The Cognitive Performance of Patients with Multiple Sclerosis during Periods of High and Low Fatigue

Brett A. Parmenter, M.A.[1]

Douglas R. Denney, Ph.D.*[1]

Sharon G. Lynch, M.D.[2]

[1]Department of Psychology

University of Kansas

Lawrence, Kansas 66045, USA

[2]Department of Neurology

University of Kansas Medical Center

Kansas City, Kansas 66160, USA

Running title: Fatigue and Cognitive Performance in MS

Keywords: multiple sclerosis, cognitive perforfance, fatigue

---

* Correspondence: DR Denney, Department of Psychology, 1415 Jayhawk Blvd., University of Kansas, Lawrence, Kansas, 66045, USA. Phone: 785-864-9805 Fax: 785-864-5696 e-mail: denney@ku.edu

Abstract

The objective of this study was to examine whether MS-related fatigue affects patients' cognitive performance. Thirty patients who had substantial fatigue in conjunction with MS and who reported marked diurnal variability in the severity of their fatigue were tested on two occasions: during a period of high fatigue and during a period of relatively low fatigue. The order of these sessions was counterbalanced across patients. During both sessions, patients completed a questionnaire concerning their present state of fatigue and a battery of neuropsychological tests of planning, selective attention, and paired associate learning. Although the patients confirmed greater fatigue during the period of high fatigue and felt they had performed more poorly during this period, there were no differences in cognitive performance that could be attributed to fatigue. Instead, all subjects showed improvement from the first to the second session regardless of whether the latter entailed a period of high or low fatigue. In contrast to studies reporting fatigue-related declines in MS patients' cognitive performance, we found no differences in performance when MS patients were tested during periods of high versus low fatigue. These contrasting results, stemming from differences in experimental design, are discussed in terms of their implications for assessing cognitive function in patients with MS.

Introduction

Surveys have shown that about 70% of patients with multiple sclerosis (MS) experience fatigue as a daily symptom.[1] Almost one third of these patients report fatigue to be their most disabling symptom,[2] and over half view fatigue as one of their worst symptoms.[3] Although less prevalent than fatigue, cognitive problems are also frequently encountered among patients with MS, the most common deficits involving recent memory, attention, verbal fluency, conceptual reasoning, and visual-spatial perception.[4] Rao found cognitive deficits in 43% of a general sample of MS patients, and higher percentages have been reported when patients are recruited from university-based medical centers.[5] Given the prevalence of both fatigue and cognitive impairment, it is not surprising to encounter many patients with both types of symptoms. Such individuals typically report that their cognitive functioning is adversely affected by their level of fatigue,[2,6] and these claims provided a major impetus to the present study.

In spite of such claims, simple correlational studies have failed to reveal associations between fatigue and cognitive performance,[3,7] and experimental studies in which patients' fatigue is manipulated have produced conflicting results. One approach adopted by these latter studies has been to compare patients' cognitive performance before and after pharmacological treatment of their fatigue. Improvements in attention and concentration following treatment with amantadine have been reported in two studies,[7,8] although in each study, the improvement was confined to a single measure within a much larger battery of tests. Furthermore, in only one of these studies did patents' ratings of fatigue improve relative to the placebo condition.[8] In the other study,[7] equal improvement in subjective ratings of fatigue occurred for patients on amantadine, pemoline, and placebo, and therefore the

enhancement in attention shown by the amantadine group is difficult to interpret.

Another approach has been to examine patients' cognitive performance over the course of an experimental induction of fatigue. Usually the induction procedure involves working on cognitive tasks; however, in one study [9] changes in patients' performance were examined following physical exertion. Although metabolic parameters were altered and patients reported increases in fatigue resulting from the sessions of physical exercise, no changes occurred in measures of verbal fluency, visual tracking, memory, or attention and concentration.

Johnson and her associates [10] used a 3-hour battery of neuropsychological tests to induce fatigue in patients with MS, chronic fatigue syndrome, or clinical depression and in healthy controls. Four times during the course of this testing session, subjects rated their fatigue and completed the Paced Auditory Serial Addition Test (PASAT). Subjective fatigue increased over the course of testing, and there was a particularly noteworthy increase in MS patients' fatigue ratings between the second and third assessments. Nevertheless, the performance of all four groups improved over the course of repeated testing, and this improvement was equal across the groups.

