Cognitive Flexibility in the Normal Elderly and in Persons with Dementia as Measured by the Written and Oral Trail Making Tests

Agata Kowalczyk, Skye McDonald, and Jacquelyn Cranney
University of New South Wales

Michael McMahon
C.R.A.G.S., Calvary Hospital, Sydney

This study investigated cognitive flexibility as indexed by the Written and Oral Trail Making Test (TMT) in sixteen persons with dementia and 60 normal elderly. Written and Oral TMT performances were significantly correlated with each other and with other tests of cognitive flexibility providing an index of the convergent construct validity of these tests. Part B of the Written TMT was influenced by psychomotor ability although this was diminished by the use of a ratio score B/A rather than simply Part B performance. The Oral test was not related to visual or motor skills, providing evidence for its divergent validity. Both parts of the Written and Oral TMT were found to be sensitive to cognitive decline in dementia. Performance on both Oral and Written TMT was also influenced by age, gender, education and intellectual ability. Previous norms (e.g., MOANS) for the Written TMT which are based on relatively educated elderly, tended to provide deflated scores for the community group tested here. Some additional, preliminary normative data for both tests were compiled which take into account the influences of each of these variables.

The Trail Making Test (TMT; Army Individual Test Battery, 1944) is a test of speed for visual search, attention, mental flexibility and motor function (Spreen & Strauss, 1991; LoSasso, Rapport, Axelrod & Reeder, 1998). The TMT consists of two subtests, Part A and Part B. Both subtests require simple motor/spatial skills (tracking) and basic sequencing abilities. It is assumed that Part B requires additional cognitive effort because the individual must mentally shift between two well-rehearsed sequences (numbers and letters). It has been suggested that when time to complete Part A is *relatively* much less than the time to complete Part B then the individual has difficulties in complex conceptual tracking (Lezak, 1995).

Dementia and TMT Performance

Clinicians have found that the TMT is an effective test for assessing progressive decline in dementia, even in the early stages of the disease (Greenlief, Morgolos & Erker, 1985; Lezak, 1995; Rasmusson, Zonderman, Kawas & Resnick, 1998). For example, Storandt, Botwinick, Danziger, Berg & Hughes (1984) found that a brief neuropsychological battery comprising of the Logical Memory and Mental Control subtests of Wechsler Memory Scale (WMS), Part A of the

We gratefully acknowledge Calvary Hospital for access to their patients. We also thank the research participants themselves for their time and effort and Dr. Kevin Bird for consultation on statistical analysis.

Address for Correspondence: Dr. Skye McDonald, School of Psychology, University of New South Wales, Sydney NSW 2052, Australia. E-mail: S.McDonald@unsw.edu.au

TMT and verbal fluency for letters S and P successfully discriminated persons suffering from mild senile dementia of the Alzheimer's type from healthy older persons. Tierney, Snow, Reid, Rorzitto & Fisher (1987) replicated and extended these findings by showing that the above neuropsychological battery accurately distinguished normal persons from persons with both mild senile dementia of the Alzheimer's type and other dementias.

Interpreting TMT Performance

When interpreting performance on the TMT researchers have looked at the time taken to complete Part A and Part B as well as the relationship between them. This has been defined as either a difference score (B–A) (Heaton, Nelson, Thompson, Burks & Franklin, 1985) or a ratio score B/A (Golden, Osmon, Moses & Berg, 1981). It has been argued that both measures are sensitive to the greater cognitive demands that it is assumed Part B makes (Corrigan & Hinkeldey, 1987; Lamberty, Putnam, Chatel, Bieliauskas & Adams, 1994). However, since performance on both parts and especially Part B of the TMT is affected by age, intelligence, education and possibly gender (Bornstein & Suga, 1988; Corrigan & Hinkeldey, 1987; Davies, 1968; Ernst, 1987; Gaudino, Geillser & Squires, 1995; Kennedy, 1981; Lamberty, Putnam, Rasmusson, Zonderman, Kawas & Resnick, 1998; Stanton, Jenkins, Savageau, Zyzanski & Avcoin, 1984; Stuss, Stathem & Poirer, 1987) the difference score (B–A) remains influenced by these variables. In contrast, the ratio of Part B to Part A (B/A) allows for greater control of such individual variability factors by effectively using the patient as his or her own control.

While the TMT has utility as a clinical measure for flexibility in the elderly it is confounded by "non-cognitive" performance factors such as primary visual and motor functioning (Schear & Sato, 1989) which deteriorate as part of normal biological ageing (Bradbury, 1991). Clinical interpretation of performance on the TMT is based on the assumption that Part B requires more complex cognitive processes than does Part A (Spreen & Strauss, 1991). However, this implies that the two Parts are equivalent in all respects (e.g. spatial arrangement of the circles) other than the addition of the more complex set-shifting component in Part B. But such a presumption has been challenged by assertions that apparent A-B differences could be due to factors such as increased demands in motor speed and visual search for Part B (Fossum, Holmberg & Reinvang, 1992; Gaudino et al., 1995; Woodruff, Mendoza, Dickson, Blanchard & Christenberry, 1995). Thus, the interpretation of differences in performance between Parts A and B remains controversial.

