Normative Data for New Zealand Elders on the Controlled Oral Word Association Test, Graded Naming Test, and the Recognition Memory Test

J. Anne Harvey
Wellington School of Medicine, University of Otago

Richard J. Siegert
Victoria University of Wellington

The use of psychometric tests for diagnostic purposes in neuropsychological settings is dependent on the availability of appropriate norms for interpreting an individual's performance. Most psychometric measures have been developed and standardised on American or British populations and may not be applicable to individuals or groups from other cultural backgrounds. Although there has been an increase in the development of overseas norms for elderly people, there is a notable lack of norms for elderly New Zealanders. Normative data, obtained from a community sample of 139 New Zealanders 70 to 90 years of age, are presented for the Recognition Memory Test, the Graded Naming Test, and the Controlled Oral Word Association Test. The correlations among the three tests are reported for the whole sample, and norms are presented for three different age bands. The relationship between declining functioning status, prior psychiatric disorder and cognitive functioning is evaluated as a possible indicator of cognitive decline. The impact of cognitive performance and intelligence is also considered.

The importance of using appropriate norms in forming judgments based upon psychological test results is a *sine qua non* of psychological testing (Anastasi & Urbina, 1997). When making inferences about an individual's performance upon a particular test, regardless of whether that test attempts to measure cognitive functioning, personality, or any other psychological construct, it is important to know how the individual compares with their peers. Consequently, when evaluating an individual's score on a psychological test it is essential to do so in relation to norms that tell us how people of similar age, gender, educational attainment, and ethnicity might score. This is of particular importance when evaluating cognitive abilities in elderly people, in whom the distinction between normal age-related decline and clinically significant impairment, is often not clear. With a rapidly ageing population in most technologically advanced countries, and a concomitant increase in age-related neurological disorders, the importance of developing diagnostic measures that are sensitive to cognitive dysfunction in older adults has become pronounced. In New Zealand, The National Health Committee on Health and Disability (1997) has stated that New Zealanders over 65 have a one in twenty chance of developing dementia, or a one in five chance if they are over 80. Consequently, there has been a growing demand for objective standards to aid in the diagnosis and assessment of those age-related illnesses where cognitive decline may be a primary symptom. Although such tests are widely available, their norms typically stop at just around the age where dementia starts to become relatively common. In addition, there are rarely ever norms developed for such tests with New Zealand individuals - a particular problem for language tests which are likely to be highly subject to international differences. Despite these issues, clinical psychologists are frequently asked to conduct neuropsychological assessments on elderly people, and to make judgments concerning their cognitive status in the absence of appropriate normative data.

Lezak (1987) considered the age norms available for ten of the most widely used American tests and commented that adequate age norms for the elderly were largely non-existent. However, this situation has improved somewhat over the past decade or so since Lezak made those observations. For example, the Mayo Clinic have collected population-based neuropsychological test norms for elderly people resident in Olmsted County, U.S.A. These projects are known as the Mayo's Older American Normative Studies (MOANS). Currently, MOANS normative data is available for elderly Americans ages 55 to 97 on 15 neuropsychological instruments assessing a broad spectrum of cognitive functions, some of which include; the Rey

Auditory Verbal Learning Test (RAVLT; Rey, 1964; Ivnik, Malec, Tangalos, Peterson, Kokmen, & Kurland, 1990; Ivnik, et al, 1992a), the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981; Ivnik et al., 1992b; Malec et al., 1992), the Wechsler Memory Scale-Revised (WMS; Wechsler, 1987; Ivnik et al.,1992c). These were later expanded (Ivnik, Malec, Smith, Tangalos, & Petersen, 1996) to include norms for the Controlled Oral Word Association Test (COWAT) letter fluency and the Token Test of the Multilingual Aphasia Examination (Benton & Hamsher, 1978), the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1978), the Reading subtest of the Wide Range Achievement Test - Revised (WRAT-R; Jastak & Wilkinson, 1984), the American version of the National Adult Reading Test (AMNART; Grober & Sliwinski, 1991), the Stroop Colour and Word Test (Golden, 1978), the Trail Making Test (TMT; Spreen & Strauss, 1991), and the Judgement of Line Orientation Test (JLO; Benton, Hamsher, Varney, & Spreen, 1993). More recently, MOANS norms have been provided for the Free and Cued Selective Reminding Test (Grober & Buschke, 1987; Ivnik et al., 1997) and the COWAT verbal fluency test has been extended to include norms for the semantic version (Lucas, Ivnik, Smith, Bohac, Tangalos, Graff-Radford, & Petersen, 1998). In addition, Ryan, Paolo, and Brungardt (1990) standardised the WAIS-R for American adults aged 75 years and above and the more recent third edition of the WAIS (WAIS-III) now extends to age 89 (Wechsler, 1997).