In another study, [11] samples of MS patients and healthy controls were tested for grip strength, serial recall of word lists, and concentration on a vigilance task before and after 30 minutes of work on a cognitive battery. Over the course of working on this battery, the MS patients reported increases in both cognitive and physical fatigue, whereas the controls reported no changes in either form of fatigue. Relative to controls, the patients performed more poorly on all three objective measures. However, these differences were evident both before and after the cognitive battery, and no changes occurred in the performance of either group over the course of the

battery. Similarly, another group of investigators [12] found that patients had longer reaction times than did healthy controls, but these differences were also evident both before and after the administration of a 4-hour cognitive battery. Changes in reaction times occurring over the course of the battery were the same for both patients and controls.

The two studies [13,14] that have demonstrated declines in patients' cognitive performance relative to controls have both required subjects to work on a continuous cognitive task throughout the induction period instead of a battery of different tasks. In one of these studies, [13] the performance of MS patients with and without mild cognitive impairment and that of healthy controls was examined during three 5-minute blocks of a vigilance task. Throughout the task, patients with cognitive impairment performed more slowly than unimpaired patients or controls. Relative to controls, the rate of performance for the MS patients declined from the first to the second and third blocks of the task. Unfortunately no subjective ratings of fatigue were collected in this study, and thus it is difficult to attribute the declines in patients' performance to fatigue.

Krupp and Elkins [14] also used a continuous performance task in their fatigue induction procedure, one that lasted about three hours. Subjects rated their mental and physical fatigue three times during the course of this procedure and completed a battery of cognitive tests before and after the procedure. Fatigue increased for both patients and controls during the procedure. Initially the two groups did not differ on any of the cognitive measures, but at posttest, the patients' performance had declined on tests of verbal and spatial memory, verbal fluency, and conceptual planning, while the performance of the controls had improved on these measures.

Beyond the fact that they have yielded conflicting results, experimental studies such as these may not offer the best venue in which to explore the relationship

between fatigue and cognitive functioning. As some have suggested, this relationship may be more complex in the naturalistic setting than can be reproduced through experimental manipulation of patients' fatigue.[9] Most patients experience pronounced diurnal fluctuations in their MS-related fatigue, with this symptom typically being worse in the late afternoon or early evening than in the morning.[1] Despite such fluctuations, only two experimental studies [12,14] were careful to test all subjects over the same time period (e.g., between 11 am and 3 pm). To our knowledge, no studies have explicitly examined cognitive performance during times when patients are naturally experiencing their greatest and least amounts of daily fatigue. Such a study might permit a more naturalistic assessment of patients' claims concerning greater cognitive impairment during periods of fatigue and may indicate the advisability of scheduling neuropsychological assessments around patients' daily bouts of fatigue.

The selection of cognitive measures for the present study was influenced not only by the previously reviewed experimental studies, but also by the findings of a study examining metabolic activity in the brains of MS patients with and without fatigue. Using positron emission tomography (PET), Roelcke found lower metabolic activity in the frontal lobes and higher activity in the anterior cingulate for patients with fatigue.[15] Because the frontal lobes are associated with executive functions [16] and the anterior cingulate with selective attention, [17] we were prompted to include measures of these cognitive functions in the present battery. We also included a measure of verbal memory, based on Krupp and Elkins' [14] finding of differences on such measures before and after the induction of fatigue.

Method

Subjects

This study was approved by the Human Subjects Committee of the University of Kansas Medical Center. Patients with clinically definite MS based on Poser's [18] criteria were recruited for this study during the course of their regular appointment at the medical center's MS Clinic. All subjects had been patients of the same neurologist (Sharon G. Lynch) for at least a year. To be eligible for the study, patients had to confirm that fatigue was a prominent symptom of their illness and that their fatigue varied predictably and to a notable degree over the course of each day. Dr. Lynch evaluated the extent of the patient's disability using the Expanded Disability Status Scale (EDSS) [19] and recorded the duration of the patient's disease.