The Oral TMT

An oral paradigm of the TMT has been recently developed to provide a version that retains the sequential, multiple-conceptual-tracking component of the task while eliminating the dependence on intact vision and motor functioning (Ricker & Axelrod, 1994). This test shows similar changes across age groups to the Written TMT in both normal (Ricker & Axelrod, 1994) and mixed clinical populations (Abraham, Axelrod & Ricker, 1996). In addition, Ricker et al. (1996) found that the Oral TMT correlated significantly with other tests of executive functioning and was statistically unrelated to expressive language skills, thus providing some evidence for both convergent and divergent validity.

Normative Issues

The utility of clinical tests is critically reliant upon the adequacy of the normative base. The Oral TMT has no published norms. While the Oral TMT is associated with the Written TMT, this relationship is not exact and so the Written TMT norms are inappropriate as a reference. Furthermore, the few published norms for the written TMT are themselves inadequate for non-USA citizens. The norms of Davies (1968) have relatively large sample sizes but are quite old. More recent norms are based on either highly educated elderly (MOANS norms: Ivnik, Malec, Smith, Tangalos & Peterson, 1996) or elderly from managerial and professional occupations (Rasmusson et al., 1998). Such samples are not representative of the general community particularly in countries such as Australia, in which educational levels are overall much lower for the contemporary elderly. This is of particular importance given the significant influence that education plays on TMT performance (Bornstein & Suga, 1988).

Overview and Aims of this Study

In sum, the Written and Oral TMT have obvious applications in the assessment of cognitive flexibility in the elderly. There are, however, a number of issues which limit their current utility and which were addressed in this study.

Firstly, the extent to which differentially slow performance on Part B of the Written TMT represents purely cognitive factors is unclear. The development of the Oral TMT provides an opportunity to examine the cognitive and "non-cognitive" components that potentially influence performance on the Written version while simultaneously providing further construct validation of the Oral test. While correlation cannot be interpreted as causation, the relationships between performances on the TMT (Written and Oral) and related constructs (i.e. flexibility, alphabet recital, motor function and visual search) should highlight which processes may be considered potential contributors to performance on the two tests. This was one of the main aims of the current study.

Secondly, while it may be hypothesised that the Oral TMT, like the written version, is sensitive to dementia this has not been empirically established. A further aim of this study was to investigate the performance of subjects with dementia on the Oral TMT compared to non-demented elderly.

Thirdly, in order to examine the adequacy of TMT norms such as the MOANS for non-USA residents, the performance of a group of Australian elderly from the general community on the Written TMT was compared to the MOANS normative sample (Ivnik et al., 1996) and was examined for the influences of age, gender and education. Finally, a new set of preliminary norms for less educated elderly on both the Written and Oral versions of the TMT were constructed encompassing scores on Parts A and B as well as the two types of difference scores (A–B) and (A/B).

Method

Participants

The sample comprised of 76 elderly people living in South Eastern Sydney aged between 55 and 89 years with a mean age of 77 years (see Table 1).

Dementia group. Sixteen participants (2 males and 14 females) were individuals diagnosed with dementia. Fifteen of the persons with dementia were referred for a neuropsychological assessment to Calvary Hospital and received a positive diagnosis of dementia based upon their neuropsychological performance, medical and psychological history and medical tests. One dementia patient was identified as having dementia during treatment at an outpatient clinic at Calvary Hospital. The dementia group included patients who received the diagnosis of Alzheimer's Disease, vascular dementia and dementia of a mixed type. A diagnostically mixed sample was chosen to reflect the variety of patients commonly seen in clinical practice.

Normal elderly group. A further 65 elderly people were assessed for inclusion in the study. These included elderly patients attending a Podiatry Outpatient Clinic and spouses of people with dementia referred for neuropsychological assessment. The Podiatry Clinic was considered an appropriate place for recruitment of participants as large numbers of independently living elderly people with varying levels of general health make use of the facility. In addition, one 77-year-old person included in the normal sample had initially been referred for neuropsychological assessment but performed at an average to above average level on all tests of cognitive functioning including the TMT tasks. Her MMSE score was 29/30 and she was cleared of any diagnosis of dementia.

In order to be included in the normal elderly sample, participants fulfilled similar criteria to those used in the MOANS research (Ivnik et al., 1996) (i.e. they had to be aged 55 or over, living independently in the community with no *current* central nervous system or psychiatric conditions, no complaint of cognitive difficulty during history taking and no suggestion of physical problems that may affect cognition). Chronic medical illness was not an exclusion criterion unless it was sufficiently severe to potentially affect cognition. Their medical, neurological and psychiatric histories were elicited via a semi-structured interview. Using these criteria 5 participants were excluded because they were identified as having a confounding condition (e.g., borderline premorbid intellectual ability, illiteracy, severe health problems, severe depression, partial blindness and possible head injury).

The remaining 60 normal elderly participants were subdivided into three groups (each $N = 20$) according to their age: (1) 55 to 74 years old; (2) 75 to 83 years old; and (3) over 84 years old. These divisions were dictated by an initial attempt to replicate MOANS age groupings, but ultimately by the availability of participants.