Despite an increase in the availability of overseas norms for elderly people, researchers consistently caution the validity of norms when applied to individuals from other ethnic and cultural backgrounds (Ivnik et al., 1998; Knight, 1997; Lezak, 1995). Knight (1997, p.4) commented recently that "the use of the WAIS-R in New Zealand requires the user to accept the credentials of a test normed in the United States". Nonetheless, due to the lack of available norms developed in New Zealand, clinical psychologists typically rely on British or American norms to interpret test results of elderly New Zealanders. However, there have been some attempts in recent years, to develop age norms suitable for use in interpreting the scores of elderly New Zealanders on neuropsychological tests. Newlove (1992) reported norms for 130 healthy, community-dwelling New Zealanders, for the RAVLT. She reported norms for five-year age bands (60-64, 65-69 etc.) from 60 - 80+. Interestingly, she reported somewhat lower means on this test than Ivnik, et al., which supports the point Knight has made about not relying upon American or British norms. More recently Siegert and Cavana (1997) have reported norms for the Trail-Making Test, a popular test of attention, from the same sample of elders as Newlove. Also worthy of mention is Louden (1996), who developed a new test of semantic or remote memory (The Kiwi Famous Faces Test), and gathered norms for both younger (18-25) and older (65-80) subjects.

In summary, the use of appropriate age-related norms is essential when interpreting psychological and neuropsychological test scores. There have been some notable attempts in recent years, particularly by researchers in North America, to overcome the serious lack of norms for elderly people on many neuropsychological tests. There have also been some local attempts to remedy this situation. The purpose of the present study was to gather norms for New Zealand elders on three neuropsychological tests which each measure different cognitive abilities or functions. The tests include the Recognition Memory Test (RMT; Warrington, 1984), a test of verbal and visual recognition memory, the Graded Naming Test (GNT; McKenna & Warrington, 1983; Warrington, 1997), a test of naming or word-finding difficulties, and the Controlled Oral Word Association Test (COWAT; Benton, & Hamsher, 1989), a test of verbal fluency. This battery was chosen because each test measures different cognitive functions, is widely used by clinicans, and could easily be administered within an hour and half session without excessive demands on participants.

Method

Participants

A non-random sample of 139 elderly adults, living independently or semi-independently in the greater Wellington region, was recruited for voluntary participation. Ages ranged from 70 to 90 with a mean age of 77.24 (s.d. = 4.8). The sample comprised of 94 women (68%) and 45 men (32%). Participants were grouped in to three age brackets for normative purposes. Those between 70 and 74 years (n = 45), 75 and 79 (n = 49), and 80 and above (n = 45). A large percentage (46%, n = 75) had received some post-secondary education, and 22% (n = 30) reported no secondary education. Table 1 compares the participants' occupation levels at retirement with the 1991 census data for European men and women of working age (21 - 69 years)

Table 1. Occupation (Major Group) for European Full-Time Workers aged 21-69 (1991 census)* compared with Elderly Sample's Occupation before Retirement

Occupation	1991 census %	Sample %
Legislators, Administrators & Managers	15.3	3.6
Professionals	13.8	3.0
Technicians & Associate Professionals	12.6	28.8
Clerks	12.7	7.2
Sales and Service Workers	8.4	5.8
Agriculture and Fishery Workers	10.7	0.7
Trades Workers	12.3	15.8
Plant and Machine Operators & Assemblers	8.9	12.2
Elementary Occupations	4.7	2.9
Armed Forces	0.5	0.0
Total	100.0	100.0

* *Source: Davis, Mcleod, Ransom, & Ongley (1997).*

(Davis, Mcleod, Ransom, & Ongley, 1997). In the case of married or de facto couples, the higher family occupation was used to classify participants according to the New Zealand socioeconomic index (NZSEI; Davis et al., 1997).

Of the total sample, 14% (n = 19) were receiving some form of community care, such as Meals on Wheels, district nursing, or services from Home Care 2000. This support was most evident among those participants who were either widowed, single, living alone, or in the 80+ age group. For example, of those receiving community support, 84% (n = 16) were living alone. Only 4% (n = 3) of the sample who were living with a spouse received community support, and this was usually for nursing care. Sixteen percent (n = 22) reported they had experienced at least one depressive episode at some stage in their lives.

The majority of participants described themselves as right-handed (86%, n = 119), 9% (n = 12) were left handed, and 6% (n = 8) were classified as ambidextrous because they used each hand predominantly for two or more important functions. The demographic characteristics of the sample are summarised in Table 2. The only variables that showed significant sex differences were Years of Education and WAIS-R scores with men having more years of education than women (t = -3.22, $p < .01$), and higher WAIS-R scores (t = -3.74, $p < .01$). The majority of participants described themselves as European or Pakeha.

