Trail Making Test in assessing children with reading disabilities: A test of executive functions or content information

Neuropsykologian
erikoistumiskoulutuksen
lisensiaatintutkimus

Vesa Närhi
Psykologian laitos
Jyväskylän yliopisto
17. 2. 1999

Perceptual and Motor Skills, 1997, 84, 1355-1362.

TRAIL MAKING TEST IN ASSESSING CHILDREN WITH READING DISABILITIES: A TEST OF EXECUTIVE FUNCTIONS OR CONTENT INFORMATION[1]

VESA NÄRHI, PEKKA RÄSÄNEN, RIITTA-LEENA METSÄPELTO, AND TIMO AHONEN

Niilo Mäki Institute
University of Jyväskylä, Finland

Summary.—The speed of performance on Part A, Part B, and on an experimental version containing alphabetical series (Part A Alphabetic) of the Trail Making Test was studied with 19 children with reading disabilities and 34 controls from Grades 4 to 6. When the test was used in discriminant profile fashion, children with reading disabilities showed a deficit compared with control children on Part B relative to Part A but did not relative to the new Part A Alphabetic. The results indicate that the performance of the children with reading disabilities on Part B is likely to be affected by their slowness on the alphabetical series. Based on these results we recommend that the speed of following the alphabetical series be assessed when using Part B of the Trail Making Test.

The Trail Making Test was originally used in neuropsychological assessment of adults, and a shortened version of the test has become a part of neuropsychological batteries for children (Reitan & Wolfson, 1992; Rourke, Bakker, Fisk, & Strang, 1983). The test has two parts, Part A and Part B. In Part A the subject has to connect numbered circles in numerical sequence, in Part B the task alternates between numbers and letters. The two parts of the test require different skills, Part A being simpler and Part B requiring more information processing (Spreen & Strauss, 1991). Davison (1974) in organizing neuropsychological tests for children into definable categories places Part A under a rubric of visuospatial and visuosequential abilities, and Part B into a cluster of tests requiring concept formation, reasoning, organizational ability, and flexibility in applying principles. In factor analytic studies (O'Donnell, MacGregor, Dabrowski, Oestreicher, & Romero, 1994; Shute & Huertas, 1990) Part B has variance in common with the Wisconsin Card Sorting Test and Category Test, both requiring concept-formation.

The abilities needed to complete Part B (organized visual search, maintaining and following through a plan, set shifting, and inhibitory capacity) associate the test with the skills defined as executive functions. Such func-

[1]The preparation of this article was supported by the Finnish Academy and Niilo Mäki Foundation. We gratefully acknowledge the assistance of Virva Pitkänen, Katja Kokko, and the personnel of Niilo Mäki Institute in the data collection. Requests for reprints should be sent to Vesa Närhi, Niilo Mäki Institute, Department of Psychology, University of Jyväskylä, PL 35, 40351 Jyväskylä, Finland or by e-mail (vnarhi@nmi.jyu.fi).

tions are seen as responsible for goal-oriented behavior based on planning, employment of strategy, control of impulse, and organized search (Welsh, Pennington, & Groisser, 1991). Anatomically, executive functions have been related to functioning of the frontal lobes (Luria, 1980). By their very nature executive functions are domain-general (language, visuospatial, etc.) top-down functions. Any method of assessment, however, also includes domain-specific components as the material upon which the executive functions work, and consequently, poor performance on a test of executive function can also result from inadequate domain-specific components.

Attention Deficit Hyperactivity Disorder (ADHD) has been connected with frontal lobe dysfunction reflected in inferior executive function (e.g., Conners & Wells, 1986; Denckla, 1994). This connection has also gained empirical support (e.g., Barkley, Grodzinsky, & DuPaul, 1992; Pennington, Groisser, & Welsh, 1993), and it is within this framework that the Trail Making Test has been used in research on learning disabilities (Boucugnani & Jones, 1989; Gorenstein, Mammato, & Sandy, 1989; McGee, Williams, Moffitt, & Anderson, 1989). However, children with learning disabilities form a heterogeneous group for whom comorbidity of different learning and attention problems is common (e.g., Shaywitz & Shaywitz, 1991). Academic difficulties are associated with deficiencies in various cognitive domains, and these deficiencies may affect scores on tests of executive function and complicate their interpretation. For this reason the assessment of executive functions in these populations requires a careful assessment of the component skills of the tests of executive function.

