal of Gerontology*, *13*, 309–313.
- Robertson-Tchabo, E. A., & Arenberg, D. (1989). Assessment of memory in older adults. In T. Hunt & C. J. Lindley (Eds.), *Testing older*

adults (pp. 200-231). Kansas City, MO: Test Corporation of America.

Robertson-Tchabo, E. A., Arenberg, D., Tobin, J. D., & Plotz, J. (1986). A longitudinal study of cognitive performance in noninsulin dependent (Type II) diabetic men. *Experimental Gerontology*, 21, 459-467.

Salthouse, T. A. (1980). Age and memory: Strategies for localizing the loss. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging: Proceedings of the George A. Talland Memorial Conference* (pp. 47-66). Hillsdale, NJ: Erlbaum.

Schaie, K. W. (1983). The Seattle Longitudinal Study: A 21-year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 64-135). New York: Guilford Press.

Schaie, K. W. (1989). Perceptual speed in adulthood: Cross-sectional and longitudinal studies. *Psychology and Aging*, 4, 443-453.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist*, 49, 304-313.

Schaie, K. W., & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. *Psychology and Aging*, 8, 44-55.

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate memory. *Memory & Cognition*, 21, 168-175.

Shock, N. W., Greulich, R. C., Andres, R., Arenberg, D., Costa, P. T., Jr., Lakatta, E., & Tobin, J. D. (1984). *Normal human aging: The Baltimore Longitudinal Study of Aging* (U.S. Public Health Service Publication No. 84-2450). Washington, DC: U.S. Government Printing Office.

Terman, L. M., & Merrill, M. (1937). *Measuring intelligence: A guide to the new revised Stanford-Binet test of intelligence*. Boston: Houghton Mifflin.

Thurstone, L. L., & Thurstone, T. G. (1949). *Examiner manual for the SRA Primary Mental Abilities Test*. Chicago: Science Research Associates.

Wechsler, D. (1945). A standardized memory scale for clinical use. *Journal of Psychology*, 19, 87-95.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.

Welford, A. T. (1977). Motor performance. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 450-496). New York: Van Nostrand Reinhold.

Wilkie, F., & Eisdorfer, C. (1985). Intellectual changes. In E. Palmore, E. W. Busse, G. L. Maddox, J. B. Nowlin, & I. Siegler (Eds.), *Normal aging III: Reports from the Duke studies, 1975-1984* (pp. 190-195). Durham, NC: Duke University Press.

Appendix A

Age-at-Later-Testing Intervals: Wechsler Adult Intelligence Scale Vocabulary (Unadjusted Raw Score) and Benton Visual Retention Test (Total Errors) Means and Standard Deviations

Age at later testing (years)	Longitudinal period (years)									
	5.95-9.7				11.95-15.7 (men)		17.95-21.7 (men)		23.95-27.7 (men)	
	Women		Men		Early testing	Later testing	Early testing	Later testing	Early testing	Later testing
Vocabulary										
28-33										
M	—	—	61.1	64.4	—	—	—	—	—	—
SD	—	—	8.3	6.4	—	—	—	—	—	—
34-39										
M	62.5	64.0	62.8	64.0	—	—	—	—	—	—
SD	10.3	9.4	9.0	8.1	—	—	—	—	—	—
40-45										
M	60.1	62.9	62.4	63.5	63.2	64.5	—	—	—	—
SD	11.5	10.5	10.6	9.5	9.9	8.9	—	—	—	—
46-51										
M	65.2	66.1	64.3	64.7	62.5	63.2	63.9	65.5	—	—
SD	9.4	9.2	9.3	8.7	11.1	9.7	10.8	8.4	—	—
52-57										
M	60.8	63.9	64.5	64.8	64.9	65.3	64.0	64.8	—	—
SD	13.3	11.9	9.5	9.6	9.1	8.5	10.2	10.3	—	—
58-63										
M	62.5	63.9	66.3	66.3	64.9	65.6	66.1	66.3	63.9	64.8
SD	11.9	12.5	8.7	8.0	9.3	8.8	8.2	8.1	11.1	8.2
64-69										
M	63.3	65.1	66.9	66.4	67.8	66.9	66.5	66.8	66.3	67.8
SD	8.7	8.4	8.6	8.4	8.3	8.1	9.5	9.4	8.6	7.9
70-75										
M	63.6	64.0	66.5	66.1	66.2	65.3	68.0	66.9	66.1	65.8
SD	9.3	10.6	8.7	8.0	9.3	9.3	6.7	5.6	8.0	6.5

