Abstract

Studies of planning ability typically involve some version of the Tower of Hanoi or Tower of London (TOL). When these tests are administered to patients with multiple sclerosis (MS), the findings pertaining to planning “performance” have been conflicting. Possible reasons for failures to find deficits in planning performance among MS patients are: (a) the patients typically have relapsing–remitting MS (RRMS) of mild severity and short duration and thus little cognitive impairment relative to those with more advanced disease; (b) the problems composing the tests are too simple and differences between patients and controls are therefore obscured by ceiling effects; and (c) the scoring system typically used permits participants to earn points for successful solutions on later trials after failing the initial attempt on each problem, thereby further diluting the difference between patients and controls. The present study compared the performance of patients with both relapsing–remitting and secondary progressive disease with that of healthy controls on a more challenging version of the TOL. Patients exhibited lengthier planning times on the test, greater disparity in their average planning times from those of controls as the difficulty level increased, and greater individual variability in their planning times across the full set of problems. However, no differences in planning performance were found between patients and controls or between RRMS and secondary progressive MS patients. Performance differences in other studies may be attributable in part to the imposition of time limits for solving each problem and the disproportionately adverse effect such time limits have on patients’ performance.

Keywords: Executive function; Neuropsychological tests; Tower of London; Cognitive impairment; Information processing speed; Cognitive tempo

Introduction

Executive function is frequently mentioned as one of the cognitive domains adversely impacted by multiple sclerosis (MS; Brassington & Marsh, 1998; Drew, Tippett, Starkey, & Isler, 2008, Drew, Starkey, & Isler, 2009; Fischer et al., 1994; Foong et al., 1997, 1999; McIntosh-Michaelis et al., 1991; Rao, 1996; Ryan, Clark, Klonoff, Li, & Paty 1996; Zakzanis, 2000). Some recent studies indicate that impairment in executive function is evident very early in the course of the disease (Roca et al., 2008) and even in patients with clinically isolated syndrome (Schulz, Kopp, Kunkel, & Faiss, 2006) or isolated optic neuritis (Nilsson et al., 2008). The problem with claims such as these is that the term “executive function” encompasses a bewildering variety of cognitive operations, including the inhibition of automatic responses, fluency, abstraction, sequencing, planning, reasoning, strategy shifting, multitasking, selective attention, and resistance to interference (Chan, Shum, Touloumpoulou, & Chen, 2008). This problem becomes especially troublesome when insufficient information is given concerning the measures of executive function that have been employed in a particular study (e.g., Achiron et al., 2007; Summers et al., 2008). The likelihood of constructing a cumulative body of findings concerning such a multifaceted construct in MS or any other neurological disease would seem to be remote.

A possible way to address this problem is to focus on a subset of these operations, especially as they bear on patients’ performance on a particular type of problem-solving executive task. Two obvious examples of this approach that have been
employed with sufficient frequency within the MS literature involve the examination of rule abstraction and strategy shifting afforded by various types of card sorting tasks and the examination of planning afforded by spatial assembly tasks patterned after the Tower of Hanoi or Tower of London (TOL). In both instances, findings with respect to MS are inconsistent. Differences between MS patients and controls in performance on card sorting tasks have been reported in some studies (e.g., Beatty, Hames, Blanco, Paul, & Wilbanks 1995; Beatty & Monson, 1996; Heaton, Nelson, Thompson, Burks, & Franklin 1985; Rao, Hammekke, & Speech, 1987, Rao, Leo, Bernardin, & Unverzagt, 1991; Parmenter, Shucard, & Shucard, 2007), but not others (Denney, Lynch, & Parmenter, 2008; Denney, Lynch, Parmenter, & Horne, 2004; Denney et al., 2004; Drew et al., 2008; Helek et al., 2010; Penny, Khaleeli, Cipolott, Thompson, & Ron, 2010). Ironically, deficiencies in performance on the Wisconsin Card Sorting Test have recently been reported for patients in the very earliest stages of MS (Nilsson et al., 2008; Schulz et al., 2006), and yet patients with MS of over 30 years duration were also recently shown to perform as well as controls on this same test (Smestad, Sandvik, Landro, & Celius, 2010).

