The Oral Trail Making Test: Effects of Age and Concurrent Validity[†]

Marty Mrazik [1],*, Scott Millis [2], Daniel L. Drane [3,4]

[1]*University of Alberta, Edmonton, Alberta, Canada*
[2]*Department of Physical Medicine & Rehabilitation, Wayne State University School of Medicine, Detroit, MI, USA*
[3]*Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA*
[4]*Department of Neurology, Regional Epilepsy Center, University of Washington, School of Medicine, Seattle, Washington DC, USA*

*Corresponding author at: University of Alberta, 6-135 Education Building North, Edmonton, AB, Canada T6G 2G5. Tel.: +1-780-492-8052; fax: +1-780-492-1318.
E-mail address: mrazik@ualberta.ca (M. Mrazik).

Abstract

The oral version of the Trail Making Test (OTMT) is a neuropsychological measure that provides an assessment of sequential set-shifting without the motor and visual demands of the written TMT (WTMT). Originally purposed to serve as an oral analog of the WTMT, the OTMT provides a means to evaluate patients with physical restrictions. However, formal validity studies and available normative data remain sparse. In a sample of healthy adults ($n = 81$), a strong correlation was observed between OTMT-B and its written counterpart ($r = .62$), but the correlations were weak between OTMT-A and either written version of the TMT. OTMT-B was significantly correlated with age but not with education or gender, whereas OTMT-A was not significantly correlated with demographic factors. The WTMT to OTMT ratios observed in the current study were generally lower than previously reported and varied across age groups, suggesting that the recommended use of a uniform conversion factor to predict one performance based on the other should be cautiously undertaken. Normative data that have been stratified by age are provided as well as suggestions for using both versions of the TMT in tandem to better elucidate the nature of cognitive deficits and to aid in the localization of cerebral dysfunction.

Keywords: Neuropsychology; Normative data; Cognitive tests; Trail Making Tests; Neuropsychological assessment

Introduction

The written Trail Making Test (WTMT; Army Individual Test Battery, 1944; Reitan, 1955) is a visuomotor speeded task consisting of two parts that is a widely used neuropsychological measure among clinicians (Rabin, Barr, & Butler, 2005). WTMT-A is a simple visual scanning task that requires one to draw a line connecting consecutive numbers from 1 to 25. WTMT-B adds the dimension of dividing one's attention between unrelated conceptual sets, requiring one to draw a line connecting numbers and letters in alternating sequence. Factor analytic studies have indicated that rapid visual search, contextual set shifting, and visuospatial sequencing are major components of WTMT-B performance (Des Rosier & Kavanaugh, 1987; Fossum, Holmberg, & Reinvang, 1992). However, non-cognitive factors such as poor vision and impaired motor functioning are known to impact performance and can preclude test administration in some instances. This problem is commonly encountered among older patients, persons who have sustained orthopedic injuries, and patients experiencing certain neurologic conditions (e.g., individuals whose dominant limbs are hemiparetic following a stroke).

An oral version of the TMT (OTMT) developed by Ricker and Axelrod (1994) removes the visual and motor components of the task in order to allow for the assessment of those who are unable to complete the standard written form. The sample in the original study included 58 healthy adults who were divided into three age-groups. Results suggested significant between-group

---

[†] Part of this paper was presented at the 21st Annual Meeting of the National Academy of Neuropsychology, San Francisco, CA.

age differences using analysis of variance (ANOVA). Subsequent post hoc analysis revealed significant differences between the youngest and oldest age groups on both OTMT-A and OTMT-B, although no statistically significant differences were detected between the younger and the middle age groups. Further, correlational analyses revealed significant relationships between these two versions of the TMT. Overall, the authors presented the OTMT as a promising addition to the assessment of attention by acting as an oral analog for WTMT.

