University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Developmental Cognitive Neuroscience Laboratory - Faculty and Staff Publications | Developmental Cognitive Neuroscience Laboratory

11-1-2004

The Development of a Trial Making Test in Young Children: The TRAILS-P

Kimberly Espy
University of Nebraska, kespy2@unl.edu

Mary F. Cwik
Southern Illinois University School of Medicine,

Follow this and additional works at: http://digitalcommons.unl.edu/dcnlfacpub

Part of the Neurosciences Commons

Espy, Kimberly and Cwik, Mary F., "The Development of a Trial Making Test in Young Children: The TRAILS-P" (2004). *Developmental Cognitive Neuroscience Laboratory - Faculty and Staff Publications.* Paper 5.
http://digitalcommons.unl.edu/dcnlfacpub/5

This Article is brought to you for free and open access by the Developmental Cognitive Neuroscience Laboratory at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Developmental Cognitive Neuroscience Laboratory - Faculty and Staff Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *The Clinical Neuropsychologist* 18 (2004), pp. 411–422; doi 10.1080/138540409052416
Copyright © 2004 Taylor and Francis Ltd. Used by permission.
http://www.tandf.co.uk/journals/authors/ntcnauth.asp

Accepted for publication May 7, 2004.

The Development of a Trial Making Test in Young Children: The TRAILS-P

Kimberly Andrews Espy and Mary F. Cwik

Department of Family & Community Medicine,
Southern Illinois University School of Medicine, Carbondale, IL, USA

Corresponding author — K. A. Espy

Abstract

Preschool children have a more limited verbal repertoire, less proficient manual skills, and more variable attention spans relative to those of school age, with comparatively few neuropsychological tasks available for use in this age range. A prototypic neuropsychological test, the Trail Making Test, was adapted for use with young children, the TRAILS-P, using a developmentally salient storybook format with colorful stimuli in differing conditions with varying executive demands. The TRAILS-P was administered to 103 normally developing preschoolers between 2 and 6 years of age; 30 of these children were retested within one month to determine test reliability. Correlations among latencies to complete each condition and condition errors generally were moderate to high, suggesting coherence in test content. There also was evidence for good test–retest reliability. Latency to complete the TRAILS-P conditions differed as a function of the interaction of condition type and age group. Although the youngest children generally took more time to complete all TRAILS-P conditions, 3-year-old children were disproportionately slow to complete the condition that required shifting between stamping stimuli of two classes, with distraction by the additional presence of irrelevant stimuli. In contrast, the number of errors differed only in the 5-year-olds relative to younger children. These findings suggest that executive abilities can be assessed adequately in young children when tasks are designed to take advantage of the developmentally unique features of the preschool period.

Introduction

The Trail Making Test (TMT; Reitan, 1971) has been used extensively in adult neuropsychological research as an assessment of psychomotor speed, complex attention, and executive functions. Originally adapted from the Army Battery, the test consists of two parts. In Part A, the subject must connect consecutively numbered circles, whereas in Part B, the subject must connect consecutively numbered and lettered circles by alternating between the two sequences (e.g., connecting the circle with an "A" en-

This research was supported in part by grant R01 MH065668 from the National Institute of Mental Health, the Pediatric Neuropsychology/Developmental Cognitive Neuroscience Award from the Rita Rudel Foundation, and the Special Research Program Award from the Southern Illinois University Office of Research Development and Administration to the first author. The authors thank the participating families and undergraduate staff who assisted in various laboratory tasks associated with this study.

closed to the circle with the "1" enclosed, then to the circle with the "B" enclosed, etc.). Latency to complete each condition is scored. Importantly, if the subject makes an error while performing the task, the error is corrected. That is, the subject must return to the last correct stimulus and recommence the sequence. Therefore, the latency measure includes the time taken for correction (Lezak, 1995). The majority of psychometric studies report test–retest reliability coefficients above .60; in some studies, however, these reliabilities are in the range acceptable for use in clinical diagnosis, that is, in the .80s and .90s (Spreen & Strauss, 1991). As the TMT is included in the Halstead-Reitan Battery, early efforts using the TMT focused on the ability to detect brain damage. Lezak (1995) concluded that the TMT is highly sensitive to the effects of brain injury or damage in adults (e.g., Botwinick, Storandt, Berg, & Boland, 1988; Buchanan, Strauss, Kirkpatrick, Breier, & Carpenter, 1994; Dikmen, Machamer, Temkin, & McLean, 1990). TMT Part B also has been considered an index of prefrontal dysfunction because of its apparent requirement to flexibly shift response sets (e.g., Butters, Kaszniak, Glisky, Eslinger, & Schachter, 1994; Jarvis & Barth, 1994). In some studies (e.g., Anderson, Bigler, & Blatter, 1995; Stuss, Benson, Kaplan, Weir, & Della Malva, 1981), however, TMT Part B performance did not differ in patients with unselected brain damage that included the frontal lobe versus those with lesions that did not include frontal regions. In contrast, TMT Part B performance was correlated highly with caudate atrophy in patients with Huntington's disease (Starkstein et al., 1988). Segalowitz, Unsal, and Dwyan (1992) observed relations between the frontothalamic electrophysiological measures of the Contingent Negative Variation waveform and TMT performance in normal adults and adolescents. More recently, using modern neuroimaging methods to pinpoint lesion location, Stuss et al. (2001) studied 62 patients with single, focal lesions restricted to frontal or non-frontal regions. Patients were required to be free of severe aphasia and without detectable neglect or comorbid psychiatric or neurological disease. Patients with both right and left frontal lesions evidenced slowed latency to complete TMT Part B. Furthermore, patients with frontal lesions were 4.4 times more likely to make more than one error than were non-frontal patients, with those patients with damage to the dorsolateral prefrontal region committing the most errors.

