*Archives of Clinical Neuropsychology*, Vol. 6, pp. 287–300, 1991
Printed in the USA. 
0887-6177/91 $3.00 + .00


# First–Last Names and the Grocery List Selective Reminding Test: Two Computerized Measures of Everyday Verbal Learning

**James R. Youngjohn**

*Memory Assessment Clinics, Scottsdale, Arizona*

**Glenn J. Larrabee**

*Memory Assessment Clinics, Sarasota, Florida*

**Thomas H. Crook, III**

*Memory Assessment Clinics, Bethesda, Maryland*

*Data are presented on two computerized tests of everyday verbal learning: Paired associate learning of First–Last Names (FLN), and the Grocery List Selective Reminding Test (GLSRT). MANOVA and multiple regression analyses demonstrated that performance on FLN and GLSRT was most strongly related to age, with significant secondary associations found for gender, with females performing better than males. Additional factor analysis of FLN and GLSRT supported the construct validity of these measures by demonstrating significant associations of performance with traditional neuropsychological measures of memory and related functions, including the Paired Associate Learning and Logical Memory subtests from the Wechsler Memory Scale, the Benton Visual Retention Test, and WAIS Digit Symbol.*

Evaluation of memory processes has undergone significant change across a number of dimensions in recent years. Global indices such as the Memory Quotient (MQ) (Wechsler, 1945) have been replaced with multiple-component

Requests for reprints should be sent to James R. Youngjohn, Ph.D., Memory Assessment Clinics, Inc., 7125 East Lincoln Drive, Suite A202, Scottsdale, AZ 85253.

approaches emphasizing assessment of orientation, attention, verbal memory, visual memory, and remote memory processes (Erickson & Scott, 1977; Kaszniak, Poon, & Riege, 1986; Larrabee, 1987). Moreover, there has been a shift away from unitary scores of verbal or visual memory towards measures which reflect multiple processes of memory function, such as the California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987), and the selective reminding procedure (Buschke, 1973). Additionally, there has been an increasing emphasis in neuropsychology on ecologic relevance of memory and cognitive tasks (Cunningham, 1986).

Four basic approaches have been used in the clinical evaluation of verbal learning and memory: text recall tasks such as Logical Memory from the Wechsler Memory Scale (WMS) and its recent revision (WMS-R; Wechsler, 1945; 1987); supraspan list learning of numbers or words, with repetition of all of the test stimuli prior to each recall attempt, such as the Rey Auditory Verbal Learning Test (Rey, 1964; Lezak, 1983) and the CVLT (Delis, Kramer, Kaplan, & Ober, 1987); supraspan word lists where subjects are selectively reminded of those words not recalled on previous trials prior to each new recall attempt (Buschke, 1973; Larrabee, Trahan, Curtiss, & Levin, 1988b); and paired associate paradigms (Wechsler, 1945, 1987).

Of the aforementioned examples, only the CVLT attempts to use ecologically relevant tasks (grocery list recall). In the present paper, data are presented on two new computer-administered verbal learning tasks, First–Last Names (FLN) and Grocery List Selective Reminding (GLSRT). First, the traditional selective reminding and paired associate paradigms will be considered in greater detail, followed by discussion of the development of FLN and GLSRT.

The selective reminding paradigm (Buschke, 1973) differs from earlier verbal learning procedures, in that the entire list is repeated by the examiner only prior to the first recall attempt. Prior to subsequent recall attempts, the examiner repeats only those words not recalled on the preceding trial. Words that have been recalled on two consecutive trials are considered to have entered long-term storage (LTS), since they were recalled without a reminder. Words that are consistently recalled on each subsequent trial without reminders are considered to have entered consistent long-term retrieval (CLTR).

The selective reminding procedure provides multiple scores, allowing for a more detailed measurement of verbal learning, recall, and retrieval, and thus addresses the frequently noted criticism that past verbal memory measures have been too reductionistic, collapsing across the many levels of an individual's performance (Ruff, Light, & Quayhagen, 1989; Parsons & Prigatano, 1978; Lezak, 1983; Russell, 1981). Indeed, preliminary factor analytic studies suggest that the different measures yielded by the selective reminding procedure provide information regarding different components of the learning process (Larrabee & Levin, 1986; Perrine, Novelly, Westerveld, Spencer, & Sass, 1989).

Other advantages of the selective reminding procedure have been the collection of fairly extensive normative data (Banks, Dickson, & Plasay, 1987; Larrabee, Trahan, Curtiss, & Levin, 1988b; Ruff, Light, & Quayhagen, 1989) and the development of multiple forms (Hannay & Levin, 1985). These advantages have led to the use of the selective reminding procedure as a preferred methodology for the measurement of verbal learning processes in general (MacLeod, 1985), and the treatment effects of candidate pharmaceutical compounds for memory dysfunction associated with aging in particular (Peters & Levin, 1977, 1979; Stern, Sano, & Mayeux, 1987; Thal & Fuld, 1983; Thal, Masur, Blau, Fuld, & Klauber, 1989; Wettstein, 1983).