Table 2. Sample Demographics and Characteristics

	Age group (years) (70-74)	(75-79)	(80+)	Total
Females (n)	33	32	29	94
Age				
M	71.91	76.63	83.34	77.04
SD	1.3	1.4	3.0	5.0
Education (years)				
M	12.73	12.06	10.93	11.95
SD	5.8	3.2	1.8	4.1
NART				
M	106.42	107.50	106	106.72
SD	12.2	13.9	12.3	12.7
WAIS-R Vocab.				
M	53.58	51.28	49.76	51.62
SD	12.7	13.4	10.9	12.4
Males (n)	12	17	16	45
Age				
M	72.58	76.82	82.31	77.64
SD	1.3	1.4	2.3	4.3
Education (years)				
M	14.00	14.82	14.25	14.40
SD	4.9	4.6	4.3	4.5
NART				
M	110.50	112.18	107.19	110
SD	10.2	11.1	12.2	11.2
WAIS-R Vocab				
M	58.50	59.82	55.75	58.02
SD	8.4	7.1	7.5	7.6
Total Sample (N)	45	49	45	139

One person was of Asian ethnicity. All were fluent in English although 3 (2%) identified English as their second language. All participants had adequate eyesight (corrected or uncorrected) and hearing to cooperate and complete the tasks.

Procedure

Participants were recruited from the greater Wellington region, including Upper Hutt, Wainuiomata, Porirua and the Wairarapa. All were residing in the community either independently or semi-independently (i.e. receiving community support). Most participants were contacted through senior citizens groups, social groups, and churches. A small number of participants were contacted through a Victoria University research volunteer database, and some were friends or acquaintances of the first author. The majority of participants were interviewed in their own home. The few participants who travelled to Victoria University to participate in the study were reimbursed for travel expenses. Initially, 166 participants were interviewed. Twenty-seven were excluded from the final sample because they reported a positive history of either a stroke, transient ischaemic attack, head trauma, major psychiatric disorder, alcohol abuse, or were currently taking antidepressant or antipsychotic medication. All participants were tested by the first author. The tests were administered with close adherence to the instructions for administration given in their respective manuals. For the COWAT, where no manual exists, the instructions were as given in Spreen and Strauss (1991). Sufficient time was allowed for rest breaks between tests if required or requested.

Materials

All participants were administered the following measures, in this order:

1. **Personal Information Questionnaire.** This questionnaire asked participants for information about their gender, ethnicity, marital status, education, occupation, residential status, and current medications. They were also asked if they had any history of stroke, alcoholism, depression, major psychiatric disorder, head trauma, or epilepsy.

2. **National Adult Reading Test** (NART; Nelson, & Willison, 1992). The NART is widely used to estimate premorbid and current verbal intellectual ability. It was administered to provide an estimate of the range of general verbal intellectual ability for this sample. Participants are required to read aloud a list of 50 words and their score is based upon the number which are pronounced correctly.

3. **Wechsler Adult Intelligence Scale - Revised** (WAIS-R) Vocabulary Scale (Wechsler, 1981). The WAIS-R was also administered to provide an estimate of the range of verbal intellectual ability for this sample. It was administered in addition to the NART since it does not depend upon reading ability. Participants are asked to provide definitions for a list of 35 words.

4. **Graded Naming Test** (McKenna & Warrington, 1983; Warrington, 1997). The GNT is a standardised test

for anomia or word-finding difficulties. The test allows for individual differences in naming ability by including items of graded difficulty. It is highly correlated with general vocabulary level and when performance is compared with a clinical estimate of premorbid intelligence results can be interpreted as an indication of anomia. Participants are shown 30 line drawings and asked to name what is in the picture.

5. **Controlled Oral Word Association Test** (Benton & Hamsher, 1989; Spreen & Strauss, 1991). A standardised test of word fluency. For the phonetic version participants are asked to generate as many words as possible in 60 seconds beginning with the letters F, A, and S, successively, and to exclude proper names and different suffixes of the same words. Examples of incorrect responses were given for the letter B. The letter R is used as a practice trial before testing commences. For the semantic version participants are asked to generate as many animal names as possible in 60 seconds, followed by a trial generating names of vegetables.

6. **Recognition Memory Test** (Warrington, 1984). A standard test of verbal (words, RMTW) and visual recognition memory (faces, RMTF). The test is administered in two parts. Participants are shown each stimulus word for 3 seconds, one at a time, and asked to judge them as either pleasant or unpleasant. Following presentation in the RMTW participants are presented with a forced-choice recognition test by presenting pairs of words (50 stimulus words and 50 distractors), and asked to identify the word in each pair, that they have previously seen. In the RMTF participants are given the same instructions as in the RMTW and in the forced-choice recognition test are asked to point to the face which they had been shown in the stimulus list.

Results

All statistical analyses were conducted using the SAS package (Hactcher & Stepanski, 1994). Descriptive statistics showed that scores for some of the tests administered reflected a non-normal distribution. Hence, non-parametric statistics were used to analyse the data. Table 3 shows the correlations among the various cognitive tests administered, Age, and total Years of Education, for the entire sample of 139 older people. For the COWAT the total number of words generated on all three letters (FAS) was used to represent overall performance on the phonetic version of the task. For the semantic version, scores for both the number of animals and vegetables generated separately were used. Inspection of Table 3 shows significant negative correlations between age and performance on the RMTF, WAIS-R Vocabulary, GNT, and the Vegetable category of the verbal fluency task. Other correlations with age are all negative but not significantly so. By contrast, the variable Years of Education, correlates positively with performance on all the measures, although not significantly so for RMTF or for the verbal fluency category of Vegetables. One-way analyses of variance (*Kruskal-Wallis H* test) were conducted to test for any sex differences in scores on the two subscales of the RMT, the GNT, and the COWAT (total score FAS). There were no statistically significant sex differences on any of these measures. Consequently, all norms reported below are for males and females combined.

