New Zealand Journal of Psychology, 1990, 19, 17-23

A Test of Remote Memory for Use with New Zealand Subjects*

Fiona J. Longmore

Psychological Services, Department of Justice, Dunedin.

Robert G. Knight

Psychology Department, University of Otago

Barry E. Longmore

Psychological Medicine Department, Otago Medical School.

The Test of Remote Memory (TRM) was developed as a measure of the remote memory performance of New Zealand patients and subjects for use in both clinical and research contexts. Normative data from 189 community volunteers between the ages of 27 and 88 are presented. The results are reported for six decades (from the 1930's to the 1980's) and are separated into six cohort-age groups, each spanning 10 years. The normative data and results from several neurologically impaired patients provided encouraging evidence of the test's validity. Age of subjects consistently affected remote memory performance with global deterioration in scores being associated with increasing age.

The assessment of memory functioning is an integral part of the evaluation of cognitive deficits in neurologically impaired patients. The majority of tests used commonly in clinical practice, such as the Wechsler Memory Scale-Revised, the California Verbal Learning Test, and the Rey Auditory Verbal Learning Test (Lezak, 1983), involve the measurement of learning and retention within a clinical testing session. Such tests focus primarily on the failure of new learning or anterograde amnesia (AA). The loss of previously acquired memories, retrograde amnesia (RA), is usually assessed by the clinician more informally, often by ascertaining how much of their personal history patients can recall. Determining degree of RA is frequently of significance in monitoring recovery from traumatic head injury or in documenting the progression of a dementing disorder in elderly patients. Frequently, RA is assessed by informal methods during a mental status examination. However, questions focusing on patients' autobiographical memoranda can be problematic, particularly when there is no reliable source against which to check their responses, and informal tests of memory for public information (e.g., "name six presidents of the United States") often prove to be too difficult or insensitive when subjected to experimental scrutiny (e.g., Hamsher & Roberts, 1985).

The more formal assessment of RA requires the availability of a reliable, standardized, and valid measure of remote memory. Herrmann (1983) cited over 40 such measures in his survey of these tests, the majority of which have been developed within the past 20 years. The most frequently cited remote memory tests are those developed by Squire (1974) and Albert, Butters, and Levin (1979) in the United States, and by Sanders and Warrington (1971) in the United Kingdom. There are, however, some substantial technical problems to be overcome when constructing a useful measure of integrity of remote memory. These difficulties primarily result from a lack of control over the nature of the learning episode subsequently tested. That is, remote memory questionnaires ask for the recall or recognition of famous persons or events ("Who was Jack Manchester?" or "What was VE day?") that

*The authors gratefully acknowledge the assistance of Dr Brian Meehan of the Geography Department, University of Otago and Mr Law of Dunedin City Council in recruiting elderly subjects, and Nicola Sewell for her help in the process of data analysis. Our thanks also to Eric Shelton for his assistance in the testing of the amnesic patients. This project was funded in part by an Otago University Research Committee award.

Address all correspondence to Fiona J. Longmore, Psychological Services, Department of Justice, P.O. Box 5440, Dunedin, New Zealand.

people may or may not have learned about at the time when they were most salient.

More specifically, the validity of a remote memory test for assessing RA depends on several factors. First, it is important to select items in which the information required for correct solution was available to the subject only at a particular time. Items must be time or decade specific. For example, an item such as "Who was Adolph Hitler?" is of little value, since his enduring notoriety ensures that most of the adolescent and adult population respond correctly, despite the fact that he died over 40 years ago. This requires that test items be checked so as to ensure that only those persons who experienced an event are likely to recall it, and that those born after an event have generally not acquired knowledge of it during the course of their education and socialization. This requirement for temporal specificity can be tested by comparing scores of subjects in different age groups on decade specific items. For each item, where appropriate, there should be a relatively sharp demarcation in correct responding between those subjects who were alive at the time of a famous occurrence and those who were not.