Although years have passed, validity studies and normative data remain sparse for this measure despite evidence that it continues to be routinely employed by many neuropsychologists (Strauss, Sherman, & Spreen, 2006). In a recent book chapter, the original authors cite only a handful of validity studies for the OTMT, with most completed by their own group (Axelrod & Lamberty, 2006). Validity studies have been completed in a sample of 58 healthy adults (Ricker & Axelrod, 1994), 85 cerebrovascular accident patients (Ricker, Axelrod, & Houtler, 1996), 86 older medical patients (Ruchinskas, 2003), and a mixed clinical sample of 112 subjects with a wide variety of neurologic, medical, and psychiatric diagnoses (Abraham, Axelrod, & Ricker, 1996). Results of these validity studies suggested that the OTMT was able to discriminate between patients with different lesion sites at a similar level of sensitivity as the WTMT and established some of the basic psychometric properties of this measure (e.g., preliminary exploration of the statistical relationship between OTMT and WTMT, some demographic factors, and concurrent validity).

Arguments that the OTMT and the WTMT are analogous measures have been based on the congruence of results found in these preliminary studies. In the original study, correlations between the two versions of the TMT were large and statistically significant ($r = .68$ for TMT-A and $r = .72$ for TMT-B). Further evidence came when using the OTMT and WTMT in various clinical populations. The Abraham and colleagues (1996) study found an age-related effect for both the OTMT and the WTMT (with older subjects performing significantly worse than younger subjects) regardless of the patient group (normal, medical, or psychiatric diagnosis). Demographic variables including intelligence, education, and gender were not found to yield significant differences for either OTMT or WTMT among the groups. In addition, the authors found both OTMT-B and WTMT-B did not correlate well with measures of expressive language but did correlate well with measures of executive control. One study has reported an impact of education on OTMT performance in a sample of older medical patients (Ruchinskas, 2003). However, the authors in this study acknowledged a significant difference in the education levels among the three groups, with the control group having almost three more years of education than one of the experimental groups (neurologic patients). Results yielded the level of education as a significant predictor of passing or failing the OTMT-B.

In contrast, there are enough differences between these measures in terms of obvious perceptual and motor demands to question whether they measure the same underlying constructs. The differences in the demands of these measures are backed by a failure to observe a relationship between the oral and written versions of the TMT in some studies. For example, the Ricker and colleagues (1996) study found no significant correlation between the OTMT-B and the WTMT-B for one group of subjects (anterior strokes). This could reflect that motor deficits only hinder performance on the WTMT.

Similarly, although the initial OTMT publication found a large correlation ($r = .68$; Ricker & Axelrod, 1994) between Part A of the OTMT and the WTMT, subsequent studies have demonstrated only a weak to mild relationship ($r = .10 - .29$; Axelrod & Lamberty, 2006). Although WTMT-A requires visual scanning and psychomotor speed, the OTMT-A is really more of an overlearned, mental control task that places no demands on visual or motor functions. Creating a true analog of WTMT-A appears virtually impossible, as one cannot develop an oral sequencing task that parallels the written version.

There have also been some limitations in the existing validity studies primarily due to restrictions in sample characteristics, and normative data for the OTMT remain scant. The three groups used in the original Ricker & Axelrod (1994) study had a narrow demographic profile that placed considerable limitations on its normative properties. More specifically, Ricker and Axelrod's (1994) sample had a significant gap in the middle to older age ranges, with the age range from 40 to 70 years not well represented (i.e., the sample included college students and persons living in a retirement community). Because they were comparing groups at extreme ends of the age spectrum, it appears that this could bias their sample toward finding age-related correlations. The 112 subjects used in the Abraham and colleagues (1996) study suffered from a breadth of medical conditions with relative small sample sizes within each group to yield a very heterogeneous population, making the generalization of findings difficult. Although the convergent validity of OTMT appears promising in clinical populations, its psychometric properties within the normal population are tenuous given the limited normative information.

The purpose of the present study was to further examine the potential clinical use of the OTMT. The relative ease of administration, coupled with its potential applicability to populations with visual and/or motor limitations, makes it a viable instrument for use in a variety of clinical contexts. Yet to date, there have been no follow-up studies evaluating its psychometric properties in a nonclinical sample that is more evenly stratified by age. Therefore, we sought to explore relations between demographic variables (i.e., age, education, and gender) and performance on the OTMT. In addition, concurrent validity of the OTMT with WTMT was evaluated, and normative performance provided for the OTMT to make is useful for clinical comparison.

