Applied Neuropsychology
2002, Vol. 9, No. 2, 106–109

BRIEF REPORTS

The Trail Making Test, Part B: Cognitive Flexibility or Ability to Maintain Set?

Kathleen Bechtold Kortte

Medical University of South Carolina, Charleston, South Carolina, USA

Michael David Horner

Ralph H. Johnson Medical Center, Department of Veterans Affairs, Medical University of South Carolina, Charleston, South Carolina, USA

Whitney K. Windham

College of Charleston, Charleston, South Carolina, USA

The Trail Making Test (TMT) is a well-established test sensitive to impairment in multiple cognitive domains. There has been ambiguity about which cognitive demands are placed on the patient by TMT Part B over and above those required to perform TMT Part A. In particular, cognitive flexibility and ability to maintain a complex response set have been 2 competing hypotheses. This study preliminarily examined which of these 2 abilities primarily contributes to Part B performance. A total of 121 clinically referred Veterans Affairs patients were administered the TMT, as well as other tests of executive and other cognitive functions. Regression analyses were used to examine which tests predicted Part B performance above and beyond Part A performance. The results provide preliminary support for TMT Part B performance being more sensitive to cognitive flexibility (operationalized as Wisconsin Card Sorting Test [WCST], percent perseverative errors) than ability to maintain set (operationalized as WCST, failure to maintain set).

Key words: Trail Making Test, cognitive flexibility, set maintenance, executive function

The Trail Making Test (TMT) was originally developed as part of the Army Individual Test Battery (1944) and is a standard component of many neuropsychological batteries (Lezak, 1995; Reitan & Wolfson, 1993). It is one of the most commonly used tests because of its high sensitivity to the presence of cognitive impairment, although poor performance on the test is a relatively nonspecific finding (Lezak, 1995; Reitan, 1958; Spreen & Benton, 1965). The test has two parts. Part A requires the participant to draw lines on a page connecting 25 numbers consecutively as quickly as possible. In Part B, the participant must draw lines alternating between numbers and letters in consecutive order. Performance is assessed by the time taken to complete each trial correctly (Reitan & Wolfson, 1993). Alternate written forms of both parts of the test have also been developed and have been shown to be reliable alternatives (Franzen, Paul, & Iverson, 1996; LoSasso, Rapport, Axelrod, & Reeder,

Kathleen Bechtold Kortte is now at Johns Hopkins University, Baltimore, MD, USA. Whitney K. Windham is now at University of Central Florida, Daytona Beach, FL, USA

Requests for reprints should be sent to Michael D. Horner, Mental Health Service (116), Ralph H. Johnson VA Medical Center, 109 Bee Street, Charleston, SC 29401-5799, USA. E-mail: hornermd@musc.edu

1998; McCracken & Franzen, 1992). In addition, an oral version, which has no visual or graphomotor component, has been developed (Ricker, Axelrod, & Houtler, 1996).

Several authors have suggested that the TMT is best conceptualized as reflecting a combination of several cognitive functions (Newby, Hallenback, & Embretson, 1983). It has been frequently described as a test of complex visual scanning with a motor component (Lezak, 1995; Shum, McFarland, & Bain, 1990). Motor speed and agility have been found to make a strong contribution to success on this task (Schear & Sato, 1989). Lamberty, Putnam, Chatel, Bieliauskas, and Adams (1994) suggested that the two parts of TMT require simple motor-spatial skills and basic sequencing abilities. Factor analytic studies of the TMT have indicated that performance on both parts of the test load on a visual perceptual factor (Groff & Hubble, 1981) and an attention factor (O'Donnell, MacGregor, Dabrowski, Oestreicher, & Romero, 1994).

Studies designed to explore what makes Part B harder than Part A indicate that there are three main factors contributing to the difference. Specifically, the relative difficulty of Parts A and B may be caused by differences in symbolic complexity and spatial arrangement and to the interaction between these two factors (Fossum, Holmberg, & Reinvang, 1992; Gaudino, Geisler, & Squires, 1995). The third factor is the difference in cognitive demands between the two tasks (Gaudino et al., 1995; Spreen & Strauss, 1998). It has been proposed that Part B performance is indicative of executive function, and that difficulty with the task might reflect impaired executive control or the ability to flexibly shift the course of an ongoing activity (Arbuthnott & Frank, 2000; Lamberty et al., 1994; Pontius & Yudowitz, 1980). Others have suggested that Part B requires an ability to maintain two response sets simultaneously (Eson, Yen, & Bourke, 1978; Reitan, 1971). As a preliminary way of investigating these two competing hypotheses of cognitive flexibility and ability to maintain a complex response set, this study compared performance on the original, written version of the TMT Part B with other measures of cognitive functioning.