The second problem is the necessity for items to be of equivalent difficulty across specific decades. If any bias in item sampling occurs, such that, for example, remote items are easier than more recent ones, the apparent resistance of remote memory to decline could be an artifactual consequence of item difficulty. It is important therefore that items be homogeneous. One solution to this problem was provided by Squire and Slater (1975), who developed the Television Test. This test consisted of items testing memory for television shows that were broadcast for only one season and relied on the fact that over 90% of US households owned a TV over the relevant period. Although this procedure produced a high degree of temporal specificity and item homogeneity, it was dependent on the viewing habits of individual subjects (Harvey and Crovitz, 1979). This measure was also of limited time span, with data collection able to extend back only to the 1950s.

A third problem is that subjects must have been exposed to the information assessed by the items. At an individual level, there is no way of knowing whether subjects have failed particular items because of retention failure or because they never acquired the information tested in the first place. However, in general terms the difficulty of items and their appropriateness to a local sample needs to be carefully assessed. Just as Information subtest items from the Wechsler Adult Intelligence Scale (WAIS) do not necessarily generalize to New Zealand, items from remote memory tests developed in the United States and based on famous political or sporting figures are unlikely to be of use outside North America. Similarly, a test developed in New Zealand is likely to be of little value in assessing recent immigrants to this country.

The purpose of the present report is to describe the construction of a remote memory test appropriate for use with long-term residents of New Zealand and to report some preliminary normative data. It was intended that this test have both research and clinical utility. The clinical uses would include documentation of the nature and extent of cognitive impairment resulting from brain injury or disease and changes in memory functioning over time. Research uses would include further studies of RA in different pathological groups.

Method

Test construction

The final version of the Test of Remote Memory (TRM) was the product of several studies in which items were tested and subsequently reformulated or discarded. This programme of research is described in detail elsewhere (F. Longmore, 1989). It was decided at an early stage that names of famous persons would constitute the basis of test items and that a correct response would involve recognizing the reason for their fame from a number of alternatives. A recognition format was employed to overcome some of the problems inherent in a recall test. These include the difficulty of formulating comprehensive marking schedules and the fact that subjects vary in their willingness to guess or to express answers in which they have a low degree of confidence. To minimize the effect of guessing, eight plausible response alternatives were generated for each item. Each item took the following form:

66. Martin Donnelly is/was best known as a(n):
1. Criminal 2. Commentator 3. Reformer 4. Explorer 5. Pilot 6. Entertainer 7. Cricketer 8. Sailor

In all, 237 item names were evaluated. The final version of the TRM comprised 70 items selected from those used in the first two draft versions of the test. Because the test was designed to be used with subjects up to the age of 85, it was important that it not be too lengthy, particularly given the multi-choice response format. The 70 items selected were made up of 12 items from each of the decades beginning 1930, 1940, 1950, 1960 and 1970, and 10 items for the decade beginning 1980. It was on this version that the norma-

tive data described below were based. Of the 70 people referred to in the items of the final version, nine were persons suspected or convicted of criminal activity (e.g., Sirhan Sirhan, Peter Fulcher), 22 were people involved in sport (e.g., Yvette Williams, Tracey Austin), 19 were known for their contribution to the Arts or entertainment industry (e.g., Rowena Jackson, Eleanor Powell), nine were famed for achievement in medicine or science (e.g., Benjamin Spock, Howard Gillies), and a further 11 for a variety of other reasons (valour, political activities, etc.) Fifty-nine percent of the item names were men, and 41% were women.

Two major criteria were used for item selection. First, item names had to be well known to subjects who were at least 10 years old at the time when the famous person was best known. To determine this, draft versions of the test were circulated informally to groups of subjects aged over 40. Items in the early versions of the test known to less than 50% of the age-relevant members of the samples tested were discarded during the development process. Items were also deleted if they were obviously better known to subjects of one sex or the other within a relevant age group. Second, items that were not decade specific were discarded. This was achieved by administering an earlier version of the test to University students and item names of those persons whose period of fame was before the students were 10 years old, that were recognized by more than about 30% of the students, were discarded. For example, many students knew of Fred Perry, the 1930's tennis player. Further, it was important that recent items be well known to subjects of all ages. Alice Cooper, for example, was well known to younger subjects but not to the older groups of subjects.