Table 4 presents the scores for the three age groups, for the Words and Faces subscales of the RMT, as cumulative frequency tables. Performance on the Faces test declined significantly with increasing age ($\chi^2(2, N = 139) = 9.26$, $p = 0.0098$) but not so on the Words test ($\chi^2(2, N = 138) = 2.34$, $p = 0.3106$). The scores for the GNT are presented as cumulative frequency distributions, for the three

Table 3. Spearmans Correlation *among age, total years of education and test scores for the whole sample (n = 137-139)

Variable	Age	Education (years)	NART	WAIS-R Vocab.	GNT	FAS	Anim	Veg	RMTW
RMTF	-.27	.13	.26	.28	.20	.34	.27	.18	.30
RMTW	-.12	.39	.56	.58	.43	.44	.34	.36	
Veg	-.20	.11	.15	.17	.30	.40	.52		
Anim	-.13	.20	.29	.40	.59	.52			
FAS	-.13	.36	.46	.43	.44				
GNT	-.22	.25	.45	.49					
WAIS-R Vocab.	-.18	.62	.78						
NART	-.07	.57							
Education (Years)	-.04								

*all correlations = .18 significant at p < .05, correlations = .25 at p < .01

Note: RMTF = Recognition Memory Test - Faces; RMTW = Recognition Memory Test - Words; Veg = Vegetable semantic version of the Controlled Oral Word Association Test (COWAT); Anim = Animal semantic version of the COWAT; FAS = the actual letters used in the COWAT; GNT = Graded Naming Test; WAIS-R Vocab. = Wechsler Adult Intelligence Scale - Revised Vocabulary Scale; NART = National Adult Reading Test.

Table 4. Cumulative Frequencies for the Recognition Memory Test for Normative Sample

	RMT-Words Cumulative Frequencies (%)			RMT-Faces Cumulative Frequencies (%)		
Score (n=)	70-74 (45)	75-79 (48)	80+ (45)	70-74 (45)	75-79 (49)	80+ (45)
22					2	
26		2				
29	2				4	
30	4					4
31					8	7
32			2			9
33					10	18
34	7	4	7	7	12	24
35			11			31
36		8	16		20	36
37	11			16		42
38		14	18	24	22	49
39			22	36	33	53
40	20		24	40	39	64
41	24	21	27	42	43	67
42		25	36	47	49	69
43	27	31	42	53	67	78
44	36	40	44	67	74	84
45	40	48	47	76	82	87
46	47	60	58	84	88	91
47	56	75	67	89	96	98
48	64	79	80	98	98	100
49	82	92	91	100	100	
50	100	100	100			

Table 5. Cumulative Frequencies for the Graded Naming Test for Normative Sample

	Cumulative Frequencies (Percent)		
Score (n=)	70 – 74 (45)	75 – 79 (49)	80+ (45)
5		2	
7		2	4
9	2		
10	4		7
11		6	11
13	7	12	13
14	9		16
15	13		22
16	16	16	33
17	20	22	40
18	27	25	49
19	31	29	51
20	44	39	56
21	47	51	62
22	60	63	76
23	71	67	89
24	80	76	
25	89	86	96
26	91	94	
27	98	98	100
28	100	100	

age groups, in Table 5. There were no statistically significant differences found between the three age groups ($\chi^2(2, N = 139) = 5.80$, $p = 0.06$). There were also no statistically significant effects of age on any of the verbal fluency tasks. However, education and general verbal intelligence were influential. Table 3 shows that performance on the FAS and semantic version of animals was correlated with NART, WAIS-R Vocabulary, GNT, and Education, but not with age. Consequently, participants were then divided into four verbal IQ groups based on NART scores (<100, 100-110, 111-120, >120). Table 6 is collapsed across age groups presenting the means and standard deviations for the three phonetic trials of the COWAT, and the two semantic trials, by NART score. Differences between IQ groups were significant for the FAS and the semantic version of Animals ($\chi^2(3, N = 137) = 25.61$, $p = 0.0001$; $\chi^2(3, N = 139)$ 11.75, $p = 0.0083$ respectively), but not for the Vegetable version ($\chi^2(3, N = 139) = 6.77$, $p = 0.0796$). As expected, scores on the GNT were also significantly influenced by NART scores ($\chi^2(3, N = 139) = 24.71$, $p = 0.0001$). The RMTW was significantly influenced by NART scores ($\chi^2(3, N = 138) = 39.36$, $p = 0.0001$), There was a tendency for RMTF to be influenced by NART scores, but this did not reach significance ($\chi^2(3, N = 139) = 6.85$, $p = 0.07$).