A total of 41 individuals were asked to participate; 9 declined the invitation. Two patients were later excluded: one asked to end the experiment early and the other was found to be too cognitively impaired to complete the measures. The 30 subjects (24 females, 6 males) who completed the study ranged in age from 23 to 74 (M = 44.7). Their length of illness ranged from 1 to 37 years (M = 11.7), and their EDSS scores ranged from 1.5 to 7 (M = 4.6). All but one patient were taking MS-related medications, and medications were unchanged over the course of the study. Eight subjects reported they were experiencing an exacerbation in their symptoms at the time of the study; because their exacerbating symptoms were unchanged over the course of the study, these patients were retained.

Measures

In addition to the cognitive measures described below, subjects' general level of fatigue during the preceding week was assessed using the Fatigue Severity Scale

(FSS) [20] and the Fatigue Impact Scale (FIS), [21] and their general level of depression during the past week was assessed using the Center for Epidemiologic Studies - Depression Scale (CES-D). [22] Also, immediately before they began working on the cognitive battery, subjects rated their present state of fatigue using an abbreviated scale derived from the Profile of Mood States (POMS) [23] and consisting of the seven items composing the fatigue scale and eight filler items selected from other scales on the POMS. Subjects rated each item according to how they were feeling "right now," using a scale from 0 (not at all) to 4 (extremely). The ratings assigned to the fatigue items were summed to yield a score reflecting the subject's present state of fatigue.

The Tower of London (TOL). [24] The TOL is a test of planning that has been shown to relate to prefrontal functioning. [25] The test was computerized for the present study. In the upper portion of the screen, the computer displayed three colored disks on three pegs. In the bottom portion of the screen, the computer displayed a model with the disks in a different arrangement on the pegs. The subject's task was to move the disks in the upper display so they matched the arrangement in the bottom display and to do so in a specified number of moves. Twelve such problems were presented, graduated in difficulty from those requiring two moves to those requiring five moves to solve. Subjects were awarded points for solving each problem on their first (3 points), second (2 points), or third (1 point) attempt. These points were summed across the 12 problems to yield the TOL point score. The computer also measured the length of time between the initial presentation of each problem and the subject's first move. "Initial planning times" were averaged separately for 2-, 3-, 4-, and 5-move problems.

The Wisconsin Card Sorting Test (WCST). [26,27] The WCST is a concept formation test that has been shown to be sensitive to impairment involving the frontal lobes. [28] The WCST consists of 128 cards bearing symbols which can be sorted

according to form (stars, circles, squares, or cross), color (red, green, yellow, or blue), or number (one to four symbols on the card). In the computerized version of the test used in this study, four reference cards were shown at the top of the screen and individual cards were presented at the bottom of the screen for the subject to match to the reference cards. The computer provided feedback as to whether the placement of a card was "right" or "wrong." After ten consecutive correct responses, the operable concept changed. The sequence of correct concepts used in the test was color-form-number-color-form-number. The computer tallied the number of cards sorted before all six concepts were discovered or the number of concepts completed before all 128 cards were used, as well as the total number of errors, and the percentage of perseverative errors committed during the test.

The Stroop Color-Word Interference Test (Stroop). [29,30] The Stroop is a test of speeded information processing and attention that has been shown to relate to activation of the anterior cingulate gyrus. [17] The computerized version of the Stroop used in the present study, consisted of three 1-minute trials during which the subject first read color words (word reading), then named the color of a row of four X's (color naming), and finally, named the color of the letters of color words (color-word naming). The colors and words used in this task were red, blue, green, and yellow. In the color-word naming trial, all the stimuli were incongruent (e.g., the word "GREEN" printed in blue letters). The subject gave a verbal response to the stimulus (i.e., read the word or named the color), and the experimenter pressed the space bar to display the next stimulus. The computer timed the trial and recorded the number of stimuli completed during the trial. In addition to the word reading, color naming, and color-word naming scores, an interference score was derived by subtracting a "predicted" color-word naming score from the actual score. [31]

The Paired Associates Learning Test (PALT). The PALT was designed for the present study to assess both immediate recall and delayed recognition memory for verbal stimuli. The test began with the computer displaying a randomized sequence of eight related (e.g., "APPLE - WORM") and eight unrelated (e.g., "GLASSES - BUS") word pairs. Subjects were then shown one word from each pair and asked to recall the word with which it was paired. This procedure was repeated a second time. The computer recorded the number of related and unrelated words correctly recalled during this immediate recall phase. After the second recall trial, subjects were administered the Stroop in order to impose a delay prior to the recognition phase of the PALT. Subjects were then shown a randomized sequence of 32 words from the original word pairs and 32 additional words (i.e., "foils") and were asked to indicate whether each word was or was not included in the original word pairs. The computer recorded the number of related words, unrelated words, and foils correctly recognized during this delayed recognition phase.