## Appendix A (continued)

Age at later testing (years)	Longitudinal period (years)									
	5.95-9.7				11.95-15.7 (men)		17.95-21.7 (men)		23.95-27.7 (men)	
	Women		Men		Early testing	Later testing	Early testing	Later testing	Early testing	Later testing
	Early testing	Later testing	Early testing	Later testing						
Vocabulary										
76-81										
<i>M</i>	66.3	65.6	66.7	65.7	68.0	65.6	67.6	64.6	70.8	67.1
<i>SD</i>	7.5	6.6	8.7	9.1	5.7	5.8	8.1	8.3	5.2	5.6
82-87										
<i>M</i>	—	—	66.4	64.0	68.9	64.6	—	—	—	—
<i>SD</i>	—	—	8.0	7.8	5.8	6.9	—	—	—	—
Benton Visual Retention Test										
28-33										
<i>M</i>	—	—	2.22	2.22	—	—	—	—	—	—
<i>SD</i>	—	—	1.65	1.93	—	—	—	—	—	—
34-39										
<i>M</i>	2.92	2.96	2.46	2.56	—	—	—	—	—	—
<i>SD</i>	2.19	2.10	2.24	1.94	—	—	—	—	—	—
40-45										
<i>M</i>	2.19	3.11	3.15	3.05	2.06	2.41	—	—	—	—
<i>SD</i>	1.55	2.28	2.12	2.56	1.73	1.84	—	—	—	—
46-51										
<i>M</i>	3.04	2.43	2.72	2.89	2.90	2.78	2.20	2.37	—	—
<i>SD</i>	2.13	2.23	2.24	2.44	2.29	2.43	1.91	1.44	—	—
52-57										
<i>M</i>	3.32	3.48	3.06	3.24	2.75	2.96	2.79	3.11	—	—
<i>SD</i>	2.51	2.02	2.37	2.38	2.19	2.16	2.29	2.19	—	—
58-63										
<i>M</i>	4.20	4.47	3.54	3.76	2.98	3.03	2.69	2.68	2.78	2.96
<i>SD</i>	1.85	2.80	2.48	2.61	2.33	2.01	2.25	2.00	2.22	2.12
64-69										
<i>M</i>	4.22	5.39	4.04	4.52	3.49	4.24	2.77	3.90	2.21	3.83
<i>SD</i>	2.33	3.31	2.62	2.82	2.58	3.09	2.09	2.92	1.52	2.78
70-75										
<i>M</i>	5.11	6.59	4.67	5.93	4.41	5.88	3.84	4.93	2.86	5.37
<i>SD</i>	2.55	3.05	2.51	2.96	2.87	3.21	2.49	2.35	2.16	2.06
76-81										
<i>M</i>	5.31	7.50	5.45	7.55	4.50	6.61	4.64	6.82	3.57	6.26
<i>SD</i>	2.52	3.50	3.35	3.62	2.44	3.39	2.31	2.71	2.86	3.35
82-87										
<i>M</i>	—	—	7.03	8.49	6.52	8.52	—	—	—	—
<i>SD</i>	—	—	4.22	4.36	3.53	4.32	—	—	—	—

Note. Age at later testing was the participant's age on the later of two testings included in the longitudinal interval. Dashes indicate a sample size of less than 15 (see Table 1).

Appendixes continue on next page



## Appendix B

## Wechsler Adult Intelligence Scale Vocabulary: Percentage Distribution of Change in Unadjusted Raw Scores at Separate Age-at-Later-Testing and Longitudinal Periods