Although the selective reminding procedure represents a major advance in terms of verbal learning assessment methodology, it might still be criticized for its limited relevance to the individual's day to day existence, i.e., ecologic validity, especially since the test forms typically employed contain lists of logically unrelated words (Hannay & Levin, 1985). Two verbal learning tests developed to have greater face validity are the Shopping List Task (McCarthy, Ferris, Clark, & Crook, 1981) and the CVLT (Delis, Kramer, Kaplan, & Ober, 1987). Both tests employ grocery items as task stimuli.

In its original form, the Shopping List Task was administered in the standard format with the entire list repeated prior to each recall attempt. It was subsequently revised in order to accommodate the selective reminding procedure (Ferris & Crook, 1983). Since this task was designed for use in pharmacologic research, it is available in multiple forms.

Perhaps one of the most popular and time-honored tests of verbal learning is Paired Associate Learning from the Wechsler Memory Scale (Wechsler, 1945). On this test, a list of paired words is presented to the subject over three trials. During the recall phase, the subject is required to give the second word of each pair after the examiner gives the first. The pairs are divided into six with a strong semantic relationship (easy pairs) and four with limited or no semantic relationship (hard pairs). The variability in performance on Paired Associate Learning is almost exclusively related to the hard word pairs (Halperin, Zeitchik, Healy, Weinstein, & Ludman, 1987).

The First–Last Names Task (Ferris & Crook, 1983) was developed to provide a more face valid and ecologically valid version of the Paired Associate Learning paradigm for use in geriatric psychopharmacological research. It requires individuals to learn a list of first–last name pairs and then provide the first name that went with each last name.

The original version of the Shopping List Test and the First–Last Names Task have been modified and incorporated into a computerized everyday memory battery (Larrabee & Crook, 1988a; Larrabee & Crook, 1989). This battery combines improvements in face and ecologic validity with the advances in standardization, reliability, and ease of data storage and analysis that are possible in computerized assessment. In its initial form, the battery

placed greater emphasis on everyday visual memory, with measures including facial memory and object location recall (Crook & Larrabee, 1988; Crook, Salama, & Gobert, 1986; Crook, Youngjohn, & Larrabee, 1990). Recently, two measures assessing everyday verbal learning have been added to the battery: a 15-item grocery list presented over five trials using the selective reminding paradigm, based on the Shopping List Task (GLSRT; McCarthy, Ferris, Clark, & Crook, 1981), and six first–last name pairs (FLN) presented over five trials using the paired associate learning format, based on the First–Last Names Task (Ferris & Crook, 1983). Initial factor analytic data demonstrate that these measures do indeed load on an everyday verbal memory factor (Larrabee & Crook, 1989).

The present study examines the performance of a sample of over 1500 participants, grouped by age (18–39, 40–49, 50–59, 60–69, and 70+) and gender, on the GLSRT and FLN. Performance on the GLSRT and FLN is analyzed in relation to age, gender, years of education, the raw score of the Vocabulary subtest from the Wechsler Adult Intelligence Scale (WAIS) (Wechsler, 1955), and the Geriatric Depression Scale (GDS; Yesavage et al. 1983).

Also, performance on the GLSRT and FLN is compared to performance on a number of traditional neuropsychological tests, including Logical Memory and Paired Associate Learning from the Wechsler Memory Scale (Wechsler, 1945), the Trail-Making Tests A & B, Digit Symbol from the WAIS (Wechsler, 1955), the Controlled Oral Word Association Test from the Benton Multilingual Aphasia Examination (Benton & Hamsher, 1976), and the Benton Visual Retention Test (Benton, 1974). These data were collected in a separate sample as baseline measures for a study of a candidate pharmaceutical compound for the treatment of Age-Associated Memory Impairment (AAMI), and are employed in the present study to replicate and expand the construct validity data reported on the GLSRT and FLN by Larrabee and Crook (1989).

## METHODS

### *Participants*

One thousand five hundred and thirty-five normal volunteers between the ages of 17 and 92 (mean = 43.39 years, *SD* = 21.66) participated in the initial study. The sample was recruited in the Washington, D.C. and Boulder, Colorado metropolitan areas. The entire group consisted of 689 males and 845 females. They were well educated (mean = 14.76 years of education, *SD* = 2.88) and well above average on the WAIS Vocabulary subtest (mean raw score = 59.22, *SD* = 10.50). All participants were carefully screened, and those with evidence or history of physical, psychiatric, or neurologic condi-

tions that could affect memory (e.g., depression, head trauma, or stroke) were excluded from the study.