Concerning the complex nature of Part B Rourke and Finlayson (1975) have shown that the deficits associated with poor performance on Part B differ from those on Part A for children with learning disabilities. Those children who show deficits on both Part B and Part A exhibited higher Verbal IQ on the WISC relative to Performance IQ, and children whose performance on Part A was within a normal range but who scored more poorly on Part B exhibited higher Performance than Verbal IQs. Also, the other measures used indicated that the deficits of the former group were confined to visuospatial and visuomotor skills, while the latter group showed deficits in verbal and language-related abilities. Rourke and Finlayson concluded that for the children whose performance on both Part A and Part B was deficient, the limiting factors may be visuospatial and visuomotor skills, and for the children who had poor performance on Part B but normal performance on Part A the crucial factor may be deficient verbal skills. Not explicitly stated, but inferred was that, if the more basic skills required on Part B (those used on Part A) had not been controlled, the interpretation of poor performance on Part B would have been obscured by deficient domain-specific functions.

The main component in Part B which is not controlled in Part A is the speed of following the alphabetical series. The possibility that this may play a significant role arose in a study by Närhi and Ahonen (1995) who used Part B as a measure of executive functions to study cognitive deficits associated with reading disabilities and ADHD. There were no differences between the studied groups (reading disabled, ADHD, comorbid, reading disabled + ADHD vs Clinical controls) on measures of executive function. This contrasted with the hypothesis that executive function deficits are specific to children with pure ADHD. The performance of all groups on measures of executive function was lower than the normative mean for their ages. It seemed probable that the inferior performance on Part B was due to various factors in different clinical groups. For the children with reading disability, it was hypothesized that their difficulties on the task might reflect problems in following the alphabetical series.

It seems that no single test can be considered to measure specifically the functions of the frontal lobes (for review see Reitan & Wolfson, 1994). Concerning Part B of the Trail Making Test evidence is convincing that the test is sensitive to disordered brain functions resulting from cerebral lesions, but that the test does not measure deficits specific to lesions of the frontal lobes (Reitan & Wolfson, 1995). Denckla (1994, 1996) has addressed the rationale for the assessment of executive functions. She proposed that executive functions can be assessed by using two tests within one cognitive domain, a simpler one as a baseline measure, and a more complex one taxing subjects' abilities to organize and employ strategies. If the performance on the more complex task is inferior to normative data, relative to that on the simple task, one can suppose that it is the result of deficient executive functions. The Trail Making Test seems to fit into this framework in that of the components needed to complete successfully the more complex Part B, visual and motor skills as well as the speed of following the numerical sequence all are assessed with Part A. The main unassessed component needed in Part B is the speed of following the alphabetical sequence.

Keeping in mind Denckla's (1994, 1996) suggestions about using a discriminant profile in the assessment of executive functions, we looked to see if children with reading disability are slower in following the alphabetical series relative to the numerical series and also relative to control children. If so, what is the effect of this upon the interpretation of performance by these children on the Trail Making Test Part B.

Method

The test versions were the standard Part A and Part B of the Trail Making Test (Reitan & Wolfson, 1992), and an experimental alphabetical version of Part A (Part A–Alphabetic) which includes letters from A to O, and in a training part, letters from A to H. The Part A–Alphabetic was administered in the same manner as Part A and Part B. In the training section the task was introduced to the subject, and possible mistakes were pointed out and explained. After successfully completing the training task, the test version was presented. Each mistake was immediately pointed out and the subject was instructed to continue from the point where the mistake occurred. The scores used were the times taken to complete the version. The administration and scoring procedure are described in detail by Reitan and Wolfson (1992). To ensure that the possible practice effects are similar for all children the order of presentation was held constant (Part A, Part A–Alphabetic, and Part B).