Reitan (1971, 1992) published a TMT for use with children, which has become routine in the neuropsychological assessment of school age children in fixed battery (e.g., Reitan & Wolfson, 1992) and hypothesis testing, flexible battery approaches (e.g., Fletcher & Taylor, 1984). Neyens and Aldencamp (1996) reported test–retest reliability coefficients in a sample of 59 children between the ages of 4 and 12 years of .33 and .56 for Part A and Part B, respectively. TMT performance in school-aged children is sensitive to various types of central nervous system injury (Boll, Berent, & Richards, 1977; Jaffe et al., 1993; Reitan, 1971). For example, O'Leary et al. (1983) found that school-aged children and adolescents with early onset epileptic seizures (before age 5) performed more poorly on both Parts A and B compared to those children with late onset seizures. Knights et al. (1991) reported that children with severe traumatic brain injuries took more time to complete Parts A and B of the TMT than children with mild and moderate head injuries. However, Ewing-Cobbs et al. (1998) found that TMT performance differences were due to age, but not due to the severity of the injury, in a sample of 91 children who had incurred traumatic brain injuries.

The TMT also has been used in investigations of children with learning disabilities and attention problems (e.g., Mittelmeier, Rossi, & Berman, 1989; Shue & Douglas, 1992). In a study of second grade children, low achievers took significantly longer to complete both Parts A and B than average achievers (Kops & Belmont, 1985).

Some studies of reading disabled and nonreading disabled school-aged children have found significant differences on Part B only (e.g., McManis, Figley, Richert, & Fabre, 1978; Naerhi, Rasanen, Metsapelto, & Ahonen, 1997). In many studies, there is an association between TMT performance and attention behaviors reported by parents or teachers on standardized behavior checklists (Gorenstein, Mammato, & Sandy; 1989; Moffitt & Silva, 1988; Shue & Douglas, 1992). Peruguni, Harvey, Lovejoy, Sandstrom, and Webb (2000) found group differences on Part B between children diagnosed with attention deficit hyperactivity disorder (ADHD) and controls, with marginal predictive, discriminative power. However, in other studies, the specificity of TMT performance in ADHD diagnosis has not been demonstrated (McGee, Williams, Moffitt, & Anderson, 1989; Naerhi & Ahonen, 1995).

There is a renewed interest in instruments like the TMT, as disturbances in executive control have been implicated in an expanding number of neurological, medical, psychiatric, and developmental disorders. Many clinical neuropsychologists view the functions of prefrontal systems as fractionated, that is, supporting interrelated, yet separable, cognitive processes, although considerable debate remains about the organization of executive control in both adults and children. Cognitive neuroscience investigations have focused considerable efforts on describing conflict interference and working memory processes in executive control. Recently, however, finer examination of more complex, executive abilities, such as shifting (e.g., Rogers & Monsell, 1995) and planning (e.g., Goel & Grafman, 1995) have been undertaken. Concomitantly, there has been a resurgence in the interest in developing clinical tests to assess these cognitive processes, with the goal being to better characterize the cognitive organization and outcome in children with various disorders that affect brain function. In this vein, performance on the TMT appears to depend on multiple cognitive abilities, including visual acuity and scanning, attention, and visuomotor speed, with Part B including a higher demand for executive attention, maintaining response set, planning, and flexibility (Lezak, 1995; Baron, 2004).

In young children, there remains a paucity of measures available by which to assess executive skills in the preschool period (age 2 through 5 years), despite the emergence of several psychiatric and neurodevelopmental disorders in this age range. Because children affected with these disorders are considered to have unique profiles of executive dysfunction (e.g., Pennington, 1997), tasks with varying demands to measure discriminable executive processes would be useful. Although recently developed preschool tests (e.g., Elliott, 1990; Korkman, Kirk, & Kemp, 1998) have well-developed normative bases, their ability to assess specific facets of executive control is limited. Prototypic executive tests that are used with adults and school-aged children are often of limited utility with preschool children because of their more limited verbal repertoire, poorer manual skills, and variable attention span. By using creative, colorful stimuli in a format that is familiar to children, such as a storybook (Espy, 1997), traditional executive function tasks may be able to be adapted for use with young children. Therefore, the purpose of this article was to describe the development and psychometric properties of an adaptation of the TMT for use in young children, the TRAILS-P, and to determine whether there were age-and condition-related differences in task performance.