Procedure

While the patient was in the clinic, the investigator explained the study and obtained informed consent. She then questioned the patient concerning the time of day when fatigue was customarily at its lowest and highest level. Two appointments were scheduled separated by at least a week (M = 11 days) -- both of which were conducted in the patient's home. These appointments occurred at the times the patient had designated as periods of greatest and least fatigue. The order of these appointments was randomized and counterbalanced across patients, with half the patients going from greatest fatigue to least fatigue (i.e., "high-low"), and half from least fatigue to greatest fatigue (i.e., "low-high").

During the initial testing session, subjects completed the two fatigue questionnaires (FSS and FIS) and the depression questionnaire (CES-D). They next recorded their present level of fatigue using the abbreviated version of the POMS. Subjects were then administered the battery of cognitive measures (i.e., the TOL, PALT, Stroop, and WCST. This session lasted between 1.5 and 2 hours.

During the second testing session, subjects again rated their present level of fatigue using the POMS and was readministered the TOL, PALT, and Stroop. The 12 problems composing the TOL remained the same between the two sessions as did the Stroop procedure. However, two equivalent forms of the PALT were designed for this study consisting of separate sets of word pairs. The order in which the sets were assigned to the two sessions was randomized and counterbalanced across subjects. The WCST was not readministered during the second session. Through their initial encounter with this test, subjects had an opportunity to learn the various bases upon which the cards were sorted (i.e., color, form, number) and to discover that the operable concept changed over the course of the procedure. Such learning nullified the readministration of this test.

Results

Typically, patients with MS experience greater fatigue as the day progresses. For 23 of the present patients, the time of greatest fatigue occurred later in the day than the time of least fatigue. However, seven patients reported the time of greatest fatigue to be earlier in the day than the time of least fatigue. Therefore we began by comparing these two subgroups of patients using independent sample t tests. The two subgroups did not differ on any demographic or disease related variables, on any of the self report measures of fatigue or depression, or on any of the cognitive measures.

Accordingly the patients were combined into a single sample for all subsequent analyses.

Since the initial testing occurred during a time of low fatigue for half the subjects (i.e., the low-high group) and a time of high fatigue for the other half (i.e., the high-low group), it was possible to compare the initial test data for these two groups. As seen in Table 1, the subjects initially tested during a time of high fatigue rated their present state of fatigue significantly higher than those tested during a time of low fatigue. This difference is to be expected since the POMS was used as a "state measure" of fatigue. On the other hand, the two groups did not differ in terms of the severity or impact of their fatigue during the past week, as measured by the FSS and the FIS. Nor did they differ in their level of depression on the CES-D or on any of the demographic or disease-related variables. Most notably, the two groups did not differ on any of the cognitive measures. During the initial testing session, the performance of the subjects tested at a time of high fatigue was not reliably different from that of subjects tested at a time of low fatigue.

This "between-group" comparison of subjects whose initial testing session occurred during a time of high fatigue versus a time of low fatigue was the only analysis that could be applied to scores derived from the WCST because this test was only administered during the initial session. However, the other cognitive measures could be subjected to a "within-group analysis" with greater statistical power. This involved a 2 (Order) x 2 (Session) mixed factorial analysis of variance applied to the repeated measures collected during the first and the second testing sessions. In this type of analysis, it was possible statistically to segregate "practice effects" from "fatigue effects" by focusing on the main effect for Session and the Order x Session interaction. A significant main effect for Session might occur because subjects performed better

during the second session than during the first session regardless of whether the second session occurred during a time of high fatigue or low fatigue. In this case, the outcome would reflect "practice effects." On the other hand, a significant interaction might occur because subjects in the low-high group performed better during the first session whereas those in the high-low group performed better during the second session. Such an outcome would reflect the deleterious impact of fatigue on cognitive performance. Table 2 presents the data from the two testing sessions that were subjected to this within-groups analysis.