Materials

All participants underwent the screening interview and a battery of psychological tests designed to provide (1) background information (premorbid intelligence, current cognitive functioning and levels of depressed mood) (2) general cognitive abilities (working memory, verbal and nonverbal learning) (3) cognitive flexibility (both TMT tasks and two independent measures) and (4) alphabet knowledge, motor and visual functions. The screening interview and test battery took approximately 45 to 60 minutes to complete. The battery included:

1. Background information: The Nelson Adult Reading Test (NART, Nelson, 1982), the Mini Mental State Examination (MMSE, Folstein, Folstein & McHugh, 1975), the Geriatric Depression Scale (GDS, Yesavage, Brink, Rose et al., 1983)
2. General cognitive abilities: Mental Control (Wechsler, 1987); Logical Memory I and II (Wechsler, 1987), Visual Reproduction I and II (Wechsler, 1987).
3. Cognitive flexibility: Written Trail Making Test Part A and Part B (Spreen & Strauss, 1991), Oral Trail Making Test (Ricker & Axelrod, 1994), Controlled Oral Word Association Test (COWAT, Spreen & Strauss 1991) and Digit Symbol subtest of the WAIS-R (Wechsler, 1981).
4. Tests of alphabet recital, motor and visual functions: Alphabet recital (Wechsler, 1987) Finger Tapping Test (Reitan & Davison, 1974), Visual Search Test (Lezak, 1995).

Procedure

All participants were tested individually and were allowed to work at their own pace on all tasks. All neuropsychological tests were administered following standard instructions. The Oral Trail Making Test was administered as follows: *Part A:* The participant was told “I would like you to count from 1 to 25 as quickly as you can — 1,2,3 and so on. Are you ready? Begin”. Time taken to complete this task was recorded. *Part B:* The participant was told “ Now I would like you to count again, but this time you are to switch between number and letter, so you would say 1-A-2-B-3-C and so on, until I say STOP. Are you ready? Begin.” The participant was then stopped once he/she reached 13 and the time recorded.

Assessment began with the collection of demographic information and the GDS, followed by the MMSE, Logical Memory I, Visual Reproduction I, TMT (Oral or Written), Mental Control, NART, TMT (alternative version), Digit Symbol, COWAT, Finger Tapping Test, Visual Search Task, Logical Memory II, Visual Reproduction II and the semi-structured interview. Although order effects were not found in previous studies (Abraham et al., 1996), the order of administration of the Oral and Written TMT were counterbalanced across participants. Participants who were unable to complete Part B of the Oral or Written TMT within 5 minutes were stopped and redirected to the next task.

Results

Background Characteristics of the Group

Background characteristics of the sample are depicted in Table 1. A series of *t* tests comparing all of the controls with the dementia group failed to reveal any differences in age, $t(74) = -1.70$, or education, $t(74) = 1.30$. Nor were there any differences between the groups in terms of the proportions of males to females ($\chi^2 = 0.95$). There were two left-handed participants in the control group, but their performance did not differ significantly from others in their corresponding age groups. IQ estimates in both groups were higher than would normally be expected given the generally low average level of education. This could reflect either the insensitivity of the NART, or different schooling opportunities and policies in Australia during the first half of the nineteenth century compared to contemporary practice. There was no difference between groups in estimated IQ, $t(74) = 1.43$. While no subject was identified as depressed on semi-structured interview, two of the dementia subjects and five of the normal elderly reported depressed mood (i.e. a score above 13) on the GDS. In the normal elderly 4 of these scores were in the mild range (13–15). There was no difference between the groups in terms of overall GDS scores, $t(74) = 1.23$. As expected dementia sufferers performed significantly more poorly on the MMSE , $t(74) = -4.037$, $p < .001$.

Neuropsychological Characteristics

Performance on the remaining standard neuropsychological tests for the dementia participants and for the normal elderly is shown in Table 2. Between-group differences were analysed using one-way analysis of variance (ANOVA). The dementia group performed poorly on all measures of general cognitive ability: Mental Control, $F(1,74) = 4.98$, $p < .05$, Logical Memory I, $F(1,74) = 86.66$, $p < .01$, Logical Memory II, $F(1,71) = 69.65$, $p < .01$, Visual Reproduction I, $F(1,74) = 25.82$, $p < .01$, Visual Reproduction II, $F(1,71) = 55.01$, $p < .01$. The obtained differences between dementia and control subjects remained significant when age, education, IQ and gender were entered as covariates. The people with dementia also performed significantly more poorly than the normal elderly on Digit Symbol, $F(1,68) = 7.82$, $p < .01$, and COWAT, $F(1,72) = 5.44$, $p < .05$, but their performance did not differ significantly to control subjects on alphabet recital, finger tapping, and visual search ($p > .05$).

TABLE 1
Demographic Characteristics of the Normal and Dementia Groups

Characteristic		Dementia (n = 16)	Controls (n = 60)
Gender			
	Females	14	43
	Males	2	17
Handedness			
	Right	16	58
	Left	0	2
Age (years)			
	M	74.25	77.8
	SD	6.73	7.57
	Range	55–83	59–89
Education (years)			
	M	8.81	8.11
	SD	2.32	1.82
	Range	6–16	5–15
Estimated IQ			
	M	105.00	107.55
	SD	5.72	6.48
	Range	95–114	91–117
Depression			
	M	7.75	6.12
	SD	5.07	4.56
	Range	0–20	0–22
Current cognitive function			
	M	24.56**	28.20
	SD	3.60	1.38
	Range	17–30	25–30

Note. ** $p < .01$ for comparisons of people with dementia to all control.