The focus of the present study is on planning, and here too the research is plagued by inconsistencies. Most studies utilizing various “tower” tasks report two principal types of outcome measures: a point score reflecting the efficacy of participants’ performance in solving the problems composing the task (i.e., “planning performance”) and a measure of the average length of time participants take formulating their solution before launching their first move on each problem (“planning time”). There is good agreement with respect to planning time. All studies that have included this measure report slower planning times for MS patients than controls (Arnett et al., 1997; Denney et al., 2004, Denney, Sworowski, & Lynch, 2005, 2008; Denney, Gallagher, & Lynch, 2011; Foong et al., 1997), and in some of these studies (Denney et al., 2004, 2011), the disparity in the planning times for patients and controls tends to become greater as the problems increase in difficulty.

The inconsistencies in the literature involve planning “performance.” Lower point scores for MS patients have been reported by several investigators (Arnett et al., 1997; Foong et al., 1997, 1999; Lazerson, Rombouts, Scheltens, Polman, & Barkhof, 2004), but in all of our own studies (Denney et al., 2004, 2005, 2008, 2011), we have failed to find a difference in the point scores of MS patients and controls. Furthermore, Drew and colleagues (2008) reported that the mean point score for a large sample of MS patients on the Tower Test of the Delis–Kaplan Executive Function System (Delis, Kaplan, & Kramer, 2001) was equal to that of the sample on which the test was normed.

As the one group that has consistently failed to find evidence of diminished planning performance in direct comparisons between MS patients and healthy controls, we conducted the present study as a challenge to our previous work. In each of the previous studies, we used a common 12-problem version of the TOL (Krikorian, Bartok, & Gay, 1994), with problems graduated in difficulty from those requiring only two moves to those requiring five moves to solve. Participants were allowed up to three trials for each problem and were awarded 3, 2, or 1 point if they succeeded in solving the problem on their first, second, or third attempt, respectively. In all of these studies, the mean point scores for MS patients were slightly lower than those of controls, and the means for both groups were fairly close to the maximum possible score of 36 (e.g., patients, \( M = 31.7 \); controls, \( M = 33.0 \); Denney, Gallagher, & Lynch, 2011). We speculated that perhaps the test was too easy and ceiling effects were obscuring the differences in planning ability between the groups. A related problem might lie in the point system which permitted participants to earn points on the second or the third trial, even after failing the first. Finally, many of the patients in these studies were in the relapsing–remitting stage of MS. Arnett and colleagues (1997) found that deficits in planning performance were limited to patients with chronic progressive MS. More generally, a meta-analysis of the literature by Zakzanis (2000) has indicated that frontal-executive deficits are more characteristic of the later, chronic progressive stage of MS. Therefore, the present study was designed to examine planning performance in separate samples of relapsing–remitting (RRMS) and secondary progressive (SPMS) patients, and to do so using a longer, more challenging version of the TOL. With these modifications in place, we hypothesized that differences in planning performance would be found—particularly in the group of SPMS patients.

Methods

Participants

The sample consisted of 46 patients with clinically definite MS (Poser et al., 1983) and 45 healthy controls. All the patients were under the care of the same board-certified neurologist (SGL) at the University of Kansas Medical Center, who limited recruitment to individuals with sufficient cognitive ability to provide informed consent and understand the study instructions. The following additional exclusionary criteria were used: (a) the presence of any neurological disorder other than MS; (b) past or present alcohol or substance abuse; (c) visual impairment exceeding 20/50 or color-blindness; (d) the current use of narcotics or benzodiazepines; and (e) any exacerbation of MS symptoms in the previous 30 days. The patients (34 females, 12 males) ranged between 28 and 63 years of age (\( M = 49.0 \)) and had from 12 to 19 years of education (\( M = 15.4 \)). Duration of diagnosis
ranged from 1 to 29 years ($M = 10.8$). Disability was rated by the neurologist using the Disease Steps scale (Hohol, Orav, & Weiner, 1995) and ranged from 0 to 6 ($M = 2.5$), with 65% of the patients rated 1 (“mild”), 2 (“moderate”), or 3 (“gait disability”). There were 29 patients with RRMS and 17 with SPMS.