Methods

Participants

Participants were 103 preschool children (mean age = 4.46 years, SD = 0.92) who were recruited from the rural, southern Illinois community through a local healthcare facility, childcare facilities, and by word of mouth. To be included in the analyses reported here, children had to complete all four TRAILS-P conditions; 76% of the potential subject pool met this criterion. There were 58 females and 45 males, with 75% of children re-

ported as White/Caucasian and 25% as minority ethnicity. None of these children were diagnosed with any neurological, psychiatric, or developmental disorders on the basis of parental report. Children were divided into three age groups, 3-year-olds (n = 38; mean age = 3.47 years, SD = 0.36), 4-year-olds (n = 35, mean age = 4.62, SD = 0.29), and 5-years-olds (n = 30; mean age = 5.53, SD = 0.41). Mean child verbal intelligence estimate, measured by the Picture Vocabulary subtest from the WJ-R (Woodcock & Johnson, 1989), was 103.77 (SD = 13.10).

Mean maternal education level for the sample was 14.57 years (SD = 2.39), with 95% of mothers having a high school degree or above, and 44% having a college degree or higher. Sample mean income was $35,573 (SD = $23,220) and median income was $29,000. According to the 2000 Census (U.S. Census Bureau, 2000), the percentage of persons in the State of Illinois with a high school education or above was 81%, and the percentage with a college degree or higher was 26%. State median income was $46,500. Therefore, the obtained sample is somewhat more highly educated, but with fewer economic resources than the broader state population.

The proportion of males and females did not differ among the three age groups, $\chi^2(2, N = 103) = 1.23$, $p > .54$, nor did the proportion of children of White versus minority race = ethnicity, $\chi^2(2, N = 103) = 0.46$, $p > .80$. Neither mean maternal education level, $F(2, 99) = 1.06$, $p > .34$, nor mean WJ-R Picture Vocabulary sub-test standard score, $F(2, 99) = 0.02$, $p > .97$, differed across the three age groups.

In order to determine the reliability of the TRAILS-P, 30 of these children were re-administered the task within 1 month of the initial administration (mean test–retest interval = 15.2 days, SD = 8.51). These 30 children were selected randomly for readministration. The mean age of the reliability subsample was 4.64 years (SD = 0.85). Seventy percent of the subsample was female and 30% was male; 67% of the reliability subsample was of White/Caucasian ethnicity and 33% was of minority ethnicity.

Measure

In the TRAILS-P, children were presented with a book with colorful dog characters. The children were told, "Here is a family of doggies. The littlest one is the baby dog, then the sister dog, then the brother dog. The Mommy dog is here, and the biggest dog, the Daddy dog, is right here. This dog family lives in this house." The children were instructed to identify all of the dogs, in order of size, to ensure adequate understanding. Children then were provided an inked stamp with a child-size handle for easy gripping. In Condition A (control), the children were instructed to stamp the dogs in order of size, starting with the "Baby" through to the "Daddy." Condition B (switch) involved the introduction of like-sized bones, which the child had to "match" to the dogs, that is, flexibly shift among the like-sized stimuli, in order. In order to assess the effects on task performance of reversing response contingencies, in Condition C (reversal), the child stamped the dogs in order of size, but now had to ignore the previously presented salient stimuli, the bones. Condition D (distraction) assessed the effects of distraction by intermixing cat stimuli as distractors with the target dogs and bones. Again, the child had to alternate stamping the dogs and then the relevant bones, in size order, but while ignoring the cats. For each condition, the latency to stamp all stimuli (with correction for wrong stamps as in the original TMT) and the number of errors were scored.

Procedure

Young children were administered a larger task battery that included the TRAILS-P. A trained child clinical graduate student administered the battery in a single session in a quiet room with the parent or guardian present. The parent was present because some younger children initially have difficulty separating or remaining separated for the duration of the testing session. To maintain constant testing conditions for all children, the parent was seated in the back of the testing room. Because the parent was kept occupied completing questionnaires during the child's evaluation, there was minimal observable impact on the child's performance. Breaks were used when necessary to maintain cooperation and interest. Test sessions were videotaped for the purpose of later scoring.

Design and Analysis

First, bivariate correlations were calculated using SAS V8 for Windows between the latencies and numbers of errors in each TRAILS-P condition in order to determine how the different test conditions were related. Spearman correlations were used because of the non-normal distributions of the error variables. Test–retest correlations, also using Spearman correlations that are less sensitive to violations of normality, were calculated to determine evidence for reliability. Then, to investigate the psychometric and substantive properties of the test, several mixed-factorial multivariate analyses were con-

ducted. First, an omnibus MANOVA was conducted to determine whether children of different age groups performed differently across the TRAILS-P conditions, with separate analyses for the latency and error scores. Then, several planned comparisons were conducted to investigate whether the varying condition task demands affected performance differentially in children of differing age groups. Where the omnibus mixed model (TRAILS-P conditions as a within subjects variable, age group as the between subjects variable) MANOVA was significant, performance between 1) Conditions A and B were compared among the age groups to determine the differential cost of shifting responses between two relevant stimulus sets, 2) Conditions B and C were compared among age groups to determine the differential reversal costs when requiring the child to now inhibit responding to a previously salient target stimulus class, and 3) Conditions B and D were compared among age groups to investigate the differential cost of inhibiting a newly introduced, irrelevant stimuli.