Normative sample

The usefulness of normative data from a remote memory test is ephemeral; at best such data are of value for the decade after they were collected. Testing a large stratified sample is extremely expensive and given the obsolescence of the data, probably not warranted. Therefore, although age group means are presented in this paper, primarily to document the validity of the TRM, the results from individual subjects have been tabulated to allow research or clinicians to choose for themselves subgroups of relevant comparison subjects. These data are available from the authors on request.

The normative sample comprised 189 subjects ranging in age from 27 to 88 years. An elderly sample was recruited by visiting pensioner cottages in Dunedin. Other subjects were members of service groups, senior citizens associations, and church groups in Invercargill and Dunedin. Only persons who had lived in New Zealand since the age of 10 or who had arrived in the country before 1930 were included in the sample. The 189 subjects were divided into six age-related cohorts as described in Table 1. Females predominated in the age range 57 to 88. About 66% of the subjects had been born in a large New Zealand town or city, about 30% in a small rural township or community, and 4% were born outside New Zealand. A total of 68.6% were married, 19% were widowed, 3% separated or divorced, and 9.2% were coded as never married. The normative group were generally healthy: Frequency of visits to the doctor varied, but on average was once a year or less. Four different ratings of socio-economic status (SES) were collected. Generally SES was biased towards higher classifications. The distribution of ratings that was closest to that of the New Zealand population in general and that contained the fewest missing values was that for father's occupation. This was chosen as the SES variable for later analyses.

Table 1: *Age ranges and subject numbers for the normative sample*

Age Range	Cohort	Male	Female	Total
27-36	1	22	21	43
37-46	2	18	19	37
47-56	3	16	10	26
57-66	4	11	15	26
67-76	5	3	25	34
77-88	6	5	18	23
Total		81	108	189

Patient sample

To illustrate the clinical use of the TRM, results from a small number of neurologically impaired patients have been reported below. Two patients had a presumptive diagnosis of senile dementia of the Alzheimer's type, and were attending the Day Hospital at Wakari Hospital. In addition, results are reported from three patients previously diagnosed as suffering from Korsakoff's disease as a result of prolonged alcohol abuse. These three patients had previously taken part in a more extensive study in which their amnesic deficits had been documented in detail (B. Longmore & Knight, 1988). These patients had Wechsler Memory Scale Memory Quotients (MQ) of 73, 81 and 82, and WAIS IQ/MQ discrepancies of 20, 20, and 29. For each patient tested, results from four or five subjects who matched the patient for age, SES, and health status, were selected for comparison purposes from the normative sample. Patients completed the TRM with assistance from the first author.

Results

Normative sample

Subjects were divided into six age cohorts each spanning 10 years, as in Table 1. The means, standard deviations, and percentage of correct responses for each decade and for the total test, are reported for each cohort in Table 2. Percentages of correctly answered items for each of the six cohort groups are shown together in Figure 1. As can be seen in Table 2 and Figure 1, there are

Table 2: *Means, Standard Deviations and Percentages of Correct Responses Reported by Decade and Total Scores*