The drug study sample on which the traditional neuropsychological measures were collected consisted of 219 individuals over age 50 (mean age = 62.25, *SD* = 7.79). There were 88 males and 131 females in the group. All participants met the inclusion and exclusion criteria for AAMI (Crook, Bartus, Ferris, Whitehouse, Cohen, & Gershon, 1986). Besides being at least 50 years old, they had subjectively noticed a decline in memory relative to their younger adult years, and met the following psychometric criteria: (i) at least one performance that was at least 1 *SD* below the mean for young adults (20–29 years old) on the Benton Visual Retention Test, Logical Memory from the WMS, or Paired Associate Learning from the WMS (cutoffs of 6 or less, 6 or less, and 13 or less, respectively); (ii) a raw score of at least 32 on the WAIS Vocabulary subtest; (iii) a score of 24 or higher on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) to exclude dementia; (iv) a Hachinski Ischemia Score (Rosen, Terry, Fuld, Katzman, & Peck, 1980) less than 4 (to exclude multi-infarct dementia); and (v) a Hamilton Depression Rating Scale (Hamilton, 1967) score of 12 or less (to exclude depression). In accordance with the criteria for AAMI, no other coexisting medical, psychiatric, or neurologic condition accounted for memory change. This group was also well educated (mean years of education = 14.74, *SD* = 3.00), and above average on the WAIS Vocabulary subtest (mean raw score = 61.71, *SD* = 10.86). Because this was part of a larger double-blind placebo-controlled investigation on the effects of a pharmaceutical compound on AAMI, only the baseline (drug and placebo free) data were analyzed for the present study.

### *Apparatus*

The GLSRT and FLN are administered using a Sony 19″ PVM 1910 color monitor interfaced with an AT&T 6300 computer equipped with a 20-megabyte hard-disk drive and customized computer graphics hardware. The tester is present throughout the session and sits behind and generally out of view of the participant. All test responses are made verbally and recorded by the examiner on a separate monitor, out of sight of the participant.

### *Procedure*

The FLN requires participants to read aloud a series of six first–last names, and then provide the first names that go with each of the last names. The order of presentation is different for each of the five trials. The GLSRT consists of a list of 15 supermarket food items that are presented over five trials in the selective reminding format. Delayed Recall of the grocery list is then tested after 30 minutes.

**TABLE 1**
**First–Last Name Recall per Trial and Sum Recall Across Trials by Age Group and Gender**

| Age Group (*N*) | | Trial 1 | Trial 2 | Trial 3 | Trial 4 | Trial 5 | Sum Recall Across Trials |
|---|---|---|---|---|---|---|---|
| 18–39 | Mean | 1.96 | 3.82 | 4.69 | 5.18 | 5.45 | 21.10 |
| (683) | *SD* | 1.35 | 1.54 | 1.49 | 1.32 | 1.14 | 5.66 |
| 40–49 | Mean | 1.24 | 2.96 | 3.67 | 4.18 | 4.63 | 16.68 |
| (92) | *SD* | 1.03 | 1.54 | 1.60 | 1.69 | 1.70 | 6.30 |
| 50–59 | Mean | 1.25 | 2.91 | 3.35 | 3.89 | 4.32 | 15.71 |
| (237) | *SD* | 1.00 | 1.60 | 1.81 | 1.80 | 1.69 | 6.75 |
| 60–69 | Mean | 1.13 | 2.67 | 3.15 | 3.59 | 4.14 | 14.69 |
| (334) | *SD* | 1.02 | 1.59 | 1.77 | 1.83 | 1.71 | 6.86 |
| 70+ | Mean | .82 | 2.17 | 2.50 | 3.11 | 3.50 | 12.03 |
| (189) | *SD* | .90 | 1.32 | 1.55 | 1.66 | 1.73 | 6.01 |
| Males | Mean | 1.42 | 3.00 | 3.65 | 4.16 | 4.55 | 16.77 |
| (689) | *SD* | 1.16 | 1.58 | 1.81 | 1.83 | 1.72 | 7.06 |
| Females | Mean | 1.54 | 3.30 | 3.96 | 4.45 | 4.83 | 18.08 |
| (845) | *SD* | 1.30 | 1.70 | 1.82 | 1.72 | 1.57 | 7.03 |
| Entire Population | | | | | | | |
| | Mean | 1.48 | 3.17 | 3.82 | 4.32 | 4.70 | 17.49 |
| (1534) | *SD* | 1.24 | 1.66 | 1.82 | 1.78 | 1.65 | 7.07 |

## RESULTS

### *Study 1: First–Last Names*

Table 1 depicts the performance of 1534 normal participants on FLN. A repeated measures MANOVA demonstrated a significant within-subject effect for trial (Hotellings T = 2.65, approximate $F(4, 1521) = 1006.92$, $p < .0001$). There also were significant effects for age group by trial (Hotellings T = .088, approximate $F(16, 6078) = 8.36$, $p < .0001$) and gender by trial (Hotellings T = .009, approximate $F(4, 1521) = 3.25$, $p < .012$). The age group by gender by trial effect was not significant ($p < .55$). The significant age group by trial interaction effects for FLN suggest that not only did the older age groups perform worse on each trial in an absolute sense, they learned across trials at a slower rate than the younger age groups.