Results

Latencies from each of the four TRAILS-P conditions were correlated, ranging in magnitude from .52 to .75, shown in Table 1. Condition errors also were related, ranging from .31 to .51. Latency and errors were correlated significantly within each condition (range .52 to .66), and across condition (range .25 to .47) with the exception of latency to complete Condition A and error on Condition B. Finally, the correlations between test and retest administrations were good (depicted on the diagonal of the top of Table 1), ranging from .45 to .77, with a mean value of .64 averaged across the four conditions.

The latency to complete the TRAILS-P varied as a function of the interaction of condition and age group, Wilks's $\Lambda = .85$, $F(6, 196) = 2.83$, $p < .02$. The main effects of TRAILS-P condition, Wilks's $\Lambda = .80$, $F(3, 98) = 7.99$, $p < .0001$, and age group, $F(2, 100) = 25.10$, $p < .0001$, also were significant, indicating that TRAILS-P latency differed among conditions and between age groups. The mean latencies to complete each of the four conditions for the three age groups are depicted in Table 2. Examination of the means and performance variability was consistent with different patterns of

Table 1. TRAILS-P Condition Intercorrelations and Test–Retest Reliabilities (N =103).

Condition	Condition			
	A (Control)	B (Switch)	C (Reversal)	D (Distraction)
Latency–Latency				
A	.64****			
B	.58****	.45**		
C	.63****	.52****	.77****	
D	.61****	.70****	.75****	.69****
Errors–Latency				
A	.66****			
B	.18	.52****		
C	.38****	.25*	.66****	
D	.30**	.47****	.46****	.66****
Errors–Errors				
A	–			
B	.31**	–		
C	.38****	.35***	–	
D	.32**	.46****	.51****	–

Diagonal elements represent test–retest reliabilities using Spearman correlations in the top section of the table.
* $p < .05$; ** $p < .01$; *** $p < .001$; **** $p < .0001$.

Table 2. Normative Data by Age Group and Trails-P Condition (N=103).

Measure	Trails-P condition							
	A (Control)		B (Switch)		C (Reversal)		D (Distraction)	
	M	SD	M	SD	M	SD	M	SD
3-Year-olds								
Latency (s)	42.58	20.01	37.61	16.12	36.37	14.35	44.97	19.56
Errors	1.66	1.91	1.47	1.69	2.00	2.22	2.16	2.51
4-Year-olds								
Latency (s)	32.86	19.54	24.34	16.31	20.97	8.06	25.20	13.49
Errors	1.63	1.99	0.91	1.44	1.31	1.81	1.23	1.99
5-Year-olds								
Latency (s)	21.23	17.51	20.47	13.87	20.10	14.61	19.57	11.99
Errors	0.60	1.16	0.77	1.36	0.87	1.31	0.40	0.81

(s) = seconds.

performance among conditions in the three age groups. Because the omnibus test indicated that latency to complete the TRAILSP conditions varied as a function of the interaction of condition type and age group, results from the three planned comparisons investigating the impact of the differing condition task demands were conducted.

First, to assess the effect of shifting responding between relevant stimulus sets, latency to complete Conditions A and B were compared between the 3- and 4-year-old age groups, and between the 4- and 5-year-old groups, shown in the top panel of Figure 1. The relative difference in the latencies to complete Conditions A and B was comparable in the 3- and 4-year-olds, age group by task type interaction Wilks's $\Lambda = .99$, $F(1, 71) = 0.57$, $p > .45$. Generally, it took both 3- and 4-year-old children less time to complete Condition B than Condition A, Wilks's $\Lambda = .90$, $F(1, 71) = 8.25$, $p < .001$, and 3-year-olds took more time to complete the conditions on average than did 4-year-olds, $F(1, 71) = 10.61$, $p < .002$. Latencies to complete Conditions A and B were comparable between 4- and 5-year old children, age group by task type interaction Wilks's $\Lambda = .96$, $F(1, 63) = 2.26$, $p > .10$. Again, children generally completed Condition B in less time than Condition A, Wilks's $\Lambda = .94$, $F(1, 63) = 3.82$, $p < .06$, and 4-year-olds took more time on average to complete the conditions than did 5-year-olds, $F(1, 63) = 4.89$, $p < .03$.

To assess the effect of reversing the salient response contingencies, where the child must inhibit stamping the previously salient stimulus that is now irrelevant, latencies to complete Conditions B and C were compared at each age group, depicted in the middle panel of Figure 1. The difference in the latencies to complete Condition C and B between 3- and 4-year-old children was comparable, age group by task type interaction Wilks's $\Lambda = .99$, $F(1, 71) = 0.33$, $p > .56$. Generally, the latency to complete Conditions C and B was comparable, Wilks's $\Lambda = .98$, $F(1, 71) = 1.52$, $p > .22$. However, there were age-related performance differences, as 3-year-olds took more time to complete the conditions on average than did 4-year-olds, $F(1, 71) = 27.29$, $p < .0001$. The difference in the latencies to complete Conditions C and B were comparable between 4- and 5-year-old children, age group by task type interaction Wilks's $\Lambda = .99$, $F(1, 63) = 0.55$, $p > .46$. In this age group comparison, however, latency to complete Conditions C and B was comparable, Wilks's $\Lambda = .99$, $F(1, 63) = 0.85$, $p > .35$, as was condition completion latency between 4- and 5-year-old children, $F(1, 63) = 0.78$, $p > .38$.