Materials and Methods

Participants

Participants included 81 adults (22 males, 59 females) who were participating in a comprehensive neuropsychological normative project. All participants were without history of known psychiatric or neurologic disorder and were living independently in their respective communities. No participant had a history of substance abuse or treatment, and none was currently receiving psychotropic medications. Participants were volunteers recruited from a variety of sources. They included family members of patients seen clinically at a couple of major medical centers, members of various civic organizations, and individuals recruited through the use of flyers in public locations. All participants provided written informed consent and the study was approved by an institutional review board. Participants were screened for exclusionary factors through the use of a careful, in-depth clinical interview. All participants performed within the normal range on a cognitive screening measure (i.e., Modified Cognistat, Drane, Yuspeh, Klingler, Huthwaite, & Mrazik, 2002, February, 2002; Drane et al., 2003) and on the 2nd edition of the Beck Depression Inventory (Beck, Steer, & Brown, 1996). Participants ranged in age from 20 to 90 years ($M = 48.47$, $SD = 19.56$), and education level ranged from 6 to 20 years ($M = 14.20$, $SD = 2.15$). The mean education level is somewhat high compared with the general population. Table 1 presents the demographic characteristics and OTMT performance for the entire sample.

Procedures

Test instructions. The OTMT (Ricker & Axelrod, 1994) and the WTMT (Army Individual Test Battery, 1944; Reitan & Wolfson, 1993) were administered according to standardized procedures as part of a comprehensive neuropsychological normative project. Test administration was counterbalanced in order to control for order effects.

Specifically, participants were given the following instructions prior to administration of the oral version of TMT. For OTMT-A:

"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."

For OTMT-B:

"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." (Ricker & Axelrod, 1994, pp. 48–49).

Although not mentioned in their original article (Ricker & Axelrod, 1994), Ricker and Axelrod have also reportedly redirected patients who lose their place on either of these tasks (Axelrod & Lamberty, 2006). We have found this to be a necessary procedure, as it is nearly impossible for patients to recover from a set loss error on OTMT-B without any kind of prompt. For the current study, we added the following instructions to be presented when errors were made on the OTMT:

If the patient makes an error on Part A, the examiner is to reorient them to the last correct number by saying: "You last said '[specific number],' please continue from there."

If the patient makes an error on Part B, the examiner is to reorient them to the last correct pair by saying: "You said '[specific number] [specific letter];' Continue from there."

Table 1. Demographic characteristics and means (*SD*) for OTMT performance for the entire sample ($n = 81$)

	Mean (*SD*)	Range
Age	49.62 (19.56)	20–90 years
Education	14.20 (2.15)	6–20 years
OTMT-A	6.88 (1.89)	2–12 s
OTMT-A errors	0.02 (0.15); mode = 0	0–1 errors
OTMT-B	33.91 (17.58)	12–90 s
OTMT-B errors	0.51 (0.36); mode = 0	0–3 errors
WTMT-A	29.85 (14.09)	13–85 s
WTMT-B	72.33 (35.57)	29–187 s
Gender composition	22 men, 59 women	

Notes: OTMT-A = Oral Trail Making Test, Part A; OTMT-B = Oral Trail Making Test, Part B; WTMT-A = Written Trail Making Test, Part A; WTMT-B = Written Trail Making Test, Part B. Not all participants received the WTMT ($n = 68$).

Timing is continued uninterrupted for both versions of the OTMT when instructions are given to keep them on task as described. For part B, timing is discontinued once subjects reach the number 13, identical to instructions for WTMT.

Analysis. Analyses included calculating Pearson product–moment correlations in order to examine the association of OTMT time to completion scores with demographic variables including age, gender, and education, as well as with the WTMT, Parts A and B (Army Individual Test Battery, 1944).

Results

OTMT-A time for the entire sample ranged from 2 to 12 s ($M = 6.88$, $SD = 1.89$), and OTMT-B time ranged from 12 to 90 s ($M = 33.91$, $SD = 17.58$). With respect to demographic characteristics, small and non-significant correlations were observed between OTMT-A time and age, education, and gender (rs ranged from −.01 to .17). Correlations were non-significant between OTMT-B time and education and gender. However, a significant correlation was observed between age and OTMT-B time ($r = .40$, $p = .0001$), with age accounting for 16% of the total variance in OTMT-B performance.