TABLE 2
Neuropsychological Test Performance by Normal Controls and Participants With Dementia

Neuropsychological Tests	Dementia *M (SD)*	Controls *M (SD)*
General Cognitive Function		
Mental Control	4.94* (1.24)	5.48 (0.75)
Logical Memory I	10.12** (7.14)	26.03 (5.77)
Logical Memory II	5.06** (5.99)	21.00 (6.94)
Visual Reproduction I	18.75** (7.44)	28.80 (6.92)
Visual Reproduction II	3.81** (8.36)	22.61 (9.11)
Cognitive Flexibility		
Digit Symbol	25.67** (10.47)	33.84 (9.91)
COWAT	26.75 * (12.42)	34.59 (11.75)
Basic Language, Motor and Visual Functions		
Alphabet Recital	9.71 (5.73)	7.95 (4.35)
Finger Tapping	38.37 (9.16)	40.48 (8.99)
Visual Search	20.87 (2.10)	20.10 (0.77)

Note. * $p < .05$; ** $p < .01$ for comparisons of people with dementia to all controls.

Oral TMT: Construct Analyses

Pearson product-moment correlations between the Oral and Written TMTs and Digit Symbol, COWAT, the ability to say the alphabet, finger tapping and visual search are detailed in Table 3. Both parts of the Oral TMT as well as the derived part scores (B–A and B/A) were moderately correlated with other tests of cognitive flexibility (both parts of the Written TMT, Digit Symbol and COWAT). These results support the convergent validity of the Oral TMT as a measure of cognitive flexibility. The Oral TMT was also moderately correlated with the efficiency of alphabet recital. Unexpectedly, Part A was moderately correlated with both visual search and finger tapping. However, the other indices (Part B, B–A, B/A) were not correlated, thus supporting the independence (i.e. divergent validity) of the Oral TMT from these basic motor and visual functions.

Moderate correlations were also obtained between the Written TMT and other tests of cognitive flexibility, supporting prior assumptions of convergent validity. The test (except B/A) was also moderately correlated with alphabet recital and finger tapping. Correlations with visual search were not significant.

Fisher *z* transformations were carried out to test the significance of differences between correlation coefficients for the Written versus Oral versions of the test. These analyses revealed that Digit Symbol was significantly more strongly correlated ($p < 0.05$) with both parts of the Written version compared to the Oral counterparts, $t(73) = 2.78$ for Part A; $t(73) = 3.86$ for Part B. This may reflect the common psychomotor component in Written TMT and Digit Symbol. This was not the case for the relative scores (B-A and B/A) obtained from the Written and Oral versions. Using these scores the two versions of the TMT showed a similar degree of association with Digit-Symbol. Significantly higher correlations were also obtained between COWAT and Part A (but not Part B) of the Written compared to Oral TMT, $t(73) = 4.01$ for Part A. Again, the B–A and B/A correlation patterns reduce the apparent differences between the two tests. These results suggest that the use of difference *or* ratio scores reduces differences between the constructs measured by the alternative versions of TMT.

No significant differences on Part A or B were obtained for correlations between Written and Oral TMT and alphabet recital, indicating that this automatic speech ability was equally associated with performance on both forms of the TMT. It was expected that finger tapping would correlate more strongly with both parts of the Written TMT compared to the Oral TMT. This was found for Part B, $t(73) = 2.089$, but not Part A. The significant correlation for Part B may reflect the possibly greater involvement of the frontal lobe in both Part B and finger tapping, compared to Part A.

Normal Elderly and People with Dementia on TMT

Performances on Oral and Written TMTs by persons with dementia, by all the normal control subjects and by each of the three different age groups are shown in Table 4. Analyses of variance indicated that the dementia group performed significantly slower than the control group on Part A of the Oral TMT, $F(1,74) = 10.76$, $p < .05$, Part B of the Oral TMT, $F(1,74) = 31.73$, $p < .01$, Part A of the Written TMT, $F(1,74) = 9.90$, $p < .05$, Part B of the Written TMT, $F(1,74) = 25.69$, $p < .01$, as well as on the Oral TMT difference, $F(1,74) = 30.68$, $p < .01$, and ratio, $F(1,74) = 21.17$, $p < .01$, scores and Written TMT difference, $F(1,74) = 22.29$, $p < .01$, and ratio $F(1,74) = 12.33$, $p < .01$, scores. These results indicate that both parts of the Written and Oral TMTs are sensitive to cognitive decline in dementia.