The reading disabled sample consisted of 19 children from Grades 4 to 6 referred to the child neuropsychological clinic of the Niilo Mäki Institute for the assessment of learning difficulties. From the consecutive referrals to the clinic we selected the sample with the following inclusion criteria: the child was attending regular school, either Verbal or Performance IQ of the WISC–R (Wechsler, 1974) was above 80, and the score on one of the two text reading tests was at least 1.5 standard deviations below the expected grade or age mean of the test presented to the child. The reason for using two different tests is that the text-reading test used in the clinic was changed during the data collection. Of the subjects 13 were diagnosed on the basis of a text-reading test normed for each age (Niilo Mäki Institute, 1992) and 6 were diagnosed on the basis of a text-reading test normed for each grade (Niilo Mäki Institute, 1994). For both tests a score of correctly read words per time unit was the score used in the diagnosis. The mean z score on the text-reading task for the sample was -3.3 ($SD = 1.1$). The Trail Making Test was administered as part of a neurocognitive evaluation of each child.

For the control group we selected one class from Grades 4 and 6, and those children ($n = 34$) whose parents provided a form for written consent participated. According to the teachers none of the children had learning disabilities. Also, according to the teachers, the classes selected represented an average school performance and no systematic selection in that term took place when some of the children were excluded due to missing permission from the parents. The testers were three psychology students who were carefully trained in administering the test.

Results

The means and standard deviations for times to complete the Trail Making Test versions are presented in Table 1. The WISC–R IQs of the children with reading disability were well within the normal range (Full scale IQ $M = 94.5$, $SD = 9.7$; Verbal IQ $M = 92.8$, $SD = 9.4$; Performance IQ $M = 97.7$, $SD = 11.9$). The sex distribution was much more even in the control

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR GROUPS OF TIMES TAKEN TO COMPLETE PARTS A, A-ALPHABETIC AND B OF THE TRAIL MAKING TEST

Variable	Children With Reading Disabilities $n=19$		Control Children $n=34$	
	M	*SD*	*M*	*SD*
Age, mo.	135.0	9.6	140.9	12.3
Part A	21.1	7.7	21.7	6.5
Part A-Alphabetic	35.1	17.2	27.4	12.2
Part B	55.8	23.4	41.9	15.0

group (19 boys, 15 girls) than in the reading disabled sample (18 boys, 1 girl). However, the sex differences in the control sample were statistically nonsignificant for all versions of the Trail Making Test by independent-sample, two-tailed *t* test, and this finding is consistent with other data (Reitan & Wolfson, 1992). Sex was not considered in the further analysis.

The children with reading disability were on the average six months younger than the control children. Although this difference was statistically nonsignificant on a two-tailed *t* test, to avoid the possible effects that age could have on the results we used age as a covariate in all subsequent analyses and for all main and interaction effects.

To study the speed of following the numerical and alphabetical series in the groups children's scores on Part A and Part A-Alphabetic were subjected to a 2 (group) by 2 (version) mixed-model analysis of variance design with version as a within-subject factor. Both main effects were nonsignificant, while the group by version interaction was significant ($F_{1,50}=4.34$, $p<.05$). The interaction indicated that, while performing at the same level as control children on Part A, children with reading disability were inferior on Part A-Alphabetic (see Table 1).

A similar mixed-model analysis-of-variance design is a valid way of comparing discriminant profiles of different groups on the baseline and complex versions of the task. In the analysis the interaction of group by version is a measure of whether the complex task is more difficult than the baseline measure in one group relative to the other. To study the effect of Part A and Part A–Alphabetic on the interpretation of scores on Part B separately, two 2(group) by 2(version) mixed-model analysis-of-variance designs with version as a within-subject factor were carried out, first between scores on Part B and Part A and then between scores on Part B and Part A–Alphabetic. In the analysis between Part B and Part A the main effect for group was nonsignificant, the main effect for version was significant ($F_{1,50}=16.12$, $p<.001$), and of most interest, there was a significant interaction for group by version ($F_{1,50}=6.87$, $p<.05$). Between Part B and Part A–Alphabetic the analysis showed significant main effects for group ($F_{1,50} = 5.10$, $p < .05$), and for version ($F_{1,50} = 6.51$, $p < .05$), but the interaction for group by version was not statistically significant ($F_{1,50} = 0.42$, $p = .52$). The differing results of the interaction for group by version in the two analyses show that whether children with reading disability are deficient relative to controls on Part B is dependent on the baseline measure used.