Gender	Age at later testing (years)	Age at earlier testing (years)	Change (later minus earlier score)											
			-10--17	-7--9	-5--6	-3--4	-1--2	0	1-2	3-4	5-6	7-9	10-12	13-16
Men	28-33	21-26	4	0	0	13	0	4	22	26	0	17	9	4
Women	34-39	27-32			8	8	25	8	13	8	13	13	4	
Men		27-32		6	3	9	18	7	20	15	14	6	4	
Women	40-45	33-38		4	7	7	4	7	26	7	7	19	7	4
Men		33-38	3	5	6	6	14	13	17	12	14	5	5	2
Men		27-32		4	6	9	17	6	17	20	12	9	0	2
Women	46-51	39-44		7	0	11	21	11	14	14	18	4		
Men		39-44	1	6	6	15	15	9	18	11	11	5	3	1
Men		33-38	1	7	10	7	14	12	16	12	9	7	5	1
Men		27-32		11	14	17	20	3	17	14	12	9	6	3
Women	52-57	45-50				12	4	16	20	8	16	16	8	
Men		45-50	3	6	6	11	10	14	16	16	11	6	0	1
Men		39-44	3	4	10	14	14	7	14	17	8	7	4	
Men		33-38	4	5	0	16	11	18	13	12	11	7	2	2
Women	58-63	51-56		3	3	7	10	0	40	23	10	3		
Men		51-56	1	4	9	12	18	8	23	13	6	3	1	2
Men		45-50	2	4	5	10	17	12	16	11	12	8	3	
Men		39-44	1	7	10	13	14	9	13	14	5	9	1	1
Men		33-38	7	0	7	0	26	4	19	22	4	0	7	4
Women	64-69	57-62		6	0	8	11	14	19	6	22	11	3	
Men		57-62	4	8	6	18	15	8	18	9	8	6	1	
Men		51-56	3	8	7	13	18	11	22	8	5	1	2	
Men		45-50	3	3	7	17	14	8	13	16	14	5	1	
Men		39-44		3	7	10	17	7	20	10	10	3	10	
Women	70-75	63-68	4	7	26	0	19	14	4	22	4	11	11	4
Men		63-68		4	10	16	16	14	19	13	3	4		
Men		57-62		5	10	20	20	10	12	14	3	3		
Men		51-56	7	9	5	17	16	5	17	16	3	2	3	
Men		45-50	6	6	6	11	17	11	11	9	11	11		
Men	76-81	69-74	4	7	8	12	23	9	19	8	5	4	1	
Men		63-68	10	8	18	8	18	5	18	10	0	0	3	
Men		57-62	7	7	14	29	21	4	7	7	4			
Men		51-56	17	4	9	30	13	4	17	4				
Men	82-87	75-80	5	21	10	13	13	3	20	10	0	3	3	
Men		69-74	25	5	0	10	33	0	24					

Note. Distributions were limited to samples of greater than 20 individuals (see Table 1 for sample sizes).

## Appendix C

Benton Visual Retention Test: Percentage Distribution of Change in Total Errors  
for Each Age-at-Later-Testing Group and Longitudinal Period

Gender	Age at later testing (years)	Age at earlier testing (years)	Change (later minus earlier score)										
			-6--9	-3--5	-2	-1	0	1	2	3-4	5-6	7-8	9-13
Men	28-33	21-26		13	19	17	22	17	13	4	4		
Women	34-39	27-32	4	8	8	13	17	25	13	13			
Men		27-32	3	3	12	21	25	15	10	10	2	1	
Women	40-45	33-38		4	11	19	15	7	22	15	4	4	
Men		33-38	1	14	12	20	19	10	9	13	2	1	
Men		27-32		9	7	13	32	15	11	8	6		
Men	46-51	39-44		21	21	21	4	7	14	7	4		
Men		39-44		12	10	24	12	14	10	14	5		
Men		33-38	3	10	10	21	17	16	12	8	4		
Men		27-32	3	3	11	11	29	14	23	6			
Women	52-57	45-50	4	16	4	12	12	12	16	20	4		
Men		45-50	3	6	12	16	21	17	10	12	2	1	
Men		39-44	2	12	7	13	20	21	12	11	4		
Men		33-38	2	7	2	18	32	9	18	14			
Women	58-63	51-56		6	10	13	33	17	7	7	7		
Men		51-56	1	12	15	14	15	12	15	15	2	0	1
Men		45-50	1	10	15	15	13	21	8	13	3	1	
Men		39-44		12	14	16	21	17	8	12	1		
Men		33-38	4	7	7	4	30	26	11	7	4		
Women	64-69	57-62		11	3	6	17	19	19	16	8		
Men		57-62	2	9	10	15	16	15	13	11	6	2	1
Men		51-56	2	7	8	14	21	8	14	18	4	2	1
Men		45-50		8	4	9	25	17	8	13	17		
Men		39-44		3	0	17	17	14	14	25	7	3	
Women	70-75	63-68		4	19	7	4	11	19	19	19		
Men		63-68	1	9	4	9	21	13	10	16	16	1	
Men		57-62	3	9	3	10	9	17	14	19	8	5	2
Men		51-56		12	5	9	14	16	17	17	9	2	
Men		45-50		6	3	6	9	26	37	14			
Men	76-81	69-74		9	5	9	11	9	7	25	15	7	3
Men		63-68		5	11	8	11	13	16	13	8	13	3
Men		57-62		4	0	7	11	7	29	32	7	4	
Men		51-56		9	9	9	9	17	9	17	22	0	8
Men	82-87	75-80	3	8	5	13	15	13	13	10	8	8	5
Men		69-74	5	5	14	0	14	10	10	14	14	14	

Note. Distributions were restricted to samples of greater than 20 individuals (see Table 1 for sample sizes).

Received January 13, 1993  
Revision received June 13, 1994  
Accepted July 3, 1994 ■