		Cohorts (Age range)						
		1	2	3	4	5	6	ALL
Test Decade		(27-36)	(37-46)	(47-56)	(57-66)	(67-76)	(77-88)	27-28
	M	8.48	8.13	7.68	7.92	6.91	6.59	7.71
1980s	*SD*	1.00	1.68	1.88	1.09	2.52	2.06	1.86
	%	84	81	77	79	69	66	77
	M	10.11	9.70	9.56	9.53	8.02	7.78	9.21
1970s	*SD*	1.72	1.82	1.95	1.83	3.37	3.13	2.43
	%	84	81	80	79	67	65	77
	M	8.44	9.45	9.56	9.61	8.55	7.34	8.86
1960s	*SD*	2.17	1.59	1.58	1.06	2.09	2.02	1.99
	%	70	79	80	80	71	61	74
	M	3.02	6.45	9.32	9.03	7.61	7.13	6.70
1950s	*SD*	2.10	2.64	1.84	1.70	2.73	2.30	2.29
	%	25	54	78	75	63	59	56
	M	2.13	3.21	7.84	9.73	7.67	6.65	5.71
1940s	*SD*	1.97	2.27	3.19	1.75	2.67	2.72	3.48
	%	18	27	65	81	64	55	48
	M	3.14	3.72	6.72	8.07	8.35	7.43	5.94
1930s	*SD*	1.57	1.95	2.80	2.41	3.09	2.96	3.20
	%	26	31	56	67	70	62	50
	M	36.54	41.62	51.80	55.07	48.02	43.65	45.15
Total	*SD*	7.28	8.76	10.91	6.94	14.11	13.43	10.38
	%	52	59	74	79	69	62	65

high and steady rates of recognition for item names of people famous from the time subjects were at least 10 years old. For item names from decades where subjects were aged 10 years or less, recognition rates are lower than for the item names that post-date them and higher than for those that pre-date them. Recognition rates are lowest for item names of people whose fame predated the subject's birth.

The differences in temporal gradients for the different cohorts, which are apparent in Figure 1, were further evaluated using a series of analyses of variance (ANOVA). In the first set of analyses, one-way ANOVA's were calculated for each decade of test items (1980's to 1930's) and for all items combined. Each ANOVA tested differences between the scores of subjects in each of the six cohort groups on the items from one particular decade. Seven separate ANOVA's were computed and the results are presented in Table

Figure 1: Mean Percentages Correct for each Decade for all Subjects (N = 189).

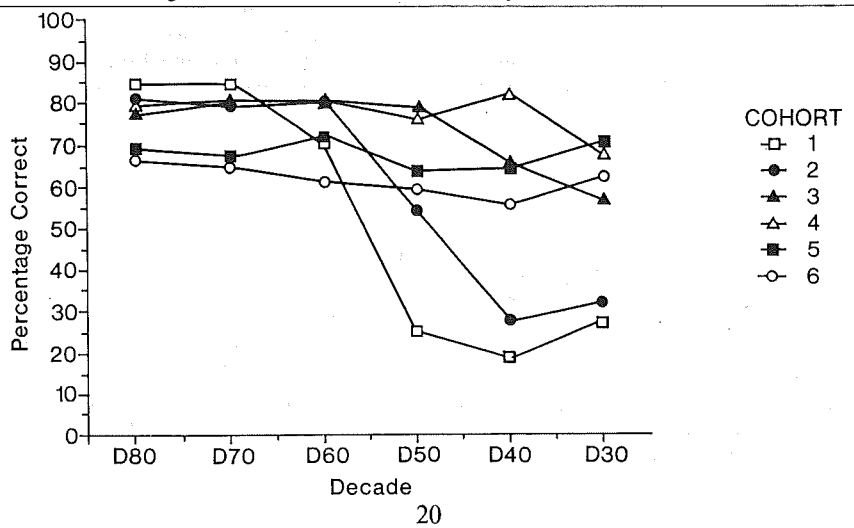


Table 3: *Results of One-way ANOVA's and Tukey's Tests of HSD for Decade and Combined Total Scores*

Test Decade	*F-value* [a]	*p*	Tukey's HSD
1980's	5.52	.0001	1.25
1970's	5.18	.0002	1.67
1960's	6.44	.0000	1.32
1950's	35.44	.0000	1.62
1940's	49.77	.0000	1.72
1930's	28.06	.0000	1.73
All	12.13	.0000	7.34

Note: HSD = Honestly Significant Difference test.
[a] Degrees of freedom 5 and 182 in each case.