The healthy controls (28 females, 17 males) were recruited from the metropolitan area of Kansas City. They had no chronic medical problems, were following no ongoing medication regimen, and met the same exclusionary criteria as the MS patients. The controls ranged in age from 22 to 58 ($M = 46.1$) and had between 12 and 18 years of education ($M = 15.5$).

**Measures**

**The Tower of London.** The computerized version of the TOL was based on the test originated by Shallice (1982) and developed further by Krikorian and colleagues (1994). In the upper portion of the screen, the computer displayed three colored disks arranged on three pegs (starting position). In the lower portion of the screen, the computer displayed a model with the disks in a different arrangement on the pegs (goal position). The participant’s task was to move the disks in the upper display so they matched the arrangement in the lower display and to do so using a specified number of moves. Moves were made by verbal dictation, the participant stating first the color of the disk and then the number of the peg to which it was to be moved. The experimenter immediately executed each move in accordance with the participant’s direction. Three sample problems (two 1-move problems and one 2-move problem) were used to introduce the task and acclimate the participant to the verbal method of responding; each sample problem was repeated until solved successfully. Twenty-four test problems were then presented, graduated in difficulty, with one 1-move, five 2-move, six 3-move, six 4-move, and six 5-move problems. There were four different starting positions for the disks in the upper display, and the starting position varied from problem to problem.

The number of moves permitted for each problem was announced prior to displaying the problem so that silence prevailed once the problem was displayed. This was important because the computer recorded a covert measure of the length of time between the display of the problem and the participant’s announcement of the color of the first disk to be moved (“planning time”). Participants were allowed three attempts to solve each problem in the specified number of moves and were awarded 3, 2, or 1 point for success on the first, second, or third attempt, respectively. The total point score was the sum of the points earned on all 24 problems. In addition, the number of problems solved correctly on the first trial was recorded.

Planning times for each trial of a problem were recorded, but only those for the first trial (regardless of whether the trial was successful) were used. These initial planning times were first examined for outliers in accordance with procedures used in previous studies (Burton, Strauss, Hultsch, Moll, & Hunter, 2006; de Frias, Dixon, Fisher, & Camicioli, 2007). Extremely slow responses, that is, those exceeding the corresponding group mean for that problem by at least 3 SD, were eliminated and replaced with imputed values determined through linear regression. The percentage of replaced values relative to the total number of initial planning times in the entire data set was small and was similar for patients (1.99%) and controls (1.76%). Mean planning times were computed for the 2M, 3M, 4M, and 5M problems as well as for the full set of 24 problems comprising the test. In addition, coefficients of variability (CoVs) were computed to reflect the individual variability in each participant’s planning times across the 24 problems. CoVs were obtained by dividing the standard deviation of each participant’s planning times by her/his overall mean planning time on those problems.

**Other measures.** Computerized versions of the Stroop Test and the Symbol Digit Modalities Test (SDMT) were administered to evaluate participants’ information processing speed. These tests are described in greater detail by Hughes, Denney, and Lynch (2011). The Stroop consisted of three 60-s trials yielding measures of participants’ performance at word reading ($W$), color naming ($C$), and naming the color of print for a set of Stroop stimuli ($S$). In each trial, the stimuli appeared individually in the center of the computer screen; the participant gave a verbal response and then pressed the spacebar to display the next stimulus. Along with the score for each trial ($W$, $C$, and $S$), two composite scores were computed: (a) the sum of the word reading and color naming scores ($W + C$), as an overall measure of processing speed on the Stroop; and (b) the relative interference score, found by dividing the difference between $C$ and $S$ by the $C$ score (Denney & Lynch, 2009).

For the SDMT (Smith, 1982), a reference key was present continuously at the top of the computer screen showing nine geometric symbols and the digit corresponding to each symbol. Items consisting of the geometric symbols alone were presented individually in the center of the screen. Participants reported the digit corresponding to each of the geometric symbols and pressed the spacebar to display the next symbol. The computer recorded the number of items completed during the single 90-s trial; the experimenter monitored the number of errors.
**Procedure**

This study was approved by the Human Subjects Committee of the University of Kansas Medical Center, and informed consent was obtained from all participants. Patients were recruited during the course of their regular clinical appointment in the MS Clinic. Controls were initially contacted in person or by phone. For half of the participants, the Stroop and SDMT were administered prior to the TOL and for the other half, these two tests followed the TOL. The full testing appointment took about 30–45 min.