MOANS Normative Data

Normal elderly participants' raw scores were converted to MOANS standard scores and percentile ranks. The mean standard scores for the normal elderly in this study were 7.9 on Part A and 8.3 on Part B compared to the expected mean of 10 developed using MOANS populations. To ensure that sub-clinical depression (as indicated by low mood on the GDS) was not a causal factor in the low mean scores of the normal elderly relative to the MOANS sample, the means were also calculated for the normal group *excluding* the five subjects with depressed mood scores. These means were essentially identical, i.e. 8.0 on Part A and 8.3 on Part B. The number of participants in each MOANS percentile rank is graphically represented in Figure 1. Visual inspection of Figure 1 Indicates that the mid-points for both Part A and Part B fall between the 11th and 28th percentile. The mean standard scores of the two populations were compared using one-sample t-tests. These analyses confirmed that the MOANS sample and the sample employed in the present study have different means ($p < .05$) on Part A, $t(59) = -6.18$, and Part B, $t(59) = -4.33$, of the Written TMT. The variances of the two samples were not statistically different.

Influence of Age, Education, IQ and Gender

In order to examine differences on the different versions of the TMT according to age, education, IQ and gender, participants were subdivided into groups based on these characteristics and this is depicted in Table 5. In order to empirically determine the relationship between these variables and

TABLE 3

Correlations Between the Oral and Written TMTs and Independent Tests of Flexibility, Basic Language, Motor and Visual Functions

	Oral TMT				Written TMT			
Tests	Part A	Part B	B–A	B/A	Part A	Part B	B–A	B/A
Written TMT – A	.430**	.429**	.417**	.259*				
Written TMT – B	.354**	.611**	.605**	.558**				
Written TMT – B–A	.260*	.573**	.571**	.576**				
Written TMT – B/A	.012	.364**	.370**	.494**				
Digit Symbol	–.294**	–.511**	–.506**	–.471**	–.577**	–.767**	–.695**	–.397**
COWAT	–.256*	–.434**	–.430**	–.370**	–.447**	–.540**	–.479**	–.250*
Efficiency of alphabet	.441**	.408**	.396**	.287**	.284**	.265*	.210*	.072
Finger Tapping	–.405**	–.108	–.093	.019	–.342**	–.312**	–.257*	–.081
Visual Search	.212*	.094	.086	–.063	.173	.176	.145	–.005

Note. * $p < .05$; ** $p < .01$.

TABLE 4
Means and Standard Deviations for Oral and Written TMT Performance (sec.) for Each Subject Group

			Group		
Neuropsychological Tests	Dementia (*N* = 16) *M (SD)*	All Controls (*N* = 60) *M (SD)*	55–74 years (*N* = 20) *M (SD*	75–83 years (*N* = 20) *M (SD)*	84 + years (*N* = 20) *M (SD)*
Written TMT – Part A	81.52* (39.57)	58.77 (20.75)	48.19 (18.75)	60.22 (20.95)	67.88 (18.43)
Written TMT – Part B	256.66** (76.21)	153.39 (71.41)	113.98 (54.47)	144.45 (65.83)	201.74 (65.99)
Written TMT differencescore	175.14** (71.57)	94.63 (57.49)	65.79 (44.85)	84.23 (48.79)	133.86 (57.11)
Written TMT ratio score	3.65** (1.65)	2.61 (0.83)	2.42 (0.85)	2.38 (0.56)	3.04 (0.90)
Oral TMT – Part A	11.08* (3.33)	8.59 (2.51)	7.65 (2.29)	9.68 (2.48)	8.42 (2.44)
Oral TMT – Part B	144.78** (112.49)	51.82 (33.14)	41.63 (25.52)	55.85 (37.73)	57.98 (34.20)
Oral TMT difference score	133.71** (111.15)	43.23 (32.96)	33.98 (25.38)	46.16 (37.56)	49.56 (34.32)
Oral TMT ratio score	13.41** (9.57)	6.30 (3.80)	5.66 (3.04)	6.00 (3.99)	7.24 (4.27)

Note. * $p < .05$; ** $p < .01$

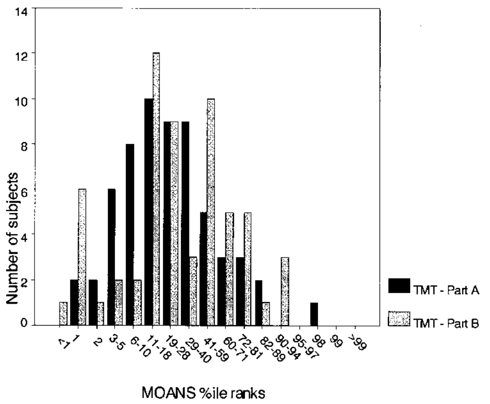

FIGURE 1
Number of control subjects in MOANS percentile ranks.

TMT performance, correlations were performed and the results of these are summarised in Table 6.

As seen in Table 6, age was moderately correlated with all indices of the TMT except Part A and B/A of the Oral TMT. This relationship remained unchanged when statistically controlling for education, estimated IQ and gender. Inspection of Table 4 reveals that TMT performance was slower for the oldest group relative to the younger two groups. The obtained results are consistent with previous research indicating that the total time to complete the TMT (especially Part B) is significantly longer for the older participants than for the younger ones. Fisher z transformations revealed that both parts of the Written TMT were significantly more strongly correlated with age compared to both parts of the Oral TMT, $t(57) = -1.89$, for Part A and $t(57) = -3.52$ for Part B. This supports findings that as people get older their psychomotor speed deteriorates and so their performance on the Written TMT is further affected. However, the B/A results suggest that the written TMT may be more age-sensitive than the Oral TMT, independent of motor function.