Table 7 presents the means for the RMT Words and Faces, the GNT, and the three verbal fluency tasks, for the three age groups by NART score. The most noticeable feature of Table 7 is the clear trend for scores on most tests to show a decline in performance with decreasing NART score.

To see if declining functional status might be a probable indicator of cognitive decline, Post Hoc analyses were conducted comparing scores between the independent and semi-independent groups on all tests. There were 19 people who were categorised as living semi-independently (responses to the question, "do you receive any community support?), and 120 categorised as independent. The majority (63%) were in the 80+ age group. The only tests that showed a significant difference were the FAS (total score) ($\chi^2(1, N = 137) = 5.9769$, $p = 0.0145$; mean semi-indept. = 31.44, s.d. = 12.3; mean indept. = 40.32, s.d. = 13.2), and the letter A ($\chi^2(1, N = 137) = 9.8784$, $p = 0.0017$; mean semi-indept. = 8.06, s.d. = 3.9; mean indept. = 12.00, s.d. = 4.9). To determine if past psychiatric disorder may be influencing cognitive test performance we compared the scores for those with a self-reported history of depression (n = 22) to those without a self-reported history of depression (n = 117). There were no significant differences between groups on any of the tests.

Discussion

In the present paper we have provided normative data for New Zealand elders on three widely used neuropsychological tests. Moreover, the three tests employed cover a

Table 6. Means for New Zealand Elders for the Controlled Oral Word Association Test by NART Score

Variable	NART Scores			
	<100 n = 29	100-110 n = 41	111-120 n = 49	> 120 n = 18
Phonetic Verbal Fluency				
F Words				
M	10.31	15.56	13.27	16.83
SD	4.3	4.0	5.2	4.4
A Words				
M	7.55	11.59	12.45	14.89
SD	4.0	4.8	4.4	3.9
S Words				
M	11.72	13.85	14.98	17.83
SD	4.5	4.3	5.1	4.7
Total FAS Words				
M	29.59	39.00	41.12	49.56
SD	11.5	10.8	13.5	11.9
Semantic Verbal Fluency				
Animal				
M	13.43*	15.46	15.49	17.56
SD	3.6	3.8	3.1	4.4
Vegetable				
M	13.16**	14.00	13.41	15.61
SD	3.2	3.9	3.3	4.1

* n = 30; ** n = 31.

Table 7. Means and Standard Deviations for Each Age Group for the RMT and GNT by NART Scores

Age Group	TEST	NART Score <100	100-110	111-120	>120
	RMTW				
	M	39.90	44.4	48.54	47.43
	SD	5.8	5.0	2.5	3.1
	RMTF				
70-74	*M*	40.5	43.27	41.39	43.14
years	*SD*	3.3	3.4	5.4	4.1
	GNT				
	M	18.5	19.67	22.62	23.29
	SD	4.5	4.7	3.5	3.5
		(n =10)	(n = 15)	(n = 13)	(n = 7)
	RMTW				
	M	41.46*	43.89	44.79	47.78
	SD	4.8	2.0	5.8	1.6
	RMTF				
75-79	*M*	38.58	38.89	42.63	42.67
years	*SD*	5.2	5.9	5.7	2.7
	GNT				
	M	18.50	20.67	21.47	22.44
	SD	5.9	4.4	3.6	4.9
		(n = 12)	(n = 9)	(n = 19)	(n = 9)
	RMTW				
	M	39.56	43.47	46.24	47.50
	SD	5.3	5.1	4.0	0.7
	RMTF				
80 +	*M*	37.78	39.65	38.41	41.50
years	*SD*	4.4	6.0	4.8	2.1
	GNT				
	M	12.56	18.77	21.41	22.00
	SD	4.5	3.8	3.8	0.00
		(n = 9)	(n = 17)	(n = 17)	(n = 2)

*n = 11.

diverse range of cognitive abilities including verbal and visual recognition memory, naming, and verbal fluency. The use of these tests with these norms will help in providing accurate and meaningful neuropsychological assessment for older New Zealanders.

As expected, we found that the GNT performance was highly correlated with verbal IQ. However, performance was also mildly correlated with age, decreasing noticeably only in 80+ age group. Our findings are in accordance with Warrington's (1997) recent standardisation of the GNT. Originally standardised on a group of 100 normal individuals aged between 18 and 76 years (McKenna & Warrington, 1983), updated norms with a larger sample demonstrate that the test has become somewhat less difficult (Warrington, 1997). This shift in test performance over time highlights the need for appropriate norms that reflect the varying educational environments and their relative contribution to performance on cognitive tests. In addition, improvements in education and higher levels of education are likely to have implications for future cohorts of elderly.

The phonetic version of the COWAT seems to be resistant to the effects of normal aging but highly correlated with education and Verbal IQ. Axelrod and Henry, (1992) also found with 80 independently living adults age 50 through 89 years consistent performance with increasing age. Our findings also confirm results found by Ivnik et al., (1996). For the semantic version we did not detect a decline with increasing age, although there was a tendency for reduced performance on the vegetable version with increasing age. Lucas et al., (1998), however, found a significant negative correlation between age, education, and performance on the semantic version. However, these findings are not directly comparable as Lucas et al., averaged over 3 semantic categories (animal, vegetable, and fruit). These conflicting findings demonstrate that some categories may be more difficult than others.