A two-way (age group by gender) ANOVA was then performed examining the sum of first–last name pairs recalled across all trials. There were significant main effects for age group ($F(4, 1524) = 121.91$, $p < .0001$) and gender ($F(1, 1524) = 28.97$, $p < .0001$). The age-by-gender interaction effect was not significant ($p < .72$). Post hoc Tukey-B analyses revealed that both the 18–39 age group and the 70+ age group were significantly different from all other

age groups. Additionally, the 60–69 age group was significantly different from the 40–49 age group.

The relationship of FLN to a number of subject variables in addition to age and gender, including years of education, WAIS Vocabulary raw score, and GDS score was examined. A stepwise multiple regression analysis was conducted examining total recall summed across all trials (see Table 2). Age entered first as the most powerful predictor, followed by Vocabulary, then followed by gender. Education and GDS did not add to the variance accounted for by the equation at the .05 alpha level.

*Study 2: Grocery List Selective Reminding Test*

A repeated measures MANOVA was performed examining each of three scores (Total Recall per trial, LTS, and CLTR) on each of the five learning trials of the GLSRT. The within subject effect for trial was significant (Hotelling's T = 3.452, approximate $F(12, 17774) = 1704.54$, $p < .0001$), as were the interaction effects for age group by trial (Hotelling's T = .077, approximate $F(48, 17774) = 9.55$, $p < .0001$), gender by trial (Hotelling's T = .006, approximate $F(12, 17774) = 2.93$, $p < .0001$), and three-way interaction effect for age group by gender by trial (Hotelling's T = .027, approximate $F(48, 17774) = 3.32$, $p < .0001$). Figure 1 graphically presents CLTR on trials 1 through 5, for age group. The significant age-group-by-trial interaction effects for the GLSRT suggest that the older age groups learned across trials at a slower rate than the younger age groups, similar to data on FLN.

**TABLE 2**
**Order of Selection and Incremental $R^2$ of Demographic Variables as Predictors of First–Last Names and Grocery List Selective Reminding Test in Stepwise Multiple Regression Analyses**

| Factor Variables | Age | | Gender | | Education | | Vocabulary | | GDS Score | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Selection Order | ($R^2$) | Selection Order | ($R^2$) | Selection Order | ($R^2$) | Selection Order | ($R^2$) | Selection Order | ($R^2$) |
| FLN Sum Recall | 1 | (.243) | 3 | (.282) | * | | 2 | (.264) | * | |
| GLSRT Total Recall | 1 | (.198) | 2 | (.279) | * | | 3 | (.291) | * | |
| GLSRT Long Term Store | 1 | (.172) | 2 | (.246) | * | | 3 | (.255) | * | |
| GLSRT Consistent Long Term Recall | 1 | (.168) | 2 | (.261) | * | | 3 | (.278) | * | |
| GLSRT Delayed Recall | 1 | (.152) | 2 | (.219) | 3 | (.225) | * | | * | |

*$p > .05$ and predictor variables therefore not selected in the stepwise multiple regression analysis.
$R^2$ includes percentage of variance accounted for by that and all previously selected predictor variables.

**TABLE 3**
**Grocery Selective Reminding Test Mean Summary Scores for Age Group and Gender**

| Age Group (*N*) | | Total Recall | Long-Term Storage | Consistent Long-Term Recall | Delayed Recall |
|---|---|---|---|---|---|
| 18–39 | Mean | 60.56 | 57.31 | 47.28 | 13.03 |
| (653) | *SD* | 7.59 | 11.39 | 15.75 | 2.13 |
| 40–49 | Mean | 57.44 | 54.92 | 38.65 | 12.33 |
| (86) | *SD* | 6.72 | 10.98 | 18.07 | 3.59 |
| 50–59 | Mean | 55.59 | 51.64 | 36.03 | 12.33 |
| (229) | *SD* | 8.73 | 13.01 | 17.07 | 2.59 |
| 60–69 | Mean | 52.10 | 45.76 | 31.92 | 11.14 |
| (328) | *SD* | 10.07 | 15.30 | 16.83 | 3.32 |
| 70 + | Mean | 46.46 | 38.73 | 24.21 | 9.82 |
| (196) | *SD* | 12.71 | 17.43 | 17.08 | 3.61 |
| Males | Mean | 53.35 | 47.06 | 34.76 | 11.27 |
| (682) | *SD* | 10.33 | 15.44 | 17.10 | 3.10 |
| Females | Mean | 58.06 | 54.91 | 41.92 | 12.68 |
| (810) | *SD* | 9.89 | 13.61 | 19.03 | 2.82 |
| Entire Population | Mean | 55.91 | 51.32 | 38.65 | 12.03 |
| (1492) | *SD* | 10.36 | 14.99 | 18.52 | 3.03 |

An additional age group by gender MANOVA was performed, this time looking at Total Recall, LTS, and CLTR, collapsed across trials (see Table 3). There were significant main effects for age group (Hotelling's T = .364, approximate $F(12, 4436) = 44.85$, $p < .0001$) and gender (Hotelling's T = .070, approximate $F(3, 1480) = 34.36$, $p < .0001$). The interaction effect was not significant ($p < .228$).