The control group was divided into those who had completed the minimal amount of secondary schooling (9 years of education or more) and those who had not (see Table 5). Level of formal education was moderately correlated with all indices of the TMT except Part A of both Oral and Written TMT. In Table 5 it is apparent that individuals with less than 9 years of formal education performed slower on Part B of the Oral and Written TMT than individuals with 9 or more years of formal education. This confirms the earlier comparison with MOANS norms. Less educated individuals are slower than their more educated counterparts especially on Part B of TMT.

Consistent with previous research, estimated pre-morbid IQ was associated with performance on TMT. All indices except Part A of the Oral TMT showed a moderate correlation. As can be seen in Table 5, subjects with an estimated Average IQ performed slower on Part B of the Oral and Written TMT than subjects with estimated High Average or Superior IQ.

Gender was also correlated with TMT performance. Specifically, as can be seen in Table 5, male subjects performed better than female subjects on Part A of the Oral TMT and Part B of the Written TMT.

TABLE 5
Means and Standard Deviations on TMTs by Education, IQ and Gender

Subgroup	Oral TMT (sec.) Part A	Oral TMT (sec.) Part B	Oral TMT (sec.) B–A	Oral TMT (sec.) B/A	Written TMT (sec.) Part A	Written TMT (sec.) Part B	Written TMT (sec.) B–A	Written TMT (sec.) B/A
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Education								
Less than 9 yrs (*n* = 28)	8.55 (2.25)	61.14 (36.08)	52.59 (35.84)	7.38 (4.16)	60.01 (20.86)	171.01 (70.84)	111 (56.78)	2.88 (0.88)
9 or more yrs (*n* = 32)	8.63 (2.80)	41.86 (26.84)	33.23 (26.69)	5.15 (3.04)	57.43 (20.92)	134.56 (68.25)	77.13 (53.81)	2.33 (0.68)
Estimated IQ								
Average (*n* = 36)	8.78 (2.66)	63.97 (37.31)	55.19 (37.05)	7.6 (4.21)	61.39 (21.69)	176.68 (75.36)	115.28(58.75)	2.87 (0.79)
High to Superior (*n* = 24)	8.30 (2.28)	33.60 (11.05)	25.3 (11.78)	4.35 (1.88)	54.82(19.02)	118.46 (48.14)	63.64 (39.45)	2.23 (0.75)
Gender								
Female (*n* = 43)	9.05 (2.66)	54.89 (35.62)	45.84 (35.57)	6.40 (4.03)	59.57 (22.07)	166.05 (73.98)	106.48 (58.3)	2.8 (0.81)
Male (*n* = 17)	7.41 (1.62)	44.06 (25.11)	36.64 (24.91)	6.04 (3.26)	56.73 (17.41)	121.37 (54.03)	64.63 (43.96)	2.16 (0.72)

TABLE 6
Correlations Between the Oral and Written TMTs and Age, Education, IQ and Gender

Characteristic	Oral TMT Part A	Oral TMT Part B	Oral TMT B–A	Oral TMT B/A	Written TMT Part A	Written TMT Part B	Written TMT B–A	Written TMT B/A
Age	.171	.248*	.237*	.193	.464**	.578**	.550**	.354**
Education	.059	–.347**	–.353**	–.370**	.048	–.244*	–.321**	–.437**
IQ	–.066	–.717**	–.716**	–.653**	–.334**	–.491**	–.488**	–.336**
Gender	–.296*	–.149	–.127	–.043	–.062	–.284*	–.331**	–.350**

Note. * $p < .05$; ** $p < .0$

Discussion

The present study compared the performance of a group of people with dementia with a group of normal elderly on two versions of the Trail Making Tests. The dementia group had received a formal diagnosis of dementia on the basis of neuropsychological, neurological and medical grounds. In contrast, the normal elderly were screened for any identifiable psychiatric or neurological condition, obvious signs of cognitive decline and for general health. The absence of dementia within this group was supported by the finding that all participants scored above the standard cut-off of 24 on the MMSE. However, it should be noted that two of the normal elderly had scores of 25 and six had scores of 26. Debate as to the most useful cut-off score for screening dementia suggests scores between 23 (Brooke & Bullock, 1999; Murden, McRae, Kaner & Bucknam, 1991) and 26 (van Gorp, Marcotte, Sultzer, Hinkin, Mahler & Cummings, 1999) for highly educated groups. However, a study specifically examining the effects of low education (8 years or less) reported significantly lower MMSE scores for low education and recommended a cut-off of 17 (Murden et al., 1991). The normal elderly who obtained scores of 25 to 26 in this sample had educational levels between 5 and 7 years. It is therefore reasonable to assume, on the basis of both the structured interview and the MMSE results that our group of normal elderly were just that, with no members suffering from undetected dementia. This is, however, instructive concerning the importance of using *both* age and education appropriate norms when assessing the elderly.