Correlational analyses revealed significant relationships between OTMT, Parts A and B, and their written counterparts (rs were .29 and .62, respectively), although the relationship was small between OTMT-A and WTMT-A. Interestingly, the correlation between the OTMT-A and the OTMT-B was also small, whereas the correlation between WMT-A and WMT-B appeared large. Table 2 presents the correlations among the demographic characteristics, oral TMT times, and written TMT times.

Given the impact of age on the OTMT, we decided to stratify our sample by age when presenting normative data. Therefore, participants were divided into six collapsed age groups with midpoint ages following the procedures outlined by Pauker (1988), as this method is recommended to maximize the clinical usefulness of normative data, especially when there is a wide range of normal subjects. Table 3 presents the mean performance of these six age groups on the OTMT and the WTMT. In addition,

Table 2. Correlation coefficients between demographic characteristics, OTMT performance, and WTMT performance

	Education	Gender	OTMT-A	OTMT-B	WTMT-A	WTMT-B
Age	.04	.05	.17	.40**	.61***	.69***
Education		−.16	−.01	−.13	−.14	−.19
Gender			.08	.08	.03	.03
OTMT-A				.22*	.29*	.31**
OTMT-B					.54***	.62***
WMT-A						.81***

Notes: OTMT-A = Oral Trail Making Test, Part A; OTMT-B = Oral Trail Making Test, Part B; WTMT-A = written Trail Making Test, Part A; WTMT-B = written Trail Making Test, Part B.
*$p < .05$.
**$p < .01$.
***$p < .001$.

Table 3. Means (*SD*) for collapsed groups on education level and time to completion performance on parts A and B of the WTMT and the OTMT (measured in seconds)

Age groups (years) (midpoint)	20–39 (29)	29–49 (39)	39–59 (49)	49–69 (59)	59–79 (69)	69–90 (79)
Age	28.8 (5.7)	38.3 (6.6)	48.2 (5.8)	59.2 (6.7)	69.3 (5.6)	76.3 (6.11)
Group size	31	31	31	27	23	18
Education	14.61 (2.33)	14.48 (2.36)	14.29 (2.28)	14.19 (2.22)	13.65 (1.77)	14.28 (2.40)
WTMT-A	21.53 (7.24)	22.44 (6.49)	24.05 (5.12)	31.48 (11.23)	38.73 (16.04)	43.63 (18.90)
WTMT-B	46.72 (17.04)	53.14 (18.62)	59.02 (17.62)	78.26 (36.96)	96.85 (36.39)	109.42 (38.25)
OTMT-A	6.25 (1.32)	6.56 (1.42)	7.02 (2.01)	7.16 (2.34)	6.98 (2.27)	7.48 (2.07)
OTMT-A err	0.00 (0.00)	0.01 (0.18)	0.01 (0.18)	0.00 (0.00)	0.01 (0.21)	0.01 (0.24)
OTMT-B	27.77 (14.80)	29.97 (15.19)	30.65 (14.18)	35.27 (16.24)	42.25 (19.16)	46.78 (20.55)
OTMT-B err	0.42 (0.62)	0.52 (0.85)	0.61 (0.92)	0.00 (0.75)	0.51 (0.73)	0.67 (0.69)

Notes: WTMT-A = Written Trail Making Test, Part A; WTMT-B = Written Trail Making Test, Part B; OTMT-A = Oral Trail Making Test, Part A; OTMT-A err = Oral Trail Making Test, Part A errors; OTMT-B = Oral Trail Making Test, Part B; OTMT-B = Oral Trail Making Test, Part B. Not all participants received the WTMT (20–39 years: $n = 24$, 29–49 years: $n = 29$, 39–59 years: $n = 25$, 49–69 years: $n = 21$, 59–79 years: $n = 22$, 69–90 years $n = 18$).

percentile distribution scores were established to outline the cumulative frequencies of test score performance for OTMT-A and OTMT-B for each of the collapsed age groups (Table 4).