Univariate comparisons of means were performed examining Total Recall, LTS, and CLTR summed across all five learning trials. There were significant age group effects for Total Recall ($F(4, 1491) = 110.25$, $p < .0001$), LTS ($F(4, 1491) = 90.90$, $p < .0001$), and CLTR ($F(4, 1491) = 97.06$, $p < .0001$). Post hoc comparisons (Tukey-B) demonstrated significantly decreased performance on all three measures from each age group to the next, with the exception of nonsignificant differences between the 40–49 and 50–59 year-old age groups (see Table 3).

Univariate comparisons were also performed for gender effects. Females performed significantly better than males on Total Recall ($F(1, 1491) = 80.74$, $p < .0001$), LTS ($F(1, 1491) = 108.84$, $p < .0001$), and CLTR ($F(1, 1491) = 57.61$, $p < .0001$; see Table 3).

An age-group-by-gender ANOVA was performed examining performance on the delayed recall trial of the GLSRT. There were significant main effects for age group ($F(4, 1454) = 64.47$, $p < .0001$) and gender ($F(1, 1454) = 102.03$, $p < .0001$). The interaction effect was marginally significant ($F(4, 1454) = 2.41$,

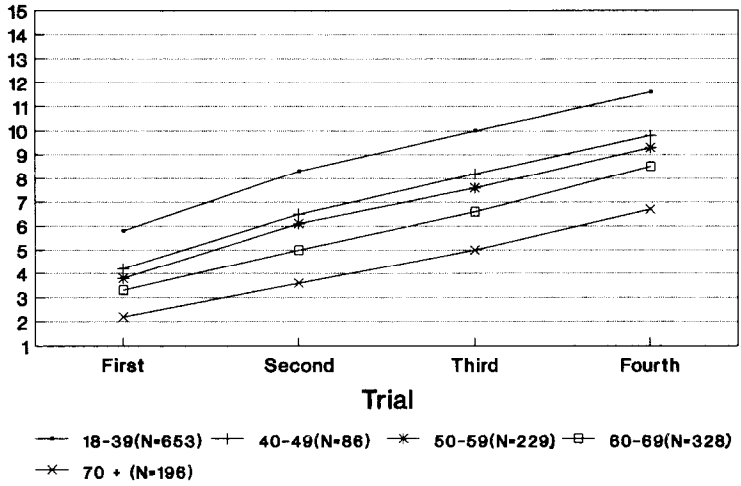


**FIGURE 1. Consistent long-term recall across learning trials 1–4 by age group.V.**

$p < .05$). post hoc analysis (Tukey-B procedure) demonstrated that the 60–69 and 70+ year olds performed significantly worse than all other age groups and that the 70+ age group did significantly worse than the 60–69 year olds. Additionally, the 50–59 age group did significantly worse than the 18–39 age group. With respect to gender effects, females were again superior to males (see Table 3).

A series of stepwise, multiple regression analyses were performed to examine the relationships between the additional subject variables and GLSRT performance (Table 2). Age was consistently selected first across the four GLSRT scores, followed by gender, as being the first and second most powerful predictors of GLSRT performance. Vocabulary was selected as the third most powerful predictor variable for GLSRT Total Recall, LTS, and CLTR. Education and GDS score were not selected as accounting for a significant level of the variance in performance on these three measures. For GLSRT Delayed Recall, education was selected as the third most powerful predictor, with Vocabulary and GDS not significantly contributing to the accuracy of the equation.

*Study 3: Relationship of FLN and GLSRT to Standard Tests*

In the third phase of our investigation we examined the relationship that performance on FLN and the GLSRT has to performance on the traditional neuropsychological measures. Total correct scores for FLN and GLSRT were included in a principal factor analysis with total correct WMS Easy and Hard Paired Associates, Logical Memory ((A+B)/2), WAIS raw scores for Vocabulary and Digit Symbol, total correct Visual Retention Test, number of words correct on Controlled Oral Word Association, and total time for

Trailmaking A and B. The first three eigen values (principal components) were 4.30, 1.44, and 1.15, accounting for 62.6% of the variance, suggesting that a three-factor solution was appropriate.