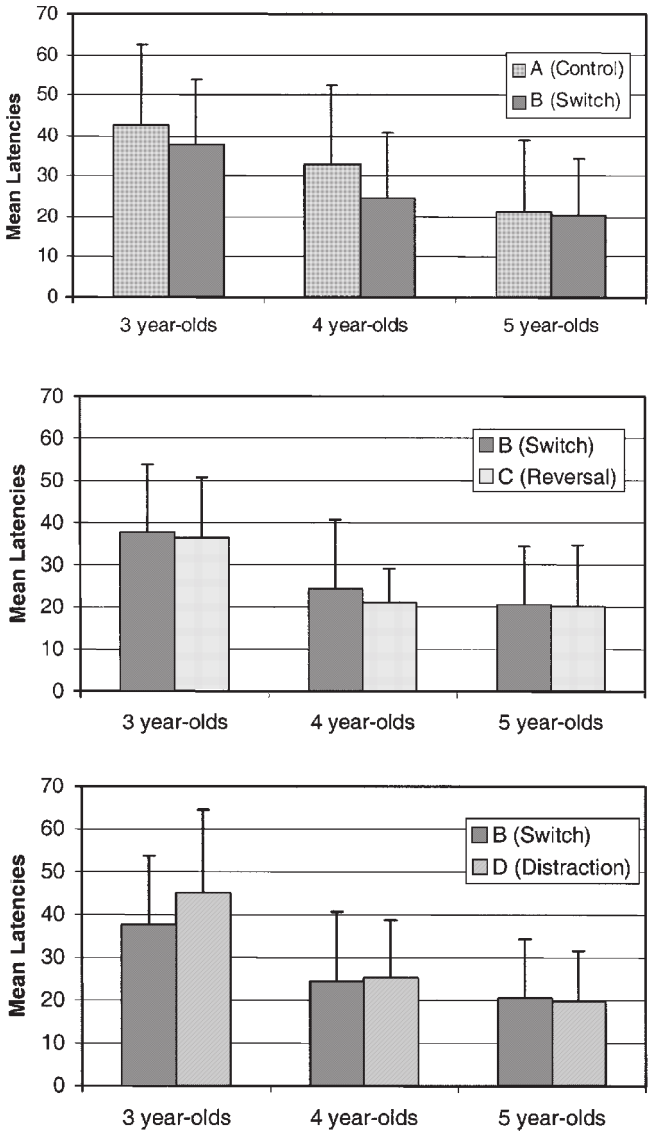

Figure 1. *Top.* Mean latency (in seconds) to complete Conditions A and B as a function of age group. *Middle.* Mean latency to complete Conditions B and C as a function of age group. *Bottom.* Mean latency to complete Conditions B and D as a function of age group.

Finally, the effects of distraction were investigated by comparing latency to complete Conditions B and D, contrasted between the 3- and 4-year-old age group and between the 4- and 5-year-old groups, as shown in the bottom panel of Figure 1. The relative difference in the latencies to complete Conditions D and B differed marginally between the 3- and 4-year-old age groups, age group by task type interaction Wilks's $\Lambda = .96$, $F(1, 71) = 3.02$, $p < .09$, where younger children took disproportionately longer to complete Condition D. Averaging across age groups, children took more time to complete Condition D than B, Wilks's $\Lambda = .94$, $F(1, 71) = 4.82$, $p < .04$, and 3-year-olds took more time to complete the conditions on average than 4-year-olds,

$F(1, 71) = 23.58$, $p < .0001$. In contrast, the difference in the latencies to complete Conditions D and B were comparable between 4- and 5-year-old children, age group by task type interaction Wilks's $\Lambda = .99$, $F(1, 63) = 0.28$, $p > .59$. Children generally completed Conditions D and B with similar latencies, Wilks's $\Lambda = .99$, $F(1, 63) = 0.00$, $p > .99$, and 4- and 5-year-olds completed the conditions in comparable latencies, $F(1, 63) = 2.37$, $p > .12$.

The number of TRAILS-P errors did not vary as a function of the interaction of condition and age group, Wilks's $\Lambda = .91$, $F(6, 156) = 1.65$, $p > .13$. Furthermore, there were no main effect differences in the number of errors among the different TRAILSP conditions, Wilks's $\Lambda = .96$, $F(3, 98) = 1.16$, $p > .32$. Across conditions, the number of errors differed between age groups, $F(2, 100) = 6.32$, $p < .003$, with 4-year-olds making more errors across conditions than 5-year-olds, $F(1, 26) = 6.00$, $p < .03$, but no differences in errors were apparent between the 3- and 4-year-old age groups, $F(1, 28) = 0.29$, $p > .59$.