While none of the participants were identified as depressed on interview it is noteworthy that 4 of the normal elderly and one of the dementia patients reported mildly depressed mood on the GDS and one normal elderly and one person with dementia scored in the severely depressed range on this measure. The inclusion of some individuals with (mainly mild) symptoms of depressed mood within the normal group suggests some may have suffered from sub-clinical depression. This is consistent with the prevalence of depression in the elderly community. Mood scores did make a small but significant contribution to the variance of the written TMT, Part B scores (r = .24, shared variance = 6%) but not to either part of the oral TMT. It is unclear the extent to which other normative studies such as the MOANS included those with sub-clinical depression. Our normative group was selected using identical criteria to the MOANS studies although the fact that this group was interviewed by the (clinically trained) researcher rather than relying upon their physician's account may have lead to a slightly different interpretation of the criteria. But more significantly, unlike this study, the MOANS study did not report formal measures of mood and so the extent of mood disturbance in the MOANS group is unknown. Regardless of whether the two groups differed in this regard, the removal of subjects in this sample who had elevated scores on the GDS made little impact on the mean of their MOANS standard scores.

This study was illuminating in terms of the construct validity of both the Written and Oral TMT. As predicted, both versions were associated with independent measures of cognitive flexibility. This is consistent with earlier studies of the construct validity of the Oral TMT as a measure of cognitive flexibility in stroke (Ricker et al., 1996) and mixed clinical populations (Abraham et al., 1996). The Written TMT was associated with not only cognitive flexibility, but also psychomotor ability. The Written TMT was not influenced by variability in visual search skills. The use of the difference scores, especially the B/A ratio, appeared to eliminate this psychomotor component and may therefore be the most useful Written TMT score for measuring cognitive flexibility *per se.* Part B and the difference scores of the Oral TMT were not related to visual or motor function and can be considered a useful alternative test of cognitive flexibility, especially for people for whom vision and/or motor deficits represent significant confounding influences.

The finding that correlations between related and unrelated constructs were generally in the expected directions was encouraging. Correlations between neuropsychological measures are rarely robust. Yet specific Fischer Z transformation tests designed to analyse the magnitude of differences between correlation coefficients confirmed that measures of convergent validity were generally more highly correlated than measures of divergent validity. This does not preclude the possibility that positive correlations were due to other, irrelevant sources of variance. The construct validity of both forms of TMT would be further strengthened if these results were replicated using other measures of both theoretically related and unrelated constructs.

Both forms of TMT were sensitive to cognitive decline in a sample of demented elderly. It should also be noted that there was no difference in the pattern of dementia versus control group results for the various Part A, Part B, B–A, B/A indices of the Oral and Written TMT.

The present results also demonstrate that MOANS normative data for the Written TMT are inappropriate for individuals with lower levels of education. In the MOANS sample (Ivnik et al., 1996) most participants had completed 12 to 15 years of formal education, whereas in the present cohort the average education was 8 years (there are *no* education-level breakdowns in the MOANS report). Furthermore, this lower level of education is typical of contemporary elderly people in many communities. Written TMT performance of the present sample was, on average, significantly poorer than that of the MOANS's sample. Consequently, when using MOANS normative data with this less educated cohort, the majority of individuals would be allocated to lower percentile ranks and possibly described as having poor problem-solving skills without taking into account this education factor. There is a clear need for TMT norms that are stratified by age, education level, IQ and gender.

In line with previous research (Abraham et al., 1996; Bornstein & Suga, 1988; Kennedy, 1981; Stanton et al., 1984), the present study emphasises the adverse impact of age on both Oral and Written TMT performance, in particular Part B. The influence of education and intelligence was also confirmed. Individuals with fewer years of formal education performed significantly slower on both Oral and Written TMT than their more educated counterparts. Similarly, those with higher intellectual ability (estimated IQ) performed faster on the Oral and Written TMT than those with lower intellectual ability.

The results revealed that gender influences Oral and Written TMT performance. Males performed significantly better than females on Part A of the Oral TMT and Part B of the Written TMT. Although the two groups did not significantly differ in their performance on the other parts of TMTs, there was a trend towards males being faster. This is contrary to previous studies reporting a female advantage on Written TMT performance (e.g. Gaudino et al., 1995). Research with larger numbers of males is needed to investigate further the reliability of gender differences in Oral and Written TMT performance.

The present study demonstrated that age, gender, education and intellectual ability must be considered when interpreting both Oral and Written TMT performance. This study enhances the normative database for both tests by providing some additional data on performance of normal elderly, taking into account the above characteristics. Further studies with larger numbers of subjects are needed to establish extensive normative data for these two tests of cognitive flexibility. Moreover, the present study employed a diagnostically mixed sample of people with dementia. Future research could determine whether there are differences in Oral TMT performance amongst groups of people with differing types of dementia.

References

Abraham, E., Axelrod, B.N., & Ricker, J.H. (1996). Application of the Oral trail making test to a mixed clinical sample. *Archives of Clinical Neuropsychology*, *11*, 697–701.

Army Individual Test Battery (1944). Manual of directions and scoring. Washington, DC: War Department, Adjutant General's Office.

Bornstein, R.A. & Suga L.J. (1988). Educational level and neuropsychological performance in healthy elderly subjects. *Developmental Neuropsychology*, *4*, 17–22.