The three-factor, varimax-rotated principal axis solution is depicted in Table 4. The first factor is a verbal learning and memory factor, with high loadings from FLN, GLSRT, and WMS Hard Associates. The second factor, with high loadings from WAIS Digit Symbol and Trailmaking A and B, represents a dimension of attention/concentration/psychomotor speed. The third factor, with a high loading from WAIS Vocabulary, represents a verbal intelligence factor.

These data show that FLN and GLSRT both measure verbal learning and memory. For FLN, 87.8% of its communality is associated with verbal learning and memory, while 9.8% is attributable to the attentional factor, and 3.5% to the verbal intellectual factor. For GLSRT, 83.3% of common variance is attributable to verbal learning and memory, 16.8% to the attentional factor, and 0.4% to verbal intellectual ability. The GLSRT communality of .73 suggests that it shares more in common with traditional tests, than does the FLN measure, with a lower communality of .41.

## DISCUSSION

Both FLN and the GLSRT were demonstrated to be highly sensitive to age-related decline of verbal learning processes. These findings are consistent with the age-related declines found by previous investigators on the traditional analogues of FLN (Paired Associate Learning; Kausler, 1982; Trahan, Larrabee,

**TABLE 4**
**Factor Analysis Of First-Last Names, Grocery List Selective Reminding Test and Standard Neuropsychological Tests**

| Variables[a] | Factors 1 | 2 | 3 | $h^2$ |
|---|---|---|---|---|
| FLN | .60 | .20 | .12 | .41 |
| GLSRT | .78 | .35 | .04 | .73 |
| WMS Easy Associates | .52 | .22 | .02 | .32 |
| WMS Hard Associates | .72 | .11 | .17 | .56 |
| WMS Logical Memory | .52 | .14 | .32 | .39 |
| WAIS Vocabulary | .18 | .08 | .97 | .98 |
| Controlled Oral Word Association | .14 | .40 | .37 | .32 |
| Benton Visual Retention | .27 | .39 | .14 | .24 |
| WAIS Digit Symbol | .23 | .66 | .07 | .49 |
| Trailmaking A | -.15 | -.78 | -.08 | .64 |
| Trailmaking B | -.25 | -.78 | -.05 | .67 |

[a]See text for descriptions.

Quintana, Goethe, & Willingham, 1989) and the GLSRT (Selective Reminding Test; Banks, Dickson, & Plasay, 1987; Larrabee, Trahan, Curtiss, & Levin, 1988b; Ruff, Light, & Quayhagen, 1989). With respect to age effects in the recall of paired associate names, a previous study (Hulicka, 1967) found a significant age decline in the ability to recall names associated with occupations. Our findings of the sensitivity of FLN to age effects also replicate those of a previous study in which FLN was one of several variables measuring name recall in the elderly (Crook & West, in press). This sensitivity to decline with age occurred on a task with apparent improved face validity which Cunningham (1986) has noted to be important in enhancing older individuals' motivation and willingness to cooperate.

Although our results demonstrated age to be the strongest of the subject variables, the relationship between FLN and GLSRT performance and the subject variables other than age is also of interest. The superiority of females over males across all measures of both FLN and the GLSRT is generally consistent with the literature on verbal learning (Banks, Dickson, & Plasay, 1987; Bolla-Wilson & Bleecker, 1986; Larrabee, Trahan, Curtiss, & Levin, 1988; Ruff, Light, & Quayhagen, 1989). The effect appeared to be somewhat larger for the GLSRT (about 1/2 *SD* superiority of females over males) than for FLN (about 1/3 *SD* difference). This minor discrepancy might be hypothesized to be in part a result of the sex-role stereotyped nature of the GLSRT task (i.e., a grocery shopping list).

Vocabulary generally had a stronger relationship to FLN and GLSRT performance than education, with the exception of Delayed GLSRT Recall. The high percentage of shared variance between education and Vocabulary raises the possibility that there may have been significant effects for education for the other GLSRT and FLN test scores, but that these were moderated when Vocabulary was entered into the equation. This hypothesis was tested in a series of stepwise multiple-regression analyses examining only the effects of age and education on the summary scores of the GLSRT and FLN, with age entering the equation first. The hypothesis was supported for Sum Recall of FLN, where the addition of education accounted for a significantly increased percentage of the variance accounted for by the equation. It was not supported for Total Recall, LTS, or CLTR of the GLSRT, where education was not selected for the equation as being a significant predictor variable at the .05 alpha level. The generally weak effects for education are consistent with those of Larrabee et al. (1988), who did not find significant education effects on the Selective Reminding Test. However, our findings of significant but modest effects for Vocabulary differ from those of Ruff et al. (1989), who did not find a significant relationship between WAIS VIQ and performance on Form I of Hannay and Levin's (1985) version of selective reminding.