At the start of each testing session, subjects rated their present state of fatigue using the abbreviated version of the POMS. An analysis of these ratings revealed a significant Order x Session interaction ($F = 42.1$; $df = 1, 24$; $p < .001$). As expected, subjects in the high-low group rated their fatigue higher in the first session ($M = 15.7$) than in the second session ($M = 8.4$; $t = 4.0$; $df = 11$; $p = .002$). Those in the low-high group rated their fatigue higher in the second session ($M = 15.4$) than in the first session ($M = 7.8$; $t = 5.2$; $df = 13$; $p < .001$).

The 2 (Order) x 2 (Session) analysis of point scores on the TOL revealed a significant main effect for Session ($F = 8.3$; $df = 1, 28$; $p = .007$). This occurred because subjects achieved higher scores on the second session ($M = 33.0$) than on the first ($M = 31.2$). The Order x Session interaction was not significant ($F = 1.4$; $df = 1, 28$; $p = .25$).

A 2 (Order) x 2 (Session) x 4 (Problem Type) analysis of variance was performed on the initial planning times for 2-, 3-, 4-, and 5-move problems on the TOL. This analysis resulted in a significant main effect for Problem Type ($F = 22.8$; $df = 3, 26$;

$p < .001$), owing to the fact that planning times lengthened as the problems required larger numbers of moves to solve. There was also a significant main effect for Session ($F = 15.0$; $df = 1, 28$; $p = .001$) because initial planning times were longer during the first session ($M = 23.3$) than during the second ($M = 18.9$). The Order x Session interaction was not significant ($F = 1.8$; $df = 1, 28$; $p = .19$).

Separate 2 (Order) x 2 (Session) analyses were performed on each of the scores derived from the Stroop. The only significant outcomes in these analyses involved the main effect for Session in the case of color-word naming scores ($F = 6.8$; $df = 1, 28$; $p = .014$) and interference scores ($F = 9.7$; $df = 1, 28$; $p = .004$). Consistent with practice effects, subjects achieved higher color-word naming scores ($M = 41.4$ vs 39.2) and evidenced greater resistance to interference ($M = 11.1$ vs 8.8) during the second testing session compared to the first.

A 2 (Order) x 2 (Session) x 2 (Word Pair) analysis of variance was performed on the recall scores for related and unrelated words on the PALT. The only significant finding was a main effect for Word Pair ($F = 77.8$; $df = 1, 23$; $p < .001$), resulting from the fact that subjects correctly recalled more of the words occurring in related pairs ($M = 14.0$) than in unrelated pairs ($M = 8.7$). Neither the main effect for Session ($F = 1.2$; $df = 1, 23$; $p = .29$) nor the Order x Session interaction ($F = 0.4$; $df = 1, 23$; $p = .54$) was significant, showing that recall scores were not subject to either practice effects or fatigue effects. Similarly, a 2 (Order) x 2 (Session) x 2 (Word Pair) analysis of variance was performed on the delayed recognition scores for related and unrelated words on the PALT. Again, the only significant finding was a main effect for Word Pair ($F = 15.2$; $df = 1, 26$; $p = .001$), with subjects recognizing more of the words occurring in related pairs ($M = 15.3$) than in unrelated pairs ($M = 14.2$). The main effect for Session approached significance ($F = 3.7$; $df = 1, 26$; $p = .07$). However, this outcome was not

due to practice effects; the mean was actually slightly lower for the second session (M = 14.5) compared to the first (M = 14.9). As in the case of the recall scores, the Order x Session interaction (F = 2.4; df = 1, 26; p = .14) for recognition scores was not significant.

In order to directly compare the various cognitive measures in terms of their sensitivity to possible fatigue effects, patients' scores during the period of high fatigue and during the period of low fatigue were compared using paired sample t tests. None of these t tests performed on the cognitive measures was significant, again indicating no significant fatigue effects in the present study. Nevertheless the t values computed on the various measures can still be compared as indices of the sensitivity of these measures to possible fatigue effects. These t values are presented in Figure 1. Because lower scores for initial planning time on the TOL and for interference on the Stroop signified better performance, the signs for the t values computed on these measures were reversed. The effect of fatigue on any particular cognitive measure is reflected by the extent to which the resultant t values are positive; a negative t value indicates that performance on the measure was actually better during the period of high fatigue. Figure 1 also presents the effect size (i.e., omega square; ω^2) for fatigue for each of these measures.