**Results**

*Initial Differences*

Table 1 presents the means and standard deviations for the participants on demographic and disease-related variables. Separate statistical tests were used to compare MS patients versus controls and to compare RRMS patients versus SPMS patients. Patients and controls did not differ significantly in gender, age, or years of education. Among the patients themselves, there were no differences between RRMS and SPMS patients in terms of gender or years of education. However, as expected, patients with SPMS were older ($t = 2.77, df = 44, p = .008$), had longer disease duration ($t = 3.35, df = 44, p = .002$), and higher disability ratings (Mann–Whitney $U$; $z = 4.3, p < .001$) than those with RRMS.

*Identification of the Most Difficult Items on the TOL*

The difficulty of each of the 24 problems comprising the TOL was determined on the basis of the percentage of participants that failed to solve each problem on the first trial. This information is provided in Table 2, along with the specification of the starting position, goal position, and minimum number of moves for each problem. Starting and goal positions are specified using the numbering system for the hexagonal problem space provided by Berg and Byrd (2002).

Ten problems stood out as being particularly difficult, with overall failure rates on the first trial ranging from 24.2% to 62.6%. $\chi^2$ analyses on each of these problems revealed no differences between the number of patients and controls who failed each problem (all $p$-values $>.22$). In addition to computing the total point score, the number of problems solved correctly on the first trial, and the initial planning time for all 24 problems, these three values were also computed for just the 10 most difficult problems of the TOL.

<table>
<thead>
<tr>
<th>Variable</th>
<th>RRMS patients ($N = 29$)</th>
<th>SPMS patients ($N = 17$)</th>
<th>$p$-value$^c$</th>
<th>All MS patients ($N = 46$)</th>
<th>Healthy controls ($N = 45$)</th>
<th>$p$-value$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (F/M)</td>
<td>21/8</td>
<td>13/4</td>
<td>1.00$^c$</td>
<td>34/12</td>
<td>28/17</td>
<td>.27$^c$</td>
</tr>
<tr>
<td>Age</td>
<td>46.7</td>
<td>52.8</td>
<td>.008$^d$</td>
<td>49.0</td>
<td>46.1</td>
<td>.15$^d$</td>
</tr>
<tr>
<td>SD</td>
<td>9.3</td>
<td>7.7</td>
<td></td>
<td>8.6</td>
<td>9.9</td>
<td></td>
</tr>
<tr>
<td>Education (years)</td>
<td>15.7</td>
<td>14.8</td>
<td>.15$^d$</td>
<td>15.4</td>
<td>15.5</td>
<td>.69$^d$</td>
</tr>
<tr>
<td>SD</td>
<td>1.7</td>
<td>2.3</td>
<td></td>
<td>2.0</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Duration of MS (years)</td>
<td>8.4</td>
<td>15.0</td>
<td>.001$^d$</td>
<td>10.8</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>SD</td>
<td>5.3</td>
<td>7.1</td>
<td></td>
<td>6.8</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Disability (Disease steps)</td>
<td>1.7</td>
<td>3.7</td>
<td>&lt;.001$^e$</td>
<td>2.5</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>SD</td>
<td>1.2</td>
<td>1.2</td>
<td></td>
<td>1.5</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Notes: RRMS = relapsing–remitting multiple sclerosis; SPMS = secondary progressive multiple sclerosis.*

$^a$Based on independent comparison of RRMS versus SPMS samples.

$^b$Based on independent comparison of MS patients versus controls.

$^c$Fisher’s exact test.

$^d$Independent sample $t$-test.