The relationship of FLN and GLSRT to standard neuropsychological tests provides further support for the construct validity of these measures. Both

FLN and GLSRT had primary loadings on a verbal learning and memory factor, with limited secondary associations with attentional and verbal intellectual factors. This replicates and extends previous factorial validity data reported by Larrabee and Crook (1989). Of additional interest is the lower communality for FLN relative to GLSRT. This suggests that FLN shares less common variance with traditional neuropsychological tests than does the GLSRT.

In conclusion, the two computerized tests of everyday verbal Learning that we have presented, FLN and GLSRT, appear to have excellent discriminant validity, with respect to their sensitivity to age and gender effects, as well as substantial construct validity, with respect to their relationship to the older, more traditional measures of neuropsychological functioning. Their improved face and ecological validity address many of the criticisms that have been directed against older memory assessment procedures (Cunningham, 1986; Erickson & Scott, 1977; Mayes, 1986) yet they remain sensitive to age-associated declines in performance. They also address Mayes' (1986) recommendation for computerization of memory assessment. These tests are currently being used in combination with other computerized everyday memory tests to evaluate treatment effects in more than a dozen clinical trials at over 30 sites in the United States and Europe for evaluation of candidate pharmacologic compounds for treatment of Alzheimer-type dementia and AAMI (Crook et al., 1986). The GLSRT and FLN have been translated into six different languages and versions of the tests are now being used in Belgium, Denmark, Finland, England, France, Italy, and Sweden. Multiple forms of both tests are available.

## REFERENCES

Banks, P. G., Dickson, A. L., & Plasay, M. T. (1987). The verbal selective reminding test: Preliminary data for healthy elderly. *Experimental Aging Research*, **13**, 203–207.

Benton, A. L. (1974). *The revised visual retention test* (4th ed.). New York: Psychological Corporation.

Benton, A. L., & Hamsher, K. de S. (1978). *Multilingual Aphasia examination.* Iowa City: University of Iowa (*Manual*, revised, 1978).

Bolla-Wilson, K., & Bleecker, M. L. (1986). Influence of verbal intelligence, sex, age, and education on the Rey Auditory Verbal Learning Test. *Developmental Neuropsychology*, **2**(3), 203–211.

Buschke, H. (1973). Selective reminding for analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior*, **12**, 543–550.

Crook, T. H., Bartus, R. T., Ferris, S. H., Whitehouse, P., Cohen, G. D., & Gershon, S. (1986). Age-associated memory impairment: Proposed diagnostic criteria and measures of clinical change—Report of a National Institute of Mental Health Work Group. *Developmental Neuropsychology*, **2**, 261–276.

Crook, T. H., & Larrabee, G. J. (1988). Interrelationships among everyday memory tests: Stability of factor structure with age. *Neuropsychology*, **2**, 1–12.

Crook, T., Salama, M., & Gobert, J. (1986). A computerized test battery for detecting and assessing memory disorders. In A. Bes, J. Cahn, S. Hoyer, J. P. Marc-Vergenes, & H. M. Wisniewski (Eds.), *Senile dementias: Early detection.* (pp. 79–85). London: John Libbey Eurotext.

Crook, T. H., & West, R. L. (in press). Name recall performance across the adult life span. British Journal of Psychology.

Crook, T. H., Youngjohn, J. R., & Larrabee, G. J. (1990). The Misplaced Objects Test: A measure of everyday visual memory. *Journal of Clinical and Experimental Neuropsychology*. **12**(6), 817–831.

Cunningham, W. R. (1986). Psychometric perspectives: Validity and reliability. In L. Poon, T. Crook, K. Davis, C. Eisdorfer, B. Gurland, A. Kaszniak, & L. Thompson (Eds.), *Handbook for clinical memory assessment of older adults* (pp. 27–31). Washington, DC: American Psychological Association.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *The California Verbal Learning Test–Research edition*. New York: Psychological Corporation.

Erickson, R. C., & Scott, M. L. (1977). Clinical memory testing: A review. *Psychological Bulletin*, **84**, 1130–1149.

Ferris, S. H., & Crook, T. (1983). Cognitive assessment in mild to moderately severe dementia. In T. Crook, S. Ferris, & R. Bartus (Eds.), *Assessment in geriatric psychopharmacology* (pp. 177–186). New Canaan, CT: Mark Powley Associates.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12**, 189–198.

Halperin, J. M., Zeitchik, E., Healy, J. M., Weinstein, L., & Ludman, W. L. (1987, February). *The development of linguistic and mnestic abilities in school-age children*. Paper presented at the 15th Annual Meeting of the International Neuropsychological Society, Washington, DC.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, **6**, 278–296.

Hannay, H. J., & Levin, H. S. (1985). Selective Reminding Test: An examination of the equivalence of four forms. *Journal of Clinical and Experimental Neuropsychology*, **7**, 251–263.

Hulicka, I. M. (1967). Age differences in retention as a function of interference. *Journal of Gerontology*, **22**, 180–184.