Methods

Participants were 121 Veterans Affairs (VA) Medical Center outpatients who were drawn from a sample of patients referred between 1996 and 2000 for neuropsychological evaluation of possible cognitive impairment. All patients who received an evaluation that included the TMT, Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993), Controlled Oral Word Association Test (COWAT; Spreen & Strauss, 1998), Wechsler Adult Intelligence Scale–Revised (WAIS–R) Digit Span subtest (Wechsler, 1981), and California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987) were included in the study. Participants ranged in age from 25 to 80 years of age ($M = 50.23$, $SD = 13.45$). Of the group, 93.4% were men, 67.8% were White, and 30.6% were African American. Educational level ranged from 5 to 18 years ($M = 12.25$, $SD = 2.62$). Current vocational status included the following: 46.2% were unemployed, 15.1% retired, 24.4% employed full-time, 6.7% employed part-time, 4.2% students, and 3.4% on disability. Of the 121 participants, 48 were found to meet criteria for a diagnosis of cognitive disorder, including 34 with cognitive disorder not otherwise specified, 5 with dementia of the Alzheimer's type, 2 with vascular dementia, and 7 with dementia not otherwise specified. The remaining patients generally either were diagnosed with various psychiatric disorders (e.g., major depressive disorder, posttraumatic stress disorder) or did not meet criteria for any Axis I diagnosis.

For the purposes of this study, ability to maintain set was operationalized by the failure to maintain set score on the WCST. Cognitive flexibility was operationalized as the percentage of perseverative errors on the WCST. (Although definitive data regarding the construct validity of these WCST scores as indexes of these functions are lacking, they are generally presumed to be measures of these constructs and were chosen as the best available of such measures.) To test whether other cognitive functions also contribute to performance on TMT Part B in addition to the abilities required to perform Part A, measures of verbal fluency, attention, and memory were also included. These were operationalized as the total score on the COWAT (verbal fluency), total raw score on the WAIS–R Digit Span subtest (attention), total score on Trial 1 through 5 of the CVLT (verbal learning), and CVLT long delay free recall (verbal memory). These tests were not assumed to cover all possible areas of cognitive function but rather were included to examine which of those functions were most closely associated with performance on Part B. Each test was administered and scored in the standardized manner described in each test's administration manual.

Results

To establish which cognitive abilities relate to performance on the TMT, Pearson correlations were conducted. As shown in Table 1, these analyses revealed that scores on TMT Part B were significantly correlated with scores on TMT Part A. Interestingly, performance on Part B was also significantly correlated with scores on all other measures of cognitive functioning except WCST failure to maintain set. To determine the relative contribution of the ability to maintain complex response set, cognitive flexibility, or other executive or other cognitive functions to the prediction of Part B performance after accounting for the variance attributable to Part A, a hierarchical regression analysis was conducted. Given that age-corrected scores were not used, the ages of the participants were entered first into the equation accounting for 22% of the variance in Part B performance. Part A was then entered and accounted for 33% of the variance in Part B performance. Above and beyond Part A, only the measure of flexibility (WCST percent perseverative errors) contributed significantly to the prediction of Part B performance (R^2 change = .07, $p < .001$).

Discussion

The results provide preliminary support for the hypothesis that TMT Part B is more sensitive to deficits in cognitive flexibility than the ability to maintain a complex response set. Specifically, the results indicate that flexibility, operationalized as WCST percent perseverative errors, was the only significant predictor of Part B performance over and above Part A performance. Although only a modest effect was found for flexibility and most of the variance was accounted for by Part A, the competing hypothesis of ability to maintain set was not found to be a significant predictor of Part B performance. In fact, the ability to maintain set (WCST, failure to maintain set) was the only measure that was not related to Part B performance. Verbal fluency, attention, verbal learning, and verbal memory were related to Part B performance but were not found to be significant predictors after accounting for the variance attributable to Part A.

The lack of a relationship between ability to maintain set and Part B may be attributable to how this construct was operationalized in this study. The WCST failure to maintain set score is the number of times the individual made an error after a sequence of five to nine correct responses. In contrast, TMT Part B has been described as requiring maintenance of two response sets simultaneously (Lezak, 1995; Reitan, 1971), which could actually require the ability to switch repeatedly between two sequences.

Further, neither flexibility nor maintenance of set is the hallmark cognitive process needed to perform Part B; rather, it appears to be some other unique construct that is not well captured by other measures typically used. TMT Part B might tap cognitive processes that are unique from most other clinical measures. Thus, these findings do not indicate that flexibility is the primary aspect of the underlying complexity of Part B but rather provide preliminary support for flexibility being one part of what makes Part B harder than Part A.