The RMT has a number of advantages with elderly people in that it is less stressful or anxiety inducing as it relies on recognition memory as opposed to free-recall. However, the usefulness of the test for the elderly has been limited as the manual does not present norms beyond age 70. In addition, many memory tests are language specific. The RMT for faces has the benefit of being independent of verbal ability and can be administered in the standard format in any language (e.g., Diesfeldt & Vink, 1989, elderly dutch-speaking sample). However, we found a significant correlation between verbal IQ and face recognition ability, suggesting the test is not independent of intelligence. Similarly, Diesfeldt and Vink, (1989) found that face recognition among the very elderly was significantly correlated with visual intelligence, as measured by the Raven's Coloured Progressive Matrices, and moderately correlated with verbal intelligence. These authors suggested that other aspects of visual perception and complex reasoning may be involved in this task. In accordance with other authors we found that face

recognition ability declines with increasing age (Smith & Winograd, 1978; Ferris, Crook, Clark, McCarthy & Rae, 1980). In our sample recognition memory for words was better than for faces and was highly correlated with intelligence, but not so for age. Collectively our findings suggest that intelligence must be considered when interpreting performance.

There are some limitations with the norms as the sample is by no means a representative cross-section of the elderly population of New Zealand. This is evident from the occupational background of the sample, when they are compared with the whole nation, as in Table 1. There is a disproportionate number of subjects in the occupational groups *Professional* and *Technicians and Associate Professionals*. There is also a curious (for a Capital city) under-representation of *Legislators, Administrators and Managers*. Less surprising, given the urban nature of the catchment area, is the lack of people in the group *Agriculture and Fishery Workers*. However, these comparisons should be viewed with caution as the two distributions in Table 1 are not easily comparable. One represents the occupational structure of the New Zealand labour force in 1991. The other represents the occupations of an elderly retired sample who have nearly all been retired for at least the past decade. Hence any differences could arguably be due at least in part to real changes in employment patterns in New Zealand. Notwithstanding this caution, our sample does look to be over-represented in some occupational groups and under-represented in others.

Another feature of the sample which suggests that it is not likely to be perfectly representative of the entire elderly population, is the high average intellectual ability. Although we did not measure I.Q. as such, both the WAIS-R Vocabulary scale and the NART, are considered to provide acceptable estimates of verbal intelligence. Our sample had an average raw score on the Vocabulary scale of 54. This translates to a scaled score of 11 and of course this would be higher for the age scaled scores for the three groups. Similarly the NART suggested an average WAIS-R IQ of 108. Thus, it seems that the sample is slightly higher than expected in terms of general intellectual ability. This may well represent some form of sampling bias present in our data collection procedure. To begin with we confined our data collection to those elderly people living independently or semi-independently in their community. In addition, while attempts were made to sample from a broad geographic catchment area within the greater Wellington region, and to recruit from suburbs thought to represent differing socioeconomic sectors of the community, we were mostly reliant upon finding volunteer participants from various community groups. Thus there was probably a tendency to select from groups of elderly people who were already among the more physically active participants in their community. In this regard, they may represent a somewhat healthier than average sample of the elderly.

In addition to age, intelligence, cultural and ethnic background, health status is another probable predictor of cognitive performance among the elderly. 'Normal' aging is associated with increasing medical illnesses that may impact on tasks of daily living and cognitive functioning (e.g., hypertension, diabetes, cerebovascular disorders; Elias, Elias, & Elias, 1990). In addition, depression is also a common feature among the elderly and often accompanies or results from the many illnesses which the elderly are subject to. Thus health status must be taken into account when evaluating elderly people. The present study evaluated changes in cognition functioning that may be due to functional decline. Test performance was compared between elders who were living in the community and receiving support from an organised support agency (semi-independent) and between those elders who reported no community support (independent). Detectable differences were noted only for the phonetic verbal fluency task. These comparisons, however, need to be viewed with caution for two reasons. Firstly, the definition of 'living independently' can vary widely. Many elderly living in the community receive regular assistance from spouses, family, friends, and neighbours as well as community organisations for a variety of reasons other than declining cognitive functioning. Secondly, the fact that most elders receiving community support were single and in the 80+ age group suggests that the need for community support reflects the degree to which couples assist each other in the face of decreasing physical ability and increasing health problems rather than just cognitive decline. Malec, Ivnik, and Smith, (1993) have also found with the MOANS population studies that although their participants have many of the medical illnesses common among elderly, there is no evidence that their cognitive performance is adversely influenced by their health problems. Malec and colleagues argue that norms for elderly people need to be representative of the health profile of the elderly population and include in their studies participants with chronic controlled medical illnesses (e.g., diabetes, hypertension, cardiac problems) and exclude only those with active neurological or psychiatric disorders. We also examined self-reported history of depression and found no significant influences on cognitive functioning. In a sample of 403 elders Malec et al., have similarly found no evidence that prior neurological or psychiatric disorders (self-reported and medical-record) has any significant effect on cognitive functioning among elderly who are living independently in the population. It is our impression that some researchers, who have reported norms in the past, have been too rigorous in excluding elderly people with almost any health problems. Our findings and those of Malec et al., suggest that composing normative data for elderly people based only on a 'super healthy' sample is unwarranted.