Kaszniak, A. W., Poon, L. W., & Riege, W. (1986). Assessing memory deficits: An information processing approach. In L. Poon, T. Crook, K. Davis, C. Eisdorfer, B. Gurland, A. Kaszniak, & L. Thompson (Eds.), *Handbook for clinical memory assessment of older adults* (pp. 27–31). Washington, DC: American Psychological Association.

Kausler, D. H. (1982). *Experimental psychology and human aging*. New York: Wiley.

Larrabee, G. J., (1987). Further cautions in interpretations of comparisons between the WAIS-R and the Wechsler Memory Scale. *Journal of Clinical and Experimental Neuropsychology*, **9**(4), 456–460.

Larrabee, G. J. & Crook, T. H. (1988a). A computerized everyday memory battery for assessing treatment effects. *Psychopharmacology Bulletin*, **24**(4), 695–697.

Larrabee, G. J., & Crook, T. H. (1989). Dimensions of everyday memory in Age-Associated Memory Impairment. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, **1**(2), 92–97.

Larrabee, G. J., & Levin, H. S. (1986). Memory self-ratings & objective test performance in a normal elderly sample. *Journal of Clinical & Experimental Neuropsychology*, **8**(3), 275–284.

Larrabee, G. J., Trahan, D. E., Curtiss, G., & Levin, H. S. (1988). Normative data for the verbal selective reminding test. *Neuropsychology*, **2**, 173–182.

Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

MacLeod, C. M. (1985). Learning a list for free recall: Selective reminding versus the standard procedure. *Memory and Cognition*, **13**(3), 233–240.

Mayes, A. R. (1986). Learning and memory disorders and their assessment. *Neuropsychologia*, **24**, 25–39.

McCarthy, M., Ferris, S. H., Clark, E., & Crook, T. (1981). Acquisition and retention of categorized material in normal aging and senile dementia. *Experimental Aging Research*, **7**, 127–135.

Parsons, O. A., & Prigatano, G. P. (1978). Methodological considerations in clinical neuropsychological research. *Journal of Consulting and Clinical Psychology*, **46**, 608–619.

Perrine, K. R., Novelly, R. A., Westerveld, M., Spencer, D. D., & Sass, K. J. (1989, February). *Factor analysis of the selective reminding test and measures of memory and intelligence.* Paper presented at the Seventeenth Annual Meeting of the International Neuropsychological Society, Vancouver, B.C., Canada.

Peters, B. H., & Levin, H. S. (1977). Memory enhancement after physostigmine treatment in the amnestic syndrome. *Archives of Neurology*, **34**, 215–219.

Peters, B., & Levin, H. S. (1979). Effects of physostigmine and lecithin on memory in Alzheimer's disease. *Annals of Neurology*, **6**, 219–221.

Rey, A. (1964). L'examen clinique en psychologie. Paris: Presses Universitaires de France.

Rosen, W. G., Terry, R. D., Fuld, P. A., Katzman, R., & Peck, A. (1980). Pathological verification of ischemic score in differentiation of dementias. *Annals of Neurology*, **7**, 486–488.

Ruff, R. M., Light, R. H., & Quayhagen, M. (1989). Selective Reminding Test: A normative study of verbal learning in adults. *Journal of Clinical and Experimental Neuropsychology*, **11**(4), 539–550.

Russell, E. W. (1981). The pathology and clinical examination of memory. In S. Filskov & T. Boll (Eds.), *Handbook of clinical neuropsychology* (pp. 287–319). New York: Wiley.

Stern, Y., Sano, M., & Mayeux, R. (1987). Effect of oral physostigmine in Alzheimer's disease. *Annals of Neurology*, **220**, 306–310.

Thal, L. J., & Fuld, P. A. (1983). Memory enhancement with oral physostigmine in Alzheimer's disease. *New England Journal of Medicine*, **308**, 720–721.

Thal, L. J., Masur, D. M., Blau, A. D., Fuld, P. A., & Klauber, M. R. (1989). Chronic oral physostigmine without lecithin improves memory in Alzheimer's disease. *Journal of the American Geriatric Society*, **37**, 42–48.

Trahan, D. E., Larrabee, G. J., Quintana, J. W., Goethe, K. E., & Willingham, A. C. (1989). Development and clinical validation of an expanded paired-associate test with delayed recall. *The Clinical Neuropsychologist*, **3**, 169–183.

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

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

Wechsler, D. (1987). *Wechsler Memory Scale–Revised. Manual.* New York: Psychological Corporation.

Wettstein, A. (1983). No effect from double-blind trial of physostigmine and lecithin in Alzheimer's disease. *Annals of Neurology*, **13**, 210–212.

Yesavage, J., Brink, T., Rose, T., Lum, O., Huang, O., Adey, V., & Leirer, V. (1983). Development and validation of a geriatric depression scale: A preliminary report. *Journal of Psychiatric Research*, **17**, 37–49.