These findings are limited because the results have not been cross-validated in another sample. Also, because the study sample was drawn from a VA popu-

Table 1. *Pearson Correlations of Scores on Trail Making Test Part B With Scores on Other Cognitive Measures*

	TMT Part B	TMT Part A	WCST %PE	WCST FTMS	COWAT	CVLT Total	CVLT LDFR
TMT Part A	.72**	——					
WCST %PE	.59**	.51**	——				
WCST FTMS	.02	.03	–.09	——			
COWAT	–.35**	–.30**	–.25**	–.07	——		
CVLT Total	–.53**	–.52**	–.43**	–.09	.39**	——	
CVLT LDFR	–.43**	–.44**	–.38**	–.10	.34**	.85**	——
WAIS–R Digit Span	–.27**	–.22*	–.17	–.23*	.43**	.33**	.30**

Note: TMT = Trail Making Test; WCST %PE = Wisconsin Card Sorting Test percent perseverative errors; WCST FTMS = Wisconsin Card Sorting Test failure to maintain set; COWAT = Controlled Oral Word Association Test; CVLT Total = California Verbal Learning Test total score on Trials 1 through 5; CVLT LDFR = California Verbal Learning Test long delay free recall; WAIS–R = Wechsler Adult Intelligence Scale–Revised.
*$p < .05$. **$p < .01$.

lation, the findings might not be generalizable to other populations. Nevertheless, these findings indicate that the TMT Part B might be sensitive to deficits in cognitive flexibility but not to deficits in ability to maintain set. It thus appears that clinical interpretation of TMT Part B performance could include discussion of the individual's cognitive flexibility with consideration given to performance on other tests of flexibility.

References

Arbuthnott, K., & Frank, J. (2000). Trail Making Test, Part B as a measure of executive control: Validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology, 22,* 518–528.

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

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test: Adult Version.* San Antonio, TX: Psychological Corporation.

Eson, M. E., Yen, J. K., & Bourke, R. S. (1978). Assessment of recovery from serious head injury. *Journal of Neurology, Neurosurgery, and Psychiatry, 41,* 1036–1042.

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

Franzen, M. D., Paul, D., & Iverson, G. L. (1996). Reliability of alternate forms of the Trail Making Test. *Clinical Neuropsychologist, 10,* 125–129.

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 Neuropsychology, 17,* 529–535.

Groff, M. G., & Hubble, L. M. (1981). A factor analytic investigation of the Trail Making Test. *Clinical Neuropsychology, 3,* 11–13.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting (WCST) manual revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Lamberty, G. J., Putnam, S. H., Chatel, D. M., Bieliauskas, L. A., & Adams, K. M. (1994). Derived Trail Making Test indices: A preliminary report. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 7,* 230–234.

Lezak, M. D. (1995). *Neuropsychological Assessment.* 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.

McCracken, L. M., & Franzen, M. D. (1992). Principal-components analysis of the equivalence of alternate forms of the Trail Making Test. *Psychological Assessment, 4,* 235–238.

Newby, R. F., Hallenback, C. E., & Embretson, S. (1983). Confirmatory factor analysis of four general neuropsychological models with a modified Halstead–Reitan Battery. *Journal of Clinical Neuropsychology, 5,* 115–133.

O'Donnell, J. P., MacGregor, L. A., Dabrowski, J. J., Oestreicher, J. M., & Romero, J. J. (1994). Construct validity of neuropsychological tests of conceptual and attentional abilities. *Journal of Clinical Psychology, 50,* 596–600.

Pontius, A. A., & Yudowitz, B. S. (1980). Frontal lobe system dysfunction in some criminal actions as shown in the narratives test. *Journal of Nervous and Mental Disease, 168,* 111–117.

Reitan, R. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8,* 271–276.

Reitan, R. (1971). Trail Making Test results for normal and brain-damaged children. *Perceptual and Motor Skills, 33,* 575–581.

Reitan, R., & Wolfson, D. (1993). *The Halstead–Reitan Neuropsychological Test Battery: Theory and clinical interpretation.* Tucson, AZ: Neuropsychology Press.

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.

Shum, D. H. K., McFarland, K. A., & Bain, J. D. (1990). Construct validity of eight tests of attention: Comparison of normal and closed head injured samples. *The Clinical Neuropsychologist, 4,* 151–162.

Spreen, O., & Benton, A. (1965). Comparative studies of some psychological tests for cerebral damage. *Journal of Nervous and Mental Disease, 140,* 323–333.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests* (2nd ed). New York: Oxford University Press.

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

Original submission August 25, 2001
Accepted December 28, 2001