Bradbury, N. (1991). Problems of elderly people. In Dryden, W. & Rentoul (Eds). *Adult Clinical Problems*. London: Routledge.

Brooke, P. & Bullock, R. (1999) Validation of a 6 item cognitive impairment test with a view to primary care useage. *International Journal of Psychiatry, 14,* 936–40.

Corrigan J.D. & Hinkeldey N.S. (1987). Relationships between parts A and B of the trail making test. *Journal of Clinical Psychology*, *43*, 402–409.

Davies, A.D.M. (1968). The influence of age on Trail Making Test performance. *Journal of Clinical Psychology*, *24*, 95–98.

Ernst, J. (1987). Neuropsychological problem-solving skills in the elderly. *Psychology and Aging*, *2*, 363–365.

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.

Fossum, B., Holmberg, H., & Reinvang, I. (1992). Spatial and symbolic factors in performance on the Trail Making Test. *Neuropsychology*, *6*, 71–75.

Gaudino, E.A., Geisler, M.W., & Squires, N.K. (1995). Construct validity in the trail making test: What makes part B harder? *Journal of Clinical and Experimental Psychology*, *17*, 529–535.

Golden, C.J., Osmon, D.C., Moses, J.A., & Berg, R.A. (1981). *Interpretation of the Halstead-Reitan Neuropsychological Test Battery*. New York: Grune & Stratton.

Greenlief, C.L., Margolis, R.B., & Erker, G.J. (1985). Application of the trail making test in differentiating neuropsychological impairment of elderly persons. *Perceptual and Motor Skills*, *61*, 1283–1289.

Heaton, R.K., Nelson, L.M., Thompson D.S., Burks, J.S., & Franklin G.M. (1985). Neuropsychological findings in relapsing-remitting and chronic-progressive multiple sclerosis. *Journal of Consulting Clinical Psychology*, *53*, 103–110.

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

Kennedy, J.K. (1981). Age effects on trail making test performance. *Perceptual and Motor Skills*, *52*, 671–675.

Lamberty, G.J., Putnam, S.H., Chatel, D.M., Bieliauskas, L.A., & Adams, K.M. (1994). *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *7*, 230–234.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.) New York: Oxford University Press.

LoSasso, G. L., Rapport, L. J., Axelrod, B. N., & Reeder, K. P., (1998). Intermanual and alternate-form equivalence on the trail-making tests. Journal of Clinical and Experimental Neuropsychology, 20, 107–110.

Murden, R.A., McRae, T.D., Kaner, S., & Bucknam, M.E. (1991) Mini-state examination exam scores vary with education in blacks and whites. *Journal of the American Geriatrics Society, 39,* 149–55.

Nelson, H.E. (1982). *National Adult Reading Test (NART). 2nd ed. Test manual*. England: NFER-Nelson.

Rasmusson, D. X., Zonderman, A. B., Kawas, C., & Resnick, S. M. (1998). Effects of age and dementia on the Trail Making Test. *The Clinical Neuropsychologist*, 12, 169–178.

Reitan, R.M. & Davison, L.A. (1974). *Clinical neuropsychology: current status and applications*. New York: Hemisphere.

Ricker, J.H. & Axelrod, B.N. (1994). Analysis of an oral paradigm for the trail making test. *Assessment*, *1*, 47–51.

Ricker, J.H., Axelrod, B.N., & Houtler, B.D. (1996). Clinical validation of the oral trail making test. *Neuropsychiatry, Neuropsychology and Behavioral Neurology*, *9*, 50–53.

Schear, J.M. & Sato, S.D. (1989). Effects of visual acuity and visual motor speed and dexterity on cognitive test performance. *Archives of Clinical Neuropsychology*, *4*, 25–32.

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

Stanton, B.A., Jenkins, C.F., Savageau, J.A., Zyzanski, S.J., & Avcoin, R. (1984). Age and educational differences on the trail making test and Wechsler memory scales. *Perceptual and Motor Skills*, *58*, 311–318.

Storandt, M., Botwinick, J., Danziger, W., Berg, L., & Hughes, C.P. (1984). Psychometric differentiation of mild senile dementia of the Alzheimer type. *Archives of Neurology*, *41*, 497–499.

Stuss, D.T., Stethem, L.L., & Poirer, C.A. (1987). Comparison of three tests of attention and rapid information processing across six age groups. *The Clinical Neuropsychologist*, *1*, 139–152.

Tierney, M.C., Snow, G., Reid, D.W., Rorzitto, M.L., & Fisher, R.H. (1987). Psychometric differentiation of dementia. *Archives of Neurology*, *44*, 720–722.

van Gorp, W.G., Marcotte, T.D., Sultzer, D., Hinkin, C., Mahler, M., & Cummings, J.L. (1999) Screening for dementia: comparison of three commonly used instruments. *Journal of Clinical and Experimental Neuropsychology, 21,* 29–38.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale — Revised. Manual*. New York: Psychological Corporation.

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

Woodruff, G.R., Mendoza, J.E., Dickson, A.L., Blanchard, E., & Christenberry, L.B. (1995). The effects of configural differences on the Trail Making Test. *Archives of Clinical Neurology*, *10*, 408.

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