$^e$Mann–Whitney $U$-test.
Comparisons Between MS Patients and Controls on the Cognitive Measures

Table 3 presents the means and standard deviations for MS patients and controls on each of the cognitive measures. Independent sample $t$-tests were used to compare the two groups on these measures. Patients and controls did not differ in the number of problems correctly solved on the initial trial or the total number of points scored across all trials of the problems. This was true whether all 24 problems were considered or just the subset of 10 most difficult problems. However, MS patients had significantly longer planning times than controls on the full set of TOL problems ($t = 5.79, df = 89, p < .001, \text{Cohen’s } d = 1.21$) and on the subset of most difficult problems ($t = 4.45, df = 89, p < .001, \text{Cohen’s } d = 0.93$). Patients also exhibited greater individual variability than controls in their planning times across the full set of problems ($\text{CoV}: t = 2.59, df = 89, p = .014, \text{Cohen’s } d = 0.53$).

Mean planning times on the 2M, 3M, 4M, and 5M problems were examined using a 2 (Group) × 4 (Difficulty) mixed factorial ANOVA. The main effects for Group ($F = 33.27, df = 1$ and 89, $p < .001, \eta^2_p = 0.27$) and for Difficulty (Wilks’ $\lambda = 0.30, F = 69.42, df = 3$ and 87, $p < .001, \eta^2_p = 0.71$) were significant. The Group × Difficulty interaction was also significant (Wilks’ $\lambda = 0.88, F = 4.10, df = 3$ and 87, $p = .009, \eta^2_p = 0.12$). Paired comparisons applied to this interaction revealed that although patients exhibited significantly longer planning times than controls at each difficulty level, the disparity between groups increased as the difficulty level became greater. The interaction is illustrated in Fig. 1.

MS patients also completed fewer items on the SDMT ($t = 4.80, df = 78, p < .001, \text{Cohen’s } d = 1.02$) and on each trial of the Stroop (W: $t = 3.74, df = 78, p < .001, \text{Cohen’s } d = 0.84$; C: $t = 3.16, df = 78, p = .002, \text{Cohen’s } d = 0.71$; S: $t = 3.10, df = 78, p = .003, \text{Cohen’s } d = 0.69$; W + C: $t = 3.68, df = 78, p < .001, \text{Cohen’s } d = 0.82$). However, there was no difference between patients and controls in relative interference scores on the Stroop.

Comparisons between RRMS Patients and SPMS Patients on the Cognitive Measures

Table 3 also presents the means and standard deviations for the subgroups of RRMS ($N = 29$) and SPMS ($N = 17$) patients. Because these two groups differed in terms of age, disease duration, and disability, between-group comparisons were first conducted as analyses of covariance, with these three variables entered as covariates. However, neither the covariates nor the
between-group comparisons were significant. When the covariates were omitted and the analyses repeated as independent sample t-tests, the results for the comparison between RRMS and SPMS patients were virtually the same. No differences were found on any of the performance or planning time scores from the TOL or on any of the scores from the Stroop or the SDMT.

**Discussion**

Compared with previous work, the present study featured a seemingly more challenging version of the TOL test to evaluate planning abilities in patients with MS. The current version had more problems and employed four alternative starting positions for these problems. Furthermore, in addition to computing the total points earned when participants were allowed up to three trials to solve each problem, we examined the number of problems solved on just the first attempt. Finally, in addition to evaluating performance on the full set of 24 problems comprising the test, we identified the 10 most difficult problems and computed scores on just that subset of problems. This latter change was particularly useful in augmenting the difficulty level of the TOL.
Whereas the percentage of total possible points earned by all participants on the full set of 24 problems ($M = 89.8\%$) was comparable with that of the 12-item version used in our earlier studies ($M = 88.9\%$; Denney et al., 2011), this percentage was substantially lower when based on the 10 most difficult problems of the present TOL ($M = 78.5\%$).

Despite all these modifications, no differences were found between MS patients and controls on any of the scores pertaining to planning performance. Furthermore, performance was found to be just as effective for patients with SPMS as for those with RRMS. This latter finding contradicts the conclusion by Zakzanis (2000) and others that frontal-executive deficits are more characteristic of patients in the later, progressive stage of MS. The fact that several of our previous TOL studies have used RRMS patients would not seem to account for the failures to find performance differences in conjunction with MS. In all, we are inclined to believe that these failures came about because the fundamental planning skills of at least most MS patients, regardless of subtype, remain virtually intact.