Axelrod and Lamberty (2006) in their summary chapter of OTMT research suggested that future research should look at the comparison of WTMT-B to OTMT-B as a means to estimate the written version from the oral score. This would allow one to take advantage of the demographically corrected scores for the WTMT-B, which would be particularly useful for patients unable to complete the written form. Their work suggested that the WTMT-B to OTMT-B ratio was approximately 2.5. However, calculations of the WTMT-B to OTMT-B ratio with our data revealed an overall ratio of 2.1 for the entire sample, with a range 1.7–2.3 for the different collapsed age groups (Table 5). These scores are substantially lower than 2.5 ratio that was previously reported and suggest that this finding may not stable enough to allow for reliable predictions from one version of the test to the other. Although there was a clear tendency for older age groups to exhibit higher ratios in our sample, the overall model did not reach statistical significance ($r = .22$, $p = .06$). Table 5 also includes the mean estimated WTMT-B score for each of the collapsed age groups, as well as a corresponding demographically corrected T-score based on the Heaton, Miller, Taylor, and Grant (2004) demographically corrected norms.

Finally, cumulative frequencies were generated to address the issue of clinical significance in our population sample. Specifically, cumulative frequencies of the WTMT-B to OTMT-B ratio for each of the collapsed age groups and for the entire sample are presented in Table 6. This will allow one giving both versions of the test to get a better idea of how common disparities are between the two versions.

Discussion

Our results suggest that OTMT-B shows a large relationship with both Parts A and B of the WTMT, but indicate that the OTMT-A is only weakly related to either written task. In addition, OTMT-A shows little relationship with OTMT-B despite

Table 4. Group means, standard deviations, and selected percentile distribution of OTMT scores (in seconds) for midpoint age groups

Age groups (years) (midpoint)	20–39 (29)	29–49 (39)	39–59 (49)	49–69 (59)	59–79 (69)	69–90 (79)
Group size	31	31	31	27	23	18
OTMT-A	6.25 (1.3)	6.56 (1.4)	7.02 (2.0)	7.16 (2.3)	6.98 (2.3)	7.48 (2.1)
9th percentile	8.0	8.0	9.6	9.6	9.1	9.1
25th percentile	7.4	7.3	7.6	8.2	8.9	8.4
50th percentile	6.0	6.0	6.4	6.4	7.1	8.0
75th percentile	5.0	5.3	5.7	5.5	5.3	5.9
OTMT-B	27.77 (14.8)	29.97 (15.2)	30.65 (14.1)	35.27 (16.2)	42.25 (19.1)	46.78 (20.6)
2nd percentile	55.0	58.0	58.0	60.0	73.0	82.3
9th percentile	44.0	53.0	53.0	56.0	69.7	73.0
25th percentile	33.0	34.1	34.0	42.0	52.9	52.9
50th percentile	22.0	22.9	24.2	30.1	32.1	36.1
75th percentile	16.0	17.2	20.0	23.6	25.0	31.4

Notes: We suggest using the 9th percentile score distribution as a cutoff point for possible impairment. Of note, however, cutpoints may vary on the basis of the screening context.

Table 5. Summary of written-to-oral TMT-B ratio scores with estimated mean raw and demographically corrected T-scores by collapsed age groups

	Age groups (years) (midpoint)					
	20–39 (29)	29–49 (39)	39–59 (49)	49–69 (59)	59–79 (69)	69–90 (79)
WTMT-B/OTMT-B ratio[a]	1.7	1.8	1.9	2.2	2.3	2.3
Mean Est. WTMT-B scores[b]	47.2	53.9	58.2	77.6	97.2	107.6
Estimated WTMT-B T-scores[c]	54	54	53	48	48	49

Notes: WTMT-B = Written Trail Making Test, Part B; OTMT-B = Oral Trail Making Test, Part B. Not all Participants received the WTMT (20–39 years: $n = 24$, 29–49 years: $n = 29$, 39–59 years: $n = 25$, 49–69 years: $n = 21$, 59–79 years: $n = 22$, 69–90 years: $n = 18$).
[a]The correlation between age and ratio scores approached statistical significance ($r = .22$, $p = .06$).
[b]Mean estimated WTMT-B scores derived from calculating the OTMT-B mean from within a collapsed age group and multiplying the mean by the WTMT-B/OTMT-B ratio.
[c]Estimated T-scores compiled means from collapsed groups and these translated scores to T-scores based on demographically corrected norms from Heaton and colleagues (2004).