Discussion

Psychometrically, the TRAILS-P performed well. The latencies among the conditions were related strongly, as would be expected given the general stability of individual differences in psychomotor speed in childhood (Baron, 2004). In addition, latencies and errors in each condition also were related, consistent with the well-known association between response speed and selection accuracy, even in very young participants. Preschool children are more variable in their effort, attention, and concentration. In this context, the average test–retest correlation of .64 is notable. The reliability of the latency to complete Condition B was lower than that of the other conditions, indicating that performance was more variable between test and retest for this condition. Although shifting between extra-dimensional sets develops rapidly in this age range (e.g., Espy, Bull, & Martin, 2004; Jacques & Zelazo, 2001), it is unlikely that significant development occurred during the average 2-week interval that contributed to reduced performance reliability. More likely, a larger item set, to better sample simple shifting behavior, would lead to improved condition test–retest reliability.

The substantive differences in the latencies to complete the TRAILS-P conditions in children of differing age groups suggest that there are developmental differences in the cognitive processes required to meet the differing task demands of the conditions. Comparing simple stimulus identification (Condition A) from that requiring simple shifting among extra-dimensional sets, the youngest children took more time to complete the conditions than middle age groups, who in turn took more time than the 5-year-olds. Note, however, the general reduction in latency to complete Condition B relative to Condition A, consistent with strong practice effects that carry over between these two conditions, apparent in both the 3- and 4-year-old, and the 4- and 5-year-old comparisons. These results suggest that the practice effects of the general reduction in stamping latency between Conditions B and A may have obscured any potential difference in any costs associated with shifting. A randomized condition order could be used to test this hypothesis further; however, such a procedure is unlikely to be adopted in the clinical assessment context. Another option might be to present Condition A twice before proceeding to Condition B, to reduce the magnitude of the practice effect on Condition B performance. In Condition B, children have to shift between two dimensions that are related more strongly in an associative manner (dog to bone), unlike the dimensions of the TMT (letters and numbers). This stronger associative relation between dimensions also might have contributed to the attenuated switching costs observed here.

In the comparisons designed to test developmental differences in the abilities to respond to reversed contingencies, such that children had to inhibit stamping the previously salient target class, performance was comparable between the conditions across age groups. Regardless of age group, children were able to respond to the simple change in response contingencies. The lack of age group differences in performance suggests that these abilities develop earlier in ontogeny and are fully mature at the ages observed here. Alternatively, the reversal task might have been too simple to actively engage inhibitory processes in this age range.

In contrast, the impact of distraction was limited largely to the youngest of children. The youngest children took correspondingly more time to complete Condition D in comparison to Condition B than would be expected based on age or condition. In the two older age groups, latency to complete the conditions was comparable. In the context of developmental reductions in psychomotor speed, the younger child took disproportionate time to complete the condition that included distraction. Interestingly, these younger children were particularly affected by distraction, in the context of general improvements in psychomotor speed. Such manipulations are important to better understand the relative impact of distractor and conflict interference on the developmental organization of executive abilities.

There were not consistent differences in the number of errors among TRAILS-P conditions, nor were the errors on a given condition disproportionate in an age group. The oldest children made fewer errors than the younger children, with no differences between the 3-and 4-year-olds. This developmental pattern is different from what was found for latencies, where there were either progressive or most marked differences between the youngest and two older groups. Given the strong relation between speed and accuracy, 4-year-old children are actively learning to modulate their responding, and the improvement in speed at this age comes with a cost of accuracy. It is only at age 5 that children can progressively stamp quickly and accurately, switching between relevant sets, reversing response contingencies, and maintaining focus in light of distraction.

These differences are consistent with a growing literature base concerning the development of executive control in preschool children (e.g., Diamond et al., 1997; Espy, 1997; Espy, Kaufmann, McDiarmid, & Glisky, 1999; Espy, Kaufmann, Glisky, & McDiarmid, 2001; Hughes, 1998; Jacques & Zelazo, 2001). Although these findings are compelling, a different developmental pattern might be evident on a different shifting task or in sample of different demographic characteristics. The observed pattern of development in task performance is a function of the child's abilities and cognitive proficiencies, as well as variation in task demands. What might appear to be growth in discrete cognitive abilities may actually be changes in task demands as a function of age. This issue is particularly important in the study of executive control, where task demands such as novelty, difficulty, salience, and expectancy are known to affect frontal activation, as demonstrated by studies in adults using functional magnetic resonance imaging (e.g., Barch et al., 1997; Casey et al., 2001; Petersen, van Mier, Fiez, & Raichle, 1998; Rogers, Owen, Middleton, Williams, & Pickard, 1999). Longitudinal growth modeling that accounts for growth in latent task demands and abilities is necessary to understand fully such issues.