Discussion

Children with reading disability, while performing Part A at the same speed as control children, were inferior in following the alphabetical sequence on Part A–Alphabetic. When used as a discriminant profile the comparison between Part B and Part A showed that children with reading disability show deficits relative to control children but did not when the performances on Part B and Part A–Alphabetic were compared. If only Part A had been used in conjunction with Part B, the conclusion about the group comparison on scores of different versions would have been that children with reading disability, relative to control children, have problems in set-shifting and organization. When Part A–Alphabetic was also used, it may be concluded that there is no such difference. Considered together, the differing results using the two baseline measures lead to the conclusion that the poorer performance of children with reading disability show on Part B is due to the difficulties they have in following the alphabetical series.

There are some possible artifacts that could have affected the results. The structural nonequivalence of Part A and Part B has been noted by Rossini and Karl (1994) who noted that the circles are further apart in Part B than in Part A. There is also a structural nonequivalence between Part A and Part A–Alphabetic, the circles being 17% further apart in Part A–Alphabetic than in Part A. We re-ran the analysis using a length-corrected score (time taken to complete the version divided by the length of the version). In these analyses the results remained the same. When comparing the interfering stimuli in the same way as presented by Gaudino, Geisler, and Squires (1995) (number of incorrect targets within 3 cm of the correct path) both Part A and Part A–Alphabetic had 13 interfering stimuli, suggesting that this is not the factor which makes Part A–Alphabetic more difficult. To avoid this possible artifact completely one should study the effect of versions with visually identical forms containing different series. However, it seems that these interfering factors are not responsible for the results.

Lezak (1983) proposed that the time scores obtained on the Trail Making Test may not be reliable since they are partly based on the times taken to notice an error and having it pointed out by the examiner. To study the possible effect of different testers the scores of the control group on Part A, Part A–Alphabetic, and Part B were subjected to separate analyses of variance. The effect for testers was not significant for any of the versions, suggesting that the scores did not include tester-dependent variance.

The absence of scores on the text reading test for the control sample is a clear methodological weakness of the study. Although according to the teachers there were no children with marked learning disabilities in the control group, it is possible that some of the children had mild reading difficulties. Should this be the case it seems that the present results are not seriously weakened, since the inclusion of children with reading disability in the control sample would not increase but decrease the group-dependent differences between Part A and Part A–Alphabetic.

The difficulties that the children with reading disability have on Part A–Alphabetic relative to Part A may be explainable by the difference in the numerical and alphabetical series. The numerical series is a logical sequence, the next digit following the previous one according to number, while the alphabetical series can be seen as an agreement based phonological sequelae without any inherent logic. It can be speculated that the poor automatisation of alphabetical series by the children with reading disability is a result of their poor phonological skills (see, e.g., Pennington, *et al.*, 1993).

In the light of Denckla's (1994, 1996) proposition about using discriminant profiles when assessing executive functions, the Trail Making Test has an obvious face value. Her proposition is logical; it attempts to control the various possibilities affecting performance on necessarily complex executive function tasks, and also, it follows closely the clinical decision-making procedure regarding the executive functions of a subject. Our present data do not provide a possibility to evaluate the validity of the Trail Making Test as a measure of executive functions; those studies are yet to be constructed. Based on the present results, it can be concluded that controlling for the speed of performance on the alphabetical sequence is essential when using the Trail Making Test with populations for whom reading difficulties are common.

REFERENCES

Barkley, R. A., Grodzinsky, G., & DuPaul, G. J. (1992) Frontal lobe functions in attention deficit disorder with and without hyperactivity: a review and research report. *Journal of Abnormal Child Psychology*, 20, 163-188.

Boucugnani, L. L., & Jones, R. W. (1989) Behaviors analogous to frontal lobe dysfunctions in children with Attention Deficit Hyperactivity Disorder. *Archives of Clinical Psychology*, 4, 161-173.

Conners, C. K., & Wells, K. C. (1986) *Hyperkinetic children: a neuropsychosocial approach*. Beverly Hills, CA: Sage.