As illustrated in Figure 1, the cognitive measures that were most sensitive to fatigue were the word reading score (t = 1.79, df = 29, p = .08) and color naming score (t = 1.90, df = 28, p = .07) from the Stroop. By contrast, the measures that seemed to be least affected by fatigue were the interference score from the Stroop and the recall score from the PALT (both t's < 1), both of which improved slightly during the period of high fatigue.

Discussion

As noted in the introduction, patients with MS often feel that their cognitive functioning is adversely affected by the fatigue that accompanies their disease. Although we did not collect systematic data on these impressions, many of the patients in the present study spontaneously commented that they had done more poorly when they were tested during the period of high fatigue than during the period of low fatigue. On the other hand, the performance data did not support these impressions. Despite the fact that patients rated their fatigue significantly higher during the period of high fatigue than during the period of low fatigue, there were no significant fatigue effects on any of the cognitive measures employed in this study. Similar discrepancies between subjective and objective performance have been noted for patients with chronic fatigue syndrome. [31,32,33]

Previous studies that purportedly demonstrate adverse effects of fatigue on the cognitive performance of patients with MS are plagued by numerous shortcomings. The two studies [7,8] that report an improvement in patients' performance following treatment with anti-fatigue drugs each focus on a single measure isolated from a much larger battery of cognitive tests -- with no statistical adjustment for the large number of dependent measures examined. The single measure is one involving attention, but the investigators fail to consider that drugs such as amantadine or pemoline may directly improve attention without impacting fatigue. Only one of the drug studies [9] shows an actual improvement in fatigue as well as an improvement in attention relative to a placebo condition. In neither study is there a demonstration that changes in patients' cognitive performance are actually correlated with changes in fatigue.

Only two of the six fatigue induction studies have reported evidence of adverse effects on patients' cognitive performance, and one of these two studies [13] failed to

examine whether patients' subjective levels of fatigue were actually increased over the course of the induction. The six induction studies vary a great deal with respect to the length of the fatigue induction period. However, a failure to discover fatigue effects can not be attributed to the use of too brief an induction. Curiously, the studies employing the shortest [13] and longest induction periods [14] are the two reporting an adverse effect on performance.

Krupp and Elkins [14] argued that a crucial element is that patients must be involved in a continuous testing procedure throughout the length of the induction period instead of being allowed to work on a battery of different tasks. However, they offer no explanation for why this feature would make a difference in the outcome of the study. Presumably as long as the induction causes a documented increase in fatigue, the effects on cognitive performance should either be present or absent.

The failure to find significant fatigue effects might be due to insufficient statistical power; however, the power of the present study would seem to be at least comparable to that of previous studies. A direct comparison is complicated by the fact that the present study employed largely a "within-group" design. This design has inherently greater statistical power than the "between-group" design used in most of the other studies. With subjects serving as there own controls, the effective sample size in the present study is 60. Only one other study [13] employed a larger sample size.

With uniformly nonsignificant fatigue effects found in the present study, the opportunity to compare the sensitivity of different types of cognitive measures to fatigue was of course limited. To the extent that such comparisons were possible, we found little evidence to support Roelcke's [15] suggestion that "frontal lobe" measures of executive functions such as planning would be more adversely affected by fatigue and that those involving the sort of selective deployment of attention that has been linked to

activity in the anterior cingulate gyrus would be relatively spared. A principal measure of executive function used in this study was the Tower of London, and the point scores and initial planning times from this measure revealed only a modest impact of fatigue. The two measures that showed the greatest fatigue effects involved the word reading and color naming trials of the Stroop, both measures relating to attention and speeded information processing.