Generally, these findings demonstrate the feasibility of adapting prototypical executive function tasks, such as the TMT, for use in young children. The TRAILS-P is unique in this regard, using engaging stimuli with an age appropriate manual response. Tasks must be adapted for use with young children carefully, simultane-

ously considering the developmental abilities in this age range, idiosyncratic interests of young children, and limits in attention that make this age group unique. Preschool children do not represent "smaller, less able" children, just as children are not "little adults." Although the TRAILS-P appears to have good psychometric properties and performance varied as a function of condition task demands and age group, evidence for convergent and discriminant validity with other standardized instruments must be demonstrated before more widespread clinical application is undertaken. However, based on the psychometric results presented here, the TRAILS-P may offer promise as a tool to assess the processes involved in executive control in young children with neurological, psychiatric, and developmental disorders.

References

Anderson, C. V., Bigler, E. D., & Blatter, D. D. (1995). Frontal lobe lesions, diffuse damage, and neuropsychological functioning in traumatic brain-injured patients. Journal of Clinical and Experimental Neuropsychology, 17, 900–908.

Barch, D., Braver, T., Nystron, L., Forman, S., Noll, D., & Cohen, J. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. Neuropsychologia, 35, 1373–1380.

Baron, I. S. (2004). Neuropsychological evaluation of the child. New York: Oxford University Press.

Boll, T. J., Berent, S., & Richards, H. (1977). Tactile-perceptual functioning as a factor in general psychological abilities. Perceptual and Motor Skills, 44, 535–539.

Botwinick, J., Storandt, M., Berg, L., & Boland, S. (1988). Senile dementia of the Alzheimer's type: Subject attrition and testability in research. Archives of Neurology, 45, 493–496.

Buchanan, R. W., Strauss, M. E., Kirkpatrick, C. H., Breier, A., & Carpenter, J. (1994). Neuropsychological impairments in deficit and non-deficit forms of schizophrenia. Archives of General Psychiatry, 51, 801–811.

Butters, M. A., Kaszniak, A. W., Glisky, E. L., Eslinger, P. J., & Schachter, D. L. (1994). Recency discrimination deficits in frontal lobe patients. Neuropsychology, 8, 343–353.

Casey, B. J., Forman, S., Franzen, P., Berkowitz, A., Braver, T., Nystrom, L., et al. (2001). Sensitivity of prefrontal cortex to changes in target probability: A functional MRI study. Human Brain Mapping, 13,26–33.

Diamond, A., Prevor, M. B., Callender, G., & Druin, D. P. (1997). Prefrontal cortex cognitive deficits in children treated continuously for PKU. Monographs of the Society for Research in Child Development, 62,1–205.

Dikmen, S., Machamer, J., Temkin, N., & McLean, A. (1990). Neuropsychological recovery in patients with moderate to severe head injury: 2-Year follow-up. Journal of Clinical and Experimental Neuropsychology, 12, 507–512.

Elliott, C. D. (1990). The nature and structure of children's abilities: Evidence from the Differential Ability Scales. Journal of Psychoeducational Assessment, 8, 376–390.

Espy, K. A. (1997). The Shape School: Assessing executive function in preschool children. Developmental Neuropsychology, 13, 495–499.

Espy, K. A., Bull, R. B., & Martin, J. M. (2004). Measuring the development of executive control with the Shape School. Manuscript submitted for publication.

Espy, K. A., Kaufmann, P. M., Glisky, M. L., & McDiarmid, M. D. (2001). New procedures to assess executive functions in preschool children. The Clinical Neuropsychologist, 15, 46–58.

Espy, K. A., Kaufmann, P. M., McDiarmid, M. D., & Glisky, M. L. (1999). Executive functioning in preschool children: Performance on A-not- B and other delayed response format tasks. Brain and Cognition, 41, 178–199.

Ewing-Cobbs, L., Prasad, M., Fletcher, J. M., Levin, H. S., Miner, M. E., & Eisenberg, H. M. (1998). Attention after pediatric traumatic brain injury: A multidimensional assessment. Child Neuropsychology, 4,35–48.

Fletcher, J. M., & Taylor, H. G. (1984). Neuropsychological approaches to children: Towards a developmental neuropsychology. Journal of Clinical Neuropsychology, 6,39–56.

Goel, V. & Grafman, J. (1995). Are the frontal lobes implicated in "planning" functions? Interpreting data from the Tower of Hanoi. Neuropsychologia, 33, 623–642.

Gorenstein, E. E., Mammato, C. A., & Sandy, J. M. (1989). Performance of inattentive-overactive children on selected measures of prefrontal- type function. Journal of Clinical Psychology, 45, 619–632.

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. British Journal of Developmental Psychology, 16, 233253.

Jacques, S. & Zelazo, P. D. (2001). The Flexible Item Selection Task (FIST): A measure of executive function in preschoolers. Developmental Neuropsychology, 20, 573591.

Jaffe, K. M., Fay, G. C., Polissar, N. L., Martin, K. M., Shurtleff, H. A., Rivara, J. M. B., et al. (1993). Severity of pediatric traumatic brain injury and neurobehavioral recovery at one year: A cohort study. Archives of Physical Medicine and Rehabilitation, 74, S87S95.