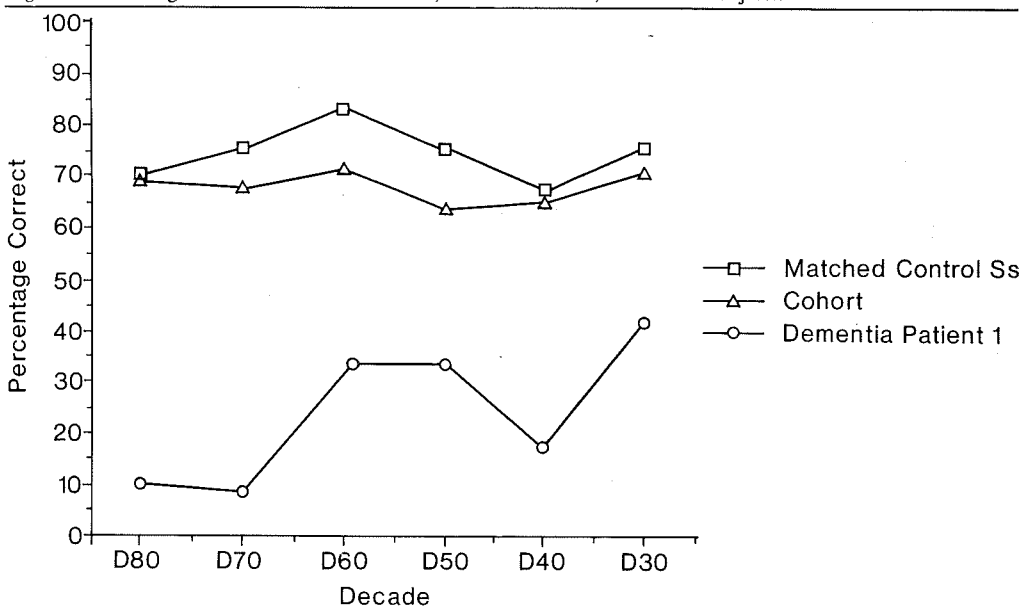
3. *Post hoc* analyses using Tukey's Honestly Significant Difference (HSD) test were subsequently computed. Critical values of Tukey's HSD at the 5% level of significance are detailed in Table 3. There were highly statistically significant differences between age groups for all decades of item names. The Tukey's tests revealed two trends. Firstly, age cohorts 1 (age 37-36) and 2 (age 37-47) recognised fewer item names from the decades 1930-1950, which were during or before their childhood, than did subjects from the other four age cohorts. Secondly, the most elderly subjects, cohorts 5 (age 67-76) and 6 (age 77-88) performed relatively poorly when compared to younger subjects on items from recent decades (1960-1980) presumably equally familiar to subjects from all 6 cohorts.

In the second set of analyses, a series of six one-way repeated measures ANOVA's were computed, one for each of the six cohorts, using a within subjects design with the six decade totals (1930s to 1980s) as dependent variables. Scores for Decade 1980 were prorated to make them equivalent (maximum = 12) to the other decade totals. The magnitudes of the F-values decreased for the ANOVAs of the higher aged cohorts. Highly significant differences between decade scores were found for age cohort 1, F (5,210) = 273.6, $p < .001$, age cohort 2, F (5,180) = 132.9, $p < .001$, age cohort 3, F (5,120) = 14.59, $p < .001$, and age cohort 4 F (5,125) = 4.12, $p < .002$, but not for cohorts 5 (F = 1.67) and 6 (F = 1.90). These results reflect the increasing temporal gradients for cohorts 4 to 1, and the absence of temporal gradients for cohorts 5 and 6.

Cronbach's alpha coefficients were calculated as a measure of the internal consistency of items. Alpha for the 70-item test was .92; for the 12 item subtests from each of the decades from 1930 to 1970, the alpha coefficients were .78, .86, .79, .69, and .74 respectively. For the subtest comprising the 10 items from the 1980s, alpha equalled .65 indicating a more modest level of reliability.