Table 6. Cumulative percent of written-to-oral TMT-B ratio scores

Ratios	Age groups (years)						Total
	20–39	29–49	39–59	49–69	59–79	69–90	
6.70	—	—	—	4.8	9.5	—	1.5
5.50	5.0	3.6	—	—	—	—	3.1
4.40	—	—	—	9.5	14.3	5.6	5.4
4.20	—	—	—	—	—	11.1	6.2
4.00	—	—	—	—	—	16.7	6.9
3.40	—	7.1	5.0	14.3	23.8	—	10.0
3.30	10.0	10.7	—	—	—	22.2	13.1
3.20	—	14.3	—	19.0	33.3	—	15.4
3.10	15.0	—	—	—	—	27.8	17.7
3.00	—	—	—	23.8	42.9	—	19.2
2.90	—	17.9	15.0	28.6	—	—	22.3
2.80	—	21.4	20.0	33.3	—	—	24.6
2.70	—	28.6	35.0	42.9	52.4	—	30.8
2.60	—	—	—	47.6	61.9	44.4	35.4
2.50	—	32.1	40.0	—	—	50	38.5
2.40	—	35.7	—	52.4	—	—	40.0
2.30	—	39.3	—	—	—	—	40.8
2.20	20.0	46.4	50.0	57.1	—	61.1	47.7
2.10	25.0	50.0	55.0	61.9	71.4	72.2	54.6
2.00	30.0	53.6	60.0	—	—	—	56.9
1.90	40.0	60.7	65.0	71.4	81.0	—	63.1
1.80	55.0	67.9	70.0	76.2	—	77.8	70.0
1.70	65.0	71.4	—	81.0	90.5	83.3	74.6
1.60	75.0	75.0	75.0	85.7	100	88.9	81.5
1.50	80.0	78.6	—	—	—	—	83.1
1.40	—	—	—	—	—	100	86.2
1.30	—	—	—	—	—	—	—
1.20	85.0	85.7	85.0	90.5	—	—	90.8
1.10	—	92.9	95.0	95.2	—	—	95.4
1.00	—	—	100	100.0	—	—	96.9
0.90	95.0	96.4	—	—	—	—	98.5
0.70	100.0	100.0	—	100.0	—	—	100.0

large correlations between the two written tasks both in this study and across prior studies involving the WTMT (Strauss et al., 2006). Taken together, these findings suggest that OTMT-A is not an equivalent task to WMT-A. It is probable that the OTMT-A is less strongly related to its written counterpart due to the relative simplicity of the task. It only requires patients to repeat an overlearned rote sequence, which produces a restricted range in scores. In contrast, WTMT-A requires other cognitive abilities such as planning, visual scanning and tracking, and visuomotor functioning. To some degree, these differences also exist between the OTMT-B and the WTMT versions. However, there is a strong relationship between the OTMT-B and the WTMT versions, which is perhaps due to a presumed dependence on mental flexibility and divided attention (although external validation of the shared underlying neurocognitive constructs is still lacking; Axelrod & Lamberty, 2006).

Results clearly identified the moderating effect of age on performance of both the OTMT-B and the WTMT-B, although this difference was more pronounced for the written version. There was a trend for completion times to get slower for OTMT-B and WTMT-B across each of our six age groups. It is likely that the additional visuomotor demands of WTMT-B versus OTMT-B contribute to the slightly greater impact of age on the written version. The absence of a significant impact of age on OTMT-A performance once again highlights the relative simplicity and rote nature of this task and again distinguishes it from its written counterpart.

Dissociations between performance on the OTMT-B and WTMT-B are highlighted by both the differential impact of age on these measures and by the modest correlations observed between these tasks. This is actually consistent with growing evidence that attentional functioning can vary across different sensory modalities in an individual subject (Benedict et al., 1998; Sinnett, Juncadella, Rafal, Azanon, & Soto-Faraco, 2007) and presumably reflects the activity of partially segregated neural networks. There is also ample evidence that dissociations across attentional tasks tapping varying sensory modalities may frequently be of diagnostic utility (Berger & Posner, 2000; Collier & Logan, 2000; Connor, Streissguth, Sampson, Bookstein, & Barr, 1999; Jonkman et al., 1997; Wood, Potts, Hall, Ulanday, & Netsiri, 2006). Therefore, rather than analogs, the OTMT and WTMT

may actually reflect somewhat complimentary tasks and administering both versions may be helpful in patients without physical handicaps. The work of Ricker and Axelrod (1994) indicated that the order in which these versions are administered does not result in confounding effects. Order effects were also absent when examined in a small subset of our current sample. This leads us to conclude that there are no confounding effects of using both versions in a single assessment.