Jarvis, P. E. & Barth, J. T. (1994). The Halstead-Reitan Neuropsychological Battery: A guide to interpretation and clinical applications. Odessa, FL: Psychological Assessment Resources.

Knights, R. M., Ivan, L. P., Ventureya, E. C., Bentivoglio, C., Stoddart, C., Winogron, W., et al. (1991). The effects of head injury in children on neuropsychological and behavioral functioning. Brain Injury, 5, 339–351.

Kops, C. & Belmont, I. (1985). Planning and organizing skills of poor school achievers. Journal of Learning Disabilities, 18, 8–14.

Korkman, M., Kirk, U., & Kemp, S. (1998). NEPSY: A developmental neuropsychological assessment manual. San Antonio, TX: The Psychological Corporation.

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

McGee, R., Williams, S., Moffitt, T. E., & 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.

McManis, D. L., Figley, C., Richert, M., & Fabre, T. (1978). Memory-For-Designs, Bender-Gestalt, Trail Making Test, and WISC-R performance of retarded and adequate readers. Perceptual and Motor Skills, 46, 443–450.

Mittelmeier, C., Rossi, J. S., & Berman, A. (1989). Discriminative ability of the Trial Making Test in young children. International Journal of Clinical Neuropsychology, 11, 163–166.

Moffitt, T. E., & Silva, P. A. (1988). Self-reported delinquency, neuropsychological deficit, and history of Attention Deficit Disorder. Journal of Abnormal Child Psychology, 16, 553–569.

Naerhi, V., & Ahonen, T. (1995). Reading disability with or without attention deficit hyperactivity disorder: Do attentional problems make a difference? Developmental Neuropsychology, 11, 337–349.

Naerhi, V., Rasanen, P., Metsapelto, R. L., & Ahonen, T. (1997). Trail Making Test in assessing children with reading disabilities: A test of executive functions or content information. Perceptual and Motor Skills, 84, 1355–1362.

Neyens, L. G. J. & Aldenkamp, A.P. (1996). Stability of cognitive measures in children of average ability. Child Neuropsychology, 2, 161–170.

O'Leary, D. S., Lovell, M. R., Sackallares, J. C., Berent, S., Giordani, B., Seidenberg, M., et al. (1983). Effects of age of onset of partial and generalized seizures on neuropsychological performance in children. The Journal of Nervous and Mental Disease, 171, 624–629.

Pennington, B. F. (1997). Dimensions of executive functions in normal and abnormal development. In N. A. Krasnegor, G. R. Lyon, & P. S. Goldman-Rakic (Eds.), Development of the prefrontal cortex (pp. 265–281). Baltimore: Paul H. Brookes Publishing Co., Inc.

Perugini, E. M., Harvey, E. A., Lovejoy, D. W., Sandstrom, K., & Webb, A.H. (2000). The predictive power of combined neuropsychological measures for Attention-Deficit/Hyperactivity Disorder in children. Child Neuropsychology, 6, 101–114.

Petersen, S., van Mier, M., Fiez, J., & Raichle, M. (1998). The effects of practice on the functional anatomy of task performance. Proceedings of the National Academy of Sciences USA, 95, 853–860.

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

Reitan, R. M. (1992). Trail Making Test: Manual for administration and scoring. South Tuscon, AZ: Reitan Neuropsychology Laboratory.

Reitan, R. M. & Wolfson, D. (1992). A short screening examination for impaired brain functions in early school-age children. Clinical Neuropsychologist, 6, 287–294.

Rogers, R., Owen, A., Middleton, H., Williamson, E., & Pickard, J. (1999). Choosing between small, likely rewards and large unlikely rewards activates inferior and orbital prefrontal cortex. Journal of Neuroscience, 19, 9029–9038.

Rogers, R. D. & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. Journal of Experimental Psychology: General, 124, 207–231.

Segalowitz, S., Unsal, A., & Dywan, J. (1992). CNV evidence for the distinctiveness of frontal and posterior neural processes in a traumatically brain injured population. Journal of Clinical and Experimental Neuropsychology, 14, 545–565.

Shue, K. L. & Douglas, V. I. (1992). Attention deficit hyperactivity disorder and the frontal lobe syndrome, Brain and Cognition, 20, 104–124.

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

Starkstein, S. E., Brandt, J., Folstein, S. R., Strauss, M., Berthier, M. L., Pearlson, G. D., et al. (1988). Neuropsychological and neuroradiological correlates in Huntington's disease. Journal of Neurology, Neurosurgery, and Psychiatry, 51, 1259–1263.

Stuss, D. T., Benson, D. F., Kaplan, E. F., Weir, W. S., & Della Malva, C. (1981). Leucotomized and non-leucotomized schizophrenics: Comparison on tests of attention. Biological Psychiatry, 16, 1085–1100.

Stuss, D. T., Bisschop, S. M., Alexander, M. P., Levine, B., Katz, D., & Izukawa, D. (2001). The Trail Making Test: A study in focal lesion patients. Psychological Assessment, 13, 230–239.

Woodcock, R. & Johnson, M. B. (1989). Woodcock Johnson Psycho-Educational Battery-Revised. Itasca, IL: Riverside Publishing Company.