What does differ for these patients is the amount of time they require to formulate their solutions to the TOL problems. As in all previous studies (Arnett et al., 1997; Denney et al., 2004, 2005, 2008, 2011; Foong et al., 1997), we found longer initial planning times for patients compared with controls. This was true for all problems combined, as well as for the subset of the 10 most difficult problems, and separately for 2M, 3M, 4M, and 5M problems.

It is important to note that these differences in planning time occurred in conjunction with differences in the speed of information processing on the Stroop and the SDMT. Scores on the SDMT were related to initial planning times on the TOL. The combined correlation for patients and controls based on Fisher’s $r$-to-$z$ transformation was $-0.390$ ($p < .001$)—although the comparable correlations using scores on the Stroop were not significant. In previous papers (e.g., Denney et al., 2011), we have distinguished between explicit and covert measures of information processing speed and have shown that slower responding by MS patients is evident on both types of measures. It is certainly fair to question whether the differences in planning or decision times occurring on covertly timed tasks such as the TOL should be considered as differences in information processing speed. These measures are perhaps better characterized as the customary “cognitive tempo” participants assume on tasks in which the emphasis is on responding correctly and no apparent importance is attached to the length of time they take to formulate their responses. The difference is analogous to the distinction between sprinting and jogging, with sprinting being more purely a matter of speed and therefore more readily seen as an “ability” and jogging being invested with more personal choice on the part of the runner. Though different, the speed with which one sprints nevertheless bears some relationship with the speed with which one jogs, and furthermore when aging or disease comes to diminish one, it also diminishes the other. Our past research shows that MS has a dramatic impact on both cognitive speed and cognitive tempo (Denney et al., 2011).

The present results also replicated our previous finding that the disparity in planning times between patients and controls becomes greater as the problems increase in difficulty (Denney et al., 2004, 2011). In studies using explicitly timed tasks, interactions of this kind are often termed the “complexity effect” (De Sonneville et al., 2002; Lengenfelder et al., 2006; Parmenter,
Zivadinov, et al., 2007; Reicker, Tombaugh, Walker, & Freedman, 2007) and are attributed to the greater burden on working memory posed by more difficult tasks. Working memory burden is also involved in the present instance, although another factor may simply be the number of moves the participant must rehearse in arriving at his/her plan. Without appealing to working memory burden, the present interaction could be viewed as consistent with the notion that patients’ rehearsal of each move simply occurs more slowly. Thus, as the number of moves required by the problems becomes greater, the difference in planning time between patients and controls also increases.

In addition to lengthier planning times, we found that the planning times of MS patients were characterized by greater individual variability than those of controls. Although we have recently reported greater individual variability in MS patients’ speed of information processing on explicitly timed tasks (Bodling, Denney, & Lynch, 2010), this is the first study showing that their decision times on covertly timed tasks are also characterized by greater variability. Recent studies of dementia (Burton et al., 2006; Dixon et al., 2007; Duchek et al., 2009; Hultsch, MacDonald, Hunter, Levy-Bencheton, & Strauss, 2000), Parkinson’s disease (Burton et al., 2006; de Frias et al., 2007), traumatic brain injury (Collins & Long, 1996; Hetherington, Stuss, & Finlayson, 1996; Stuss, Murphy, Binns, & Alexander, 2003), and epilepsy (Bruhn & Parsons, 1977) indicate that individual variability may be a sensitive marker of central nervous system dysfunction. The neural underpinnings of individual variability are unclear. Researchers have broadly referred to increased “neural noise” (e.g., Dixon et al., 2007), a concept that has also been used to characterize the cognitive changes associated with MS itself (Kail, 1997, 1998). A more specific theory attributes this variability to axon demyelination that results in less stable signal transduction (Bunce et al., 2007). A recent neuroimaging study of healthy older adults showed the degree of white matter lesioning to be correlated with individual variability, but not with average response times or overall performance on a variety of cognitive tests (Bunce et al., 2007). These findings are interesting in light of the association of MS with white matter lesions, demyelination, and axonal degeneration.