A clinical comparison between the oral and written administrations may also allow an examiner to better clarify factors contributing to the patient's performance in some situations. For example, combining the comparative data obtained from the OTMT and WTMT with a thorough examination of motor and visuoperceptual functions can help to elucidate the possible role of perceptual, motor, and attentional deficits on tasks requiring more complex cognitive integration. Moreover, in patients lacking perceptual and/or motor deficits, or in whom the contribution of such deficits can be controlled, it may be possible to determine if deficits in attention are limited to one or more sensory modalities (e.g., auditory versus visual).

Our results support previous research demonstrating age-related differences in performance on both the WTMT and the OTMT (Drane, Yuspeh, Huthwaite, & Klingler, 2002; Ernst, 1987; Rasmusson, Zonderman, Kawas, & Resnick, 1998; Ricker & Axelrod, 1994). These results provide additional evidence that complex processing speed may slow down and cognitive flexibility may diminish with normal aging, thereby supporting the use of appropriate normative data that controls for age when using the OTMT.

Our data also revealed a trend for older subjects to have higher WTMT-B to OTMT-B ratios, although results were not statistically significant between groups. A higher ratio in older individuals could suggest that performance on the written version declines at a greater rate than the oral version. This could reflect greater impact of general physical slowing that may accompany the aging process. Previous studies have not compared this ratio across age groups, and this trend bears further examination in future normative studies using the OTMT. We should also note that the calculated WTMT-B to OTMT-B ratio in our sample was lower than that reported in previous research (Axelrod & Lamberty, 2006), suggesting that conversion across forms of the TMT could potentially be problematic.

Although the sample size of the current study is modest, this represents the largest sample of normative data presented to date for this measure and provides a broader age range of subjects compared with the original normative study completed by Ricker and Axelrod (1994). In addition, the age ranges within each of the six groups used in this study enhance the clinical utility of the OTMT by allowing a better comparison of subjects to an appropriate normative group, as the only existing normative data lack subjects in the middle age ranges. Our sample was also well-defined, screened for existing cognitive and mood-related issues, and did not reflect a population of convenience. The current study includes a disproportionate number of females, yet gender has not been shown to affect performance in any prior study of the OTMT and has been shown to have minimal impact on the WTMT (Strauss et al., 2006). Despite the wide-range of education levels reflected in our current sample, most individuals tended to have had at least a high-school education. Therefore, it is possible that we could have underestimated the impact of education on this measure. Only one prior study has reported an impact of education (Ruchinskas, 2003). However, as there were significant differences in education levels across the three groups analyzed in the Ruchinskas (2003) study, a strong possibility exists that results were confounded by differences in neurologic diagnosis. Overall, all studies to date addressing this issue have been limited and lacked adequate sampling of education levels.

Further research using both versions of the TMT in various clinical populations is warranted. Likewise, factor analyzing a battery of neurocognitive tests including the WTMT and OTMT would help to elucidate the underlying constructs measured by each version of the test. Such research approaches might also reveal more about the underlying neural substrates of these different facets of attention. Functional imaging studies exploring the use of these tasks may also prove useful in this regard as well.

Overall, although the OTMT and the WTMT do not appear to be completely analogous paradigms, it appears likely that they both require set-shifting, allowing for this construct to be assessed in patients with physical and visual limitations. In addition, differences in task demands may make them useful in a complimentary manner for establishing differences in attention across sensory modalities where such dissociations are suspected. Finally, this test is available in the public domain, making it readily available to be used by clinicians without any associated cost.

Funding

Conflict of Interest

None declared.

Acknowledgements

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). *Manuel of directions and scoring*. Washington, DC: War Department, Adjutant's Generals Office.