Analyses of variance conducted on total scores revealed no statistically significant differences between subjects on the basis of sex, F (1,187) = 2.57, p = .11, father's occupations, F (6,182) = .73, p = .62, overall state of health, F (2,186) = 2.49, p = .09, or number of years spent at school, F (9, 179) = .69, p = .72. Partial correlations were

Figure 2: Percentage Correct: Dementia Patient 1, Matched Control, and Cohort Subjects.



calculated to study the effects of these demographic variables on the association between subjects' ages and their total scores. With significance levels set at $p < .001$, to reduce the probability of Type 1 errors (because of the large number of correlations calculated), the only variable having any significant association with the test scores was age. The significant and positive direction of this association remained unchanged with the effects of other variables partialled out.

Patient sample

Data from five patients are presented to exemplify the use of the TRM and provide clinical evidence for it's validity. In Figures 2 and 3, percent correct results for each decade from the two demented patients are presented. Both patients performed poorly relative to their matched controls and their age cohort. Their results are generally suggestive of an overall rate of decline, although the data from Patient 1 suggest some sparing of remote memories. Results from the patients with alcoholic Korsakoff's disease are presented in Table 4. Patient 1 showed a general but ungraded decline in remote memory performance. Patient 2, however, revealed the typical profile of a patient with RA. His memory for remote events was unexceptional but his recognition of famous people from 1970 onwards was grossly abnormal. Patient 3 showed signs of RA, but also performed poorly on items from the 1950s.

Discussion

Results from the normative sample provided encouraging evidence for the reliability and validity of the TRM. Total scores from the 70-item test were found to be highly reliable. The reliabilities of the subtest measuring memory for items from individual decades were acceptable given the relatively small number of items per decade (a consequence of keeping the length of the test at a level likely to make its use with elderly and neurologically impaired patients realistic). Despite considerable demographic variability between subjects in the normative sample, the test was sensitive only to the effects of age. The obvious age effects supports the necessity for age-specific norms.

Generally, the patterns of age-related memory performance were consistent with those reported by Squire (1974) and Sanders and Warrington (1971): Decline in remote memory was greatest for the most elderly groups, and was most apparent for items from the most recent decades. On the present test, using recognition format for responding, the decline in performance was marked between groups with a mid-point age at testing of 60 years or less and those groups where the mid-point age was 70 or more (cohorts 5 and 6). This decline in performance was consistent over all decades of items. Further research, involving concurrent assessment of cognitive function, premobid intellectual ability, and remote memory would permit a test of whether remote memory

Figure 3: Percentage Correct: Dementia Patient 2, Matched Control, and Cohort Subjects