In summary, the sample we obtained comprised 139 elderly New Zealanders living independently or semi-independently in the community. At the same time, they are unlikely to be a perfectly representative cross-section of the entire elderly population. However, given the almost total absence of any existing norms for older New Zealanders, we consider that the present sample represents a first step towards developing an adequate local normative database for older people.

We began this article by arguing for the importance

of using appropriate age-based norms for psychological tests. The data presented have highlighted another related issue, namely the need to also consider general intellectual ability, or IQ, in interpreting any one individual's test results. In the case of one of the tests we used, the COWAT (FAS), there was no correlation with age, but significant correlations with education and verbal IQ. Similarly, the correlations between scores on the GNT and the two estimates of verbal IQ (NART, Vocab.) were both larger than the GNT's correlation with age. This is not unexpected as we would not anticipate the same level of performance on the verbal fluency task from an 80-year-old with a verbal IQ of 90 as we would expect from another 80 year old with an IQ of 120. Indeed, a score which might be considered low average and expected for the former 80-year-old, might be considered borderline and suggesting impairment for the latter. The point is an obvious one but an important one nonetheless. Moreover, it may be true for other cognitive tests and not simply the ones for which we present data in this article. For example, Siegert and Cavana (1997) reported that both age and IQ predicted a significant proportion of the variance in performance among an elderly sample on the Trail-Making Test. Thus it is important when considering an individual's test results to not only compare them with peers of similar age, but also with peers of similar intellectual abilities and educational attainment. To some extent IQ has become an unfashionable concept in recent years (eg. Gould, 1981; Lezak, 1988), but clinical psychologists who reject the notion of IQ altogether may be in danger of throwing the baby out with the bath water.

References

Anastasi, A., & Urbina, S. (1997). *Psychological Testing*. Prentice-Hall International, Inc.

Benton, A. L., & Hamsher, K. deS. (1978). *Multilingual Aphasia Examination Manual.* Iowa City: University of Iowa.

Benton, A. L., & Hamsher, K. deS. (1989). *Multilingual Aphasia Examination.* Iowa City: University of Iowa.

Benton, A. L., Hamsher, K. deS., Varney, N. R., & Spreen, O. (1993). *Contributions to neuropsychological assessment.* New York: Oxford University Press.

Davis, P., McLeod, K., Ransom, M., & Ongley, P. (1997). *The New Zealand Socioeconomic Index of Occupational Status (NZSEI).* Wellington. Statistics New Zealand.

Diesfeldt, H., & Vink, M. (1989). Recognition memory for words and faces in the very old. *British Journal of Clinical Psychology, 28,* 247-253.

Elias, M. F. Elias, J. W. Elias, P. K. (1990). Biological and health influences on behavior. In J. E. Birren, K. W. Schaie, (Eds.). *Handbook of the psychology of aging (3rd ed.). The handbooks of aging* (pp 79-102).San Diego: Academic Press.

Ferris, S. H., Crook, T., Clark, E., McCarthy, M., & Rae, D. (1980). Facial recognition memory deficits in normal ageing and senile dementia. *Journal of Gerontology, 35,* 707-714.

Golden, C. J. (1978). *The Stroop Colour and Words Test (Manual).* Chicago: Stoelting.

Gould, S.J. (1981). *The Mismeasure of Man.* New York: Norton.

Grober, E., & Buschke, H. (1987). Genuine memory deficits in dementia. *Developmental Neuropsychology, 3,* 13-36.

Grober, E., & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology, 13,* 933-949.

Hactcher, L., & Stepanski, E. J. (1994). A Step-By-Step Approach to Using SAS® System for Univariate and Multivariate Statistics. Cary, NC: SAS Institute Inc.

Ivnik, R.J., Malec, J.F., Smith, G. E., Tangalos, E.G., & Petersen, R.C. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE Token, WRAT-R Reading, AMNART, Stroop, TMT, and JLO. *The Clinical Neuropsychologist, 10,* 262 - 278.

Ivnik, R.J., Malec, J.F., Smith, G. E., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1992a). Mayo's older Americans normative studies: Updated AVLT norms for ages 56 to 97. *The Clinical Neuropsychologist, 6,* 83 - 104.

Ivnik, R.J., Malec, J.F., Smith, G. E., Tangalos, E.G., Peterson, R.C., Kokmen, E., & Kurland, L.T. (1992b). Mayo's older Americans normative studies: WAIS-R norms for ages 56 to 97. *The Clinical Neuropsychologist, 6,* (Supplement), 1-30.