Returning to the topic of planning performance, some of the inconsistency between studies concerning deficits in MS patients’ planning performance on various tower tasks may be attributable to these patients’ slower planning times. In studies where explicit limits are imposed on the amount of time to solve each problem, patients tend to exhibit poorer performance than controls. For example, Arnett and colleagues permitted participant’s only 2 min to complete each of the nine problems in their version of the Tower of Hanoi and found that the MS patients in their study solved fewer problems than controls. The impact of this time limit is also evident in the finding that the patients actually executed fewer moves per problem than the controls. In the Discussion section of their paper, Arnett and colleagues expressed concerns over the impact that the time limit may have had on participants’ performance. In a subsequent study investigating planning in samples of depressed and non-depressed MS patients, Arnett, Higginson, and Randolph (2001) had abandoned the Tower of Hanoi in favor of the psychologically sounder TOL and had jettisoned the time limit. This later study, however, did not include a sample of healthy controls.

Undoubtedly the most stringent time limit was used in the fMRI-adapted version of the TOL featured in the study by Lazeron and colleagues (2004). This test consisted of twelve 2M, 3M, and 4M (i.e., “easy”) problems and twelve 6M, 7M, and 8M (i.e., “hard”) problems. Participants were shown the starting display and the goal display, and, instead of making moves themselves, chose one of the two numbers designating the minimum number of moves that would be necessary to transform the starting display into the goal display. They were allowed only 40 s to respond to each problem. Patients responded correctly on fewer of the easy problems than controls; no differences were observed for the hard problems, probably because of floor effects. Interestingly, there was a near significant difference (p = .09) on the control task in this study, in which participants merely had to choose one of two numbers corresponding to the number of disks of a certain color that were on models simulating the starting and goal displays. Given the time limit, patients got fewer correct answers on this simple counting task than the controls. Patients’ poorer performance on both the experimental and control tasks would seem to suggest a deficit in processing speed, rather than one involving planning ability.

In fairness, it must be acknowledged that two other studies (Foong et al., 1997, 1999) report differences in planning performance between MS patients and controls on a TOL procedure that did not include an explicit time limit. The explanation for these contradictory findings must be sought elsewhere; participants in Foong’s studies were aware their decision and solution times were being recorded, and this awareness may have led them to hasten their responses. Still, half of the studies indicating deficits in MS patients’ planning performance used time limits that appear to have posed a particular disadvantage to the patients owing to their slower planning times. The five studies (Denney et al., 2004, 2005, 2008, 2011; Drew et al., 2008) that report no differences in planning performance between MS patients and controls, on the other hand, employ no time limits and avoid making the timing of participants’ responses salient in any way.

DeLuca has proposed the relative consequence model to explain how difficulties exhibited by MS patients in various cognitive domains may result from their fundamental problem with processing speed (DeLuca, Chelune, Tulsky, Lengenfelder, & Chiaravalloti, 2004). What is not addressed in this model is the point at which a “relative consequence” melds into an
“unintended consequence.” This point likely occurs where the problem in patients’ processing speed ceases to be an integral contributor to their performance deficit in the other domain and, instead, the latter performance deficit is a mere artifact of the former. In the present instance, we are inclined to view the deficit in MS patients’ planning performance as an unintended rather than a relative consequence. A study explicitly manipulating the time limit used in conjunction with the TOL would likely offer a test of this hypothesis. The literature bearing on working memory in MS contains several instances where patients’ performance is substantially improved when they are permitted additional time on the items composing the test (Demaree, DeLuca, Gaudino, & Diamond, 1999; Lengenfelder et al., 2006). The most recent of these studies (Leavitt, Lengenfelder, Moore, Chiaravalloti, & DeLuca, 2011) indicates that the improvement in working memory is limited to MS patients who are confronting items that pose the greatest burden upon working memory. The planning literature might well benefit from studies featuring the opposite manipulation, one where more restrictive time limits are imposed and the deleterious impact of these restrictions on patients’ planning performance is evaluated.

Conflict of Interest

None declared.

Acknowledgements

A paper based on this research was presented at the 5th Joint Triennial Congress of the European and Americas Committees for Treatment and Research in Multiple Sclerosis (ECTRIMS/ACTRIMS), Amsterdam, the Netherlands, October 21, 2011.

References