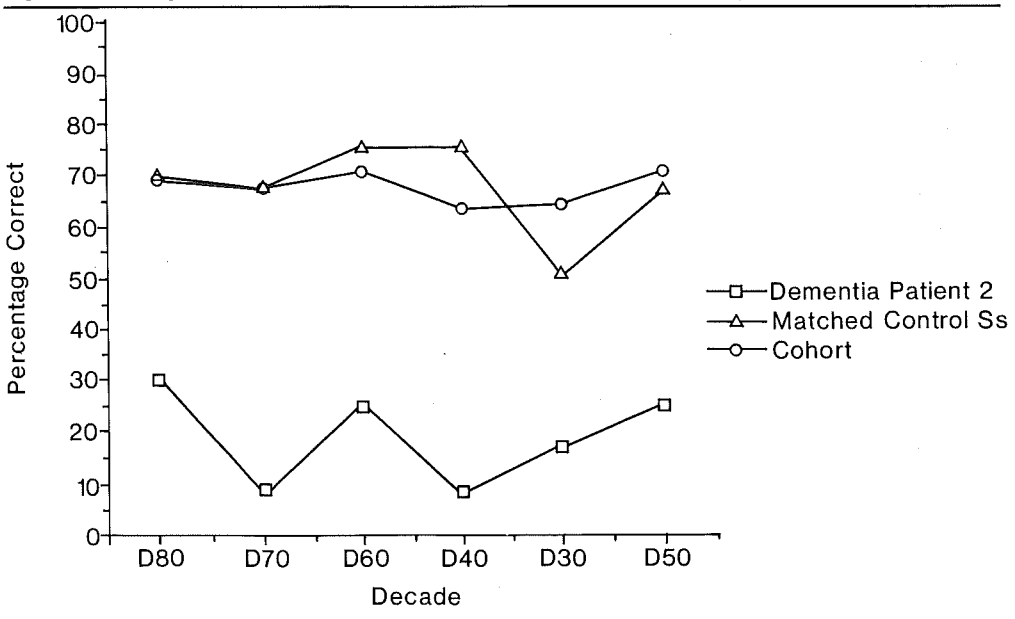


Table 4: *Percentages Correct per Decade for Korsakoff Patients, Matched Control, and Cohort Subjects*

	Decade					
	1980	1970	1960	1950	1940	1930
Subject Group						
Patient 1	60	25	42	42	25	17
Matched Controls	70	75	83	67	42	42
Cohort Group	76	79	81	77	65	55
Patient 2	10	50	67	83	75	75
Matched Controls	90	75	92	75	83	58
Cohort Group	79	79	80	75	80	66
Patient 3	10	17	75	25	75	42
Matched Controls	70	75	83	83	75	83
Cohort Group	69	66	71	63	63	69

impairment is a marker of a general decline in cognitive functioning, or as Squire (1974) has suggested, a specific age-related deficit.

Equivalence of memorability of the items from each decade is provided by the flat gradients for normative subjects in cohorts 5 and 6 (over the age of 67). The poor performance of patients with established amnesia following neurological disease adds further weight to the validity of the test. The results from the Korsakoff patients generally supported the findings of Cohen and Squire (1981) and Albert et al. (1979). Marked temporal gradients appear to be most characteristic of the RA in these patients.

In conclusion, the TRM demonstrated good content validity and can be confidently described as a test of remote memory rather than one of historical knowledge. The generally high rates of recognition by subjects of item-names famous after the subjects' childhood, make the test particularly suitable for the study of pathological patients, because the probability of floor effects obscuring important differences has been minimized. Patients generally find the test interesting to complete and appear well motivated to do well.

References

Albert, M.S., Butters, N., & Levin, J. (1979). Temporal gradients in the retrograde amnesia of patients with alcoholic Korsakoff's disease. *Archives of Neurology, 36,* 211-216.

Cohen, N.J., & Squire, L.R. (1981). Retrograde amnesia and remote memory impairment. *Neuropsychologia, 19,* 337-358.

Hamsher, K.De S., & Roberts, R.J. (1985). Memory for recent U.S. presidents in patients with cerebral disease. *Journal of Clinical and Experimental Neuropsychology, 7,* 1-13.

Harvey, M.T., & Crovitz, H.F. (1979). Television questionnaire techniques in assessing forgetting in long-term memory. *Cortex, 15,* 609-618.

Herrmann, D.J. (1983). Questionnaires about memory. In J.E. Harris & P.E. Morris (Eds.), *Everyday memory, actions, and absent-mindness* (pp.133-151). London: Academic Press.

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

Longmore, B.E., & Knight, R.G. (1988). The effect of intellectual deterioration on retention deficits in amnesic alcoholics. *Journal of Abnormal Psychology, 97,* 448-454.

Longmore, F.J. (1989). *The Otago Remote Memory Test.* Unpublished Masters thesis submitted to the University of Otago, Dunedin, New Zealand.

Sanders, H., & Warrington, E.K. (1971). Memory for remote events in amnesia patients. *Brain, 94,* 661-668.

Squire, L.R. (1974). Remote memory as affected by aging. *Neuropsychologia, 12,* 429-435.

Squire, L.R., & Slater, P.C. (1975). Forgetting in very long-term memory as assessed by an improved questionnaire technique. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 50-54.