Axelrod, B. N., & Lamberty, G. J. (2006). The Oral Trail Making Test. In A. M. Poreh (Ed.), *Neuropsychological assessment: A quantified process approach* (pp. 45–52). Lisse, the Netherlands: Swets & Zeitlinger.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory—II*. San Antonio, TX: The Psychological Corporation.

Benedict, R. H. B., Lockwood, A. H., Shucard, J. L., Shucard, D. W., Wack, D., & Murphy, B. W. (1998). Functional neuroimaging of attention in the auditory modality. *Neuroreport*, *9*, 121–126.

Berger, A., & Posner, M. I. (2000). Pathologies of brain attentional networks. *Neuroscience and Biobehavioral Reviews*, *24*, 3–5.

Collier, G. L., & Logan, G. (2000). Modality differences in short-term memory for rhythms. *Memory and Cognition*, *28*, 529–538.

Connor, P. D., Streissguth, A. P., Sampson, P. D., Bookstein, F. L., & Barr, H. M. (1999). Individual differences in auditory and visual attention among fetal alcohol-affected adults. *Alcoholism, Clinical and Experimental Research*, *23*, 1395–1402.

Des Rosier, G., & Kavanagh, D. (1987). Cognitive assessment in closed-head injury: Stability, validity, and parallel forms for two neuropsychological measures of recovery. *International Journal of Clinical Neuropsychology*, *9*, 169–173.

Drane, D., Yuspeh, L., Klingler, L., Huthwaite, J., & Mrazik, M. (2002, February). *Introduction of a total composite score for the Cognistat (NCSE). Paper presented at the Journal of the International Neuropsychological Society*.

Drane, D. L., Yuspeh, R. L., Huthwaite, J. S., & Klingler, L. K. (2002). Demographic characteristics and normative observations for Derived-Trail Making Test indices. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *15*, 39–43.

Drane, D. L., Yuspeh, R. L., Huthwaite, J. S., Klinger, L. K., Foster, L. M., Mrazik, M., et al. (2003). Healthy older adult performance on a modified version of the Cognistat (NCSE): Demographic issues and preliminary normative data. *Journal of Clinical and Experimental Neuropsychology*, *25 (1)*, 133–144.

Ernst, J. (1987). Neuropsychological problem-solving skills in the elderly. *Psychology and Aging*, *4*, 363–365.

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

Heaton, R. S., Miller, S. W., Taylor, M. J., & Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead–Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults* (Rev. ed.). Lutz, FL: Psychological Assessment Resources Inc.

Jonkman, L. M., Kemner, C., Verbaten, M. N., Koelega, H. S., Camfferman, G., van der Gaag, R. J., et al. (1997). Effects of methylphenidate on event-related potentials and performance of attention-deficit hyperactivity disorder children in auditory and visual selective attention tasks. *Biological Psychiatry*, *41 (6)*, 690–702.

Pauker, J. E. (1988). Constructing overlapping cell tables to maximize the clinical usefulness of normative test data: Rationale and an example from neuropsychology. *Journal of Clinical Psychology*, *44*, 930–933.

Rabin, L. A., Barr, W. B., & Butler, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, APA Division 40 members. *Archives of Clinical Neuropsychology*, *20 (1)*, 33–65.

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. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology*, *19*, 393–394.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead–Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation* (2nd ed.). Tucson, AZ: Neuropsychology Press.

Ricker, J. H., & Axelrod, B. N. (1994). Analysis of an oral paradigm for the Trail Making Test. *Assessment*, *1 (1)*, 51–55.

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.

Ruchinskas, R. A. (2003). Limitations of the Oral Trail Making Test in a mixed sample of older individuals. *The Clinical Neuropsychologist*, *17 (2)*, 137–142.

Sinnett, S., Juncadella, M., Rafal, R., Azanon, E., & Soto-Faraco, S. (2007). A dissociation between visual and auditory hemi-inattention: Evidence from temporal order judgments. *Neuropsychologia*, *45 (3)*, 552–560.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford University Press.

Wood, S. M., Potts, G. F., Hall, J. F., Ulanday, J. B., & Netsiri, C. (2006). Event-related potentials to auditory and visual selective attention in schizophrenia. *International Journal of Psychophysiology*, *60 (1)*, 67–75.