Davison, L. A. (1974) Current status of clinical neuropsychology. In R. M. Reitan & L. A. Davison (Eds.), *Clinical neuropsychology: current status and applications*. Washington, DC: V. H. Winston. Pp. 325-361.

Denckla, M. B. (1994) Measurement of executive function. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities*. Baltimore, MD: Paul H. Brookes. Pp. 117-142.

DENCKLA, M. B. (1996) A theory and model of executive function: a neuropsychological perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory and executive function*. Baltimore, MD: Paul H. Brookes. Pp. 263-278.

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.

GORENSTEIN, E. E., MAMMATO, C. A., & SANDY, J. M. (1980) Performance of inattentive-overactive children on selected measures of prefrontal-type function. *Journal of Clinical Psychology*, 45, 619-632.

LEZAK, M. D. (1983) *Neuropsychological assessment.* (2nd ed.) New York: Oxford Univer. Press.

LURIA, A. R. (1980) *Higher cortical functions in man.* (2nd ed.) New York: Basic Books.

MCGEE, R., WILLIAMS, S., MOFFITT, T., & ANDERSON, J. (1989) A comparison of 13-year-old boys with attention deficit and/or reading disorder on neuropsychological measures. *Journal of Abnormal Child Psychology*, 17, 37-53.

NÄRHI, V., & AHONEN, T. (1995) Reading disability with or without Attention Deficit Hyperactivity Disorder: do attentional problems make a difference? *Developmental Neuropsychology*, 11, 337-349.

NIILO MÄKI INSTITUTE. (1992) *Neuropsychological and achievement tests: local normative data for Niilo Mäki Institute–Test Battery.* Jyväskylä, Finland: Author.

NIILO MÄKI INSTITUTE. (1994) *Neuropsychological and achievement tests: local normative data for Niilo Mäki Institute–Test Battery–Revised.* Jyväskylä, Finland: Author.

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.

PENNINGTON, B. F., GROISSER, D., & WELSH, M. C. (1993) Contrasting cognitive deficits in Attention-Deficit Hyperactivity Disorder versus reading disability. *Developmental Psychology*, 29, 511-523.

REITAN, R. M., & WOLFSON, D. (1992) *Neuropsychological evaluation of older children.* Tucson, AZ: Neuropsychology Press.

REITAN, R. M., & WOLFSON, D. (1994) A selective and critical review of neuropsychological deficits and the frontal lobes. *Neuropsychology Review*, 4, 161-198.

REITAN, R. M., & WOLFSON, D. (1995) Category Test and Trail Making Test as measures of frontal lobe functions. *The Clinical Neuropsychologist*, 9, 50-56.

ROSSINI, E. D., & KARL, M. A. (1994) The Trail Making Test A and B: a technical note on structural nonequivalence. *Perceptual and Motor Skills*, 78, 625-626.

ROURKE, B. P., BAKKER, D. J., FISK, J. L., & STRANG, J. D. (1983) *Child neuropsychology: an introduction to theory, research, and clinical practice.* New York: Guilford.

ROURKE, B. P., & FINLAYSON, M. A. J. (1975) Neuropsychological significance of variations in patterns of performance on the Trail Making Test for older children with learning disabilities. *Journal of Abnormal Psychology*, 84, 412-421.

SHAYWITZ, B. A., & SHAYWITZ, S. E. (1991) Comorbidity: a critical issue in attention deficit disorder. *Journal of Child Neurology*, 6(Suppl.), 13-22.

SHUTE, G. E., & HUERTAS, V. (1990) Developmental variability of frontal lobe function. *Developmental Neuropsychology*, 6, 1-11.

SPREEN, O., & STRAUSS, E. (1991) *Compendium of neuropsychological tests: administration, norms, and commentary.* New York: Oxford Univer. Press.

WECHSLER, D. (1974) *Manual for the Wechsler Intelligence Scale for Children (Revised).* New York: Psychological Corp.

WELSH, M., PENNINGTON, B. C., & GROISSER, D. B. (1991) A normative-developmental study of executive function: a window on prefrontal function in children? *Developmental Neuropsychology*, 7, 131-149.

Accepted April 15, 1997.