Ivnik, R.J., Malec, J.F., Smith, G. E., Tangalos, E.G., Peterson, R.C., Kokmen, E., & Kurland, L.T. (1992c). Mayo's older Americans normative studies: WMS-R norms for ages 56 to 94. *The Clinical Neuropsychologist, 6,* (Supplement), 49–82.

Ivnik, R.J., Malec, J.F., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1990). The Auditory Verbal Learning Test (AVLT) norms for 55 years and older. *Journal of Clinical and Consulting Psychology, 2,* 304-312.

Ivnik, R. J., Smith, G. E., Lucas, J. A., Tangalos, E. G., Petersen, R. C., & Kokmen, E. (1997). Free and Cued Selective Reminding Test: MOANS Norms. *Journal of Clinical and Experimental Neuropsychology, 19,* 676-691.

Jastak, S., & Wilkinson, G. S. (1984). *The Wide Range Achievement Test – Revised.* Administration Manual. Wilmington, DE: Jaskak.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1978). *The Boston Naming Test.* Boston: E. Kaplan & H. Goodglass.

Knight, R. G. (1997). The Wechsler Adult Intelligence Scale - Revised in clinical neuropsychology practice. *New Zealand Journal of Psychology, 26,* 2-19.

Lezak, M. D. (1987). Norms for growing older. *Developmental Neuropsychology, 3,* 1-12.

Lezak, M. D. (1988). IQ: R. I. P. *Journal of Clinical and Experimental Neuropsychology, 10,* 351-361.

Lezak, M. D. (1995). *Neuropsychological Assessment.* 3rd ed. Oxford University Press.

Louden, K. E. (1996). *Development of a New Zealand famous faces test.* M.A. (Applied) in Clinical and Community Psychology Thesis. Victoria University of Wellington, New Zealand.

Lucas, J. A., Ivnik, R. J., Smith, G. E., Bohac, D. L., Tangalos, E. G., Graff-Radford, N. R., & Petersen, R. C. (1998). Mayo's older American normative studies: Category fluency norms. *Journal of Clinical and Experimental Neuropsychology, 20,* 194-220.

Malec, J. F., Ivnik, R.J., Smith, G. E. (1993). In R. W. Parks, R. F. Zec, R. S. Wilson, (Eds.), *Neuropsychology of Alzheimer's Disease and Other Dementias* (pp 81-111). New York: Oxford University Press.

Malec, J. F., Ivnik, R.J., Smith, G. E., Tangalos, E.G., Peterson, R.C., Kokmen, E., & Kurland, L.T. (1992). Mayo's older Americans normative studies: Utility of corrections for age and education for the WAIS-R. *The Clinical Neuropsychologist, 6*, (Supplement), 31-47.

McKenna, P., & Warrington, E.K. (1983). *Graded Naming Test.* Windsor, U. K.: NFER - Nelson.

McKenna, P., & Warrington, E.K. (1980). Testing for nominal dysphasia. *Journal of Neurology, Neurosurgery, and Psychiatry, 43*, 781-788.

National Health Committee National Advisory Committee on Health and Disability. (1997). *Guidelines for the Support and Management of People with Dementia.*

Nelson, H. E., & Willison, J. R. (1992). *The National Adult Reading Test Manual. Second Edition.* Windsor, Berks: NFER – Nelson.

Newlove, D. (1992). *The Rey Auditory Verbal Learning Test: Norms for older New Zealanders.* M.A. (Applied) in Clinical and Community Psychology Thesis. Victoria University of Wellington, New Zealand.

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

Ryan, J. J., Paolo, A. M., and Brungardt, T.M. (1990). Standardization of the Wechsler Adult Intelligence Scale - Revised for Persons 75 Years and Older. *Psychological Assessment, 2*, 404-411.

Siegert, R. J. & Cavana, C. M. (1997). Norms for older New Zealanders on the Trail-Making Test. *New Zealand Journal of Psychology, 26*, 25 - 31.

Smith, A. D., & Winograd, E. (1978). Adult age differences in remembering faces. *Developmental Psychology, 14*, 38-44.

Spreen, O., & Strauss, E. (1991). *A compendium of neuropsychological tests: Administration, norms and commentary.* New York: Oxford University press.

Warrington, E. K. (1984). *Recognition Memory Test.* Windsor, U. K.: NFER - Nelson.

Warrington, E. K. (1997). The graded naming test: A restandardisation. *Neuropsychological Rehabilitation, 7*, 143-146.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised.* NewYork: The Psychological Corporation.

Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale – Third Edition.* San Antonio: The Psychological Corporation.

Acknowledgement:
This research was made possible by a grant from the Victoria University of Wellington Science Faculty Leave and Grants Committee. We are most grateful to Ms Jaana Montgomery for her assistance with typing.

Address for correspondence:
Dr Richard Siegert
School of Psychology
Victoria University of Wellington
PO Box 600, Wellington
Email: Richard.Siegert@vuw.ac.nz