It is important to consider the difference between what might be termed "trait fatigue" and "state fatigue." The present study as well as the treatment and induction studies reviewed earlier are primarily investigations of patients' current state of fatigue and its impact on cognitive performance. Roelcke's suggestion stemmed from a study of "trait fatigue" in which patients who experienced a great deal of fatigue as a symptom of MS were compared with patients for whom fatigue was not a prominent symptom of their disease. It remains a possibility that groups differing in trait fatigue might exhibit differential performance on cognitive measures and that measures of executive functions might be especially sensitive to trait fatigue. We think this possibility is unlikely, however. As part of our ongoing program of research, we have established "local norms" for several of the measures used in the present study. When the present sample was compared to a normative sample of 100 MS patients selected from the same clinic, the only significant difference was that the 30 patients in the present sample reported significantly higher levels of trait fatigue on the FSS ($t = 3.08$, $df = 128$, $p = .003$). This difference is not surprising in light of the fact that subjects were selected for the present study on the basis of their having a substantial problem with fatigue. However, despite their experiencing more severe fatigue as a symptom of their illness, these subjects did not differ from the normative sample in their

performance on the any of cognitive measures for which normative data were available (i.e., the TOL and the Stroop).

Unlike the treatment or induction studies, the present study is not a true experiment wherein patients' levels of fatigue were actively manipulated. We would argue that what the study lacks in terms of experimental manipulation of the fatigue variable is compensated by its ecological validity. Most patients with MS experience substantial diurnal variation in their fatigue, and when we took advantage of this fact to assess patients' cognitive performance during periods of high and low fatigue, we found very little difference. From a practical standpoint, these findings call into question the necessity of scheduling neuropsychological assessments so as to avoid periods of high fatigue during the course of the patient's day.

References

1. Freal JE., Kraft GH, Coryell JK. (1984) Symptomatic fatigue in multiple sclerosis. Arch Phys Med Rehabil **65**: 135-138.
2. Krupp LB, Alvarez LA, LaRocca NG, Scheinberg LC. (1988) Fatigue in multiple sclerosis. Arch Neurol **45**: 435-437.
3. Fisk JD et al. (1994) The impact of fatigue on patients with multiple sclerosis. Can J Neurol Sci **21**: 9-14.
4. Rao SM., Leo GJ, Bernardin LB, Unverzagt F. (1991) Cognitive dysfunction in MS. I. Frequency, patterns, and predictors. Neurology **41**: 685-691.
5. Miller AE. (1996) Clinical features. In SD Cook (Ed.), Handbook of multiple sclerosis: second edition. New York: Marcel Dekker.
6. Schwartz CE, Coulthard-Morris L, Zeng Q. (1996) Psychosocial correlates of fatigue in multiple sclerosis. Arch Phys Med Rehabil **77**: 165-170.
7. Geisler MW et al. (1996) The effects of amantadine and pemoline on cognitive functioning in multiple sclerosis. Arch Neurol **53**: 183-188.
8. Cohen RA, Fisher M. (1989) Amantadine treatment of fatigue associated with multiple sclerosis. Arch Neurol **46**: 676-680.
9. Caruso LS et al. (1991) Exertional fatigue fails to affect cognitive function in multiple sclerosis. J Clin Exp Neuropsychol **13**: 74. (Abstract)
10. Johnson SK et al. (1997) The effects of fatigue on neuropsychological performance in patients with chronic fatigue syndrome, multiple sclerosis, and depression. Appl Neuropsychol **4**: 145-153.
11. Paul RH et al. (1998) Cognitive and physical fatigue in multiple sclerosis: relations between self-report and objective performance. Appl Neuropsychol **5**: 143-148.

12. Jennekins-Schinkel A, Sanders EACM, Lanser JBK, Van der Velde EA. (1988) Reaction time in ambulant multiple sclerosis. J Neurol Sci **85**: 173-180.

13. Kujala P, Portin R, Revonsuo A, Ruutianen J. (1995) Attention related performance in two cognitively different subgroups of patients with multiple sclerosis. J Neurol Neurosurg Psychiatry **59**: 77-82

14. Krupp LB, Elkins LE. (2000) Fatigue and declines in cognitive functioning in multiple sclerosis. Neurology **55**: 934-939.

15. Roelcke U et al. (1997) Reduced glucose metabolism in the frontal cortex and basal ganglia of multiple sclerosis patients with fatigue: a $_{18}$F-fluorodeoxyglucose positron emission tomography study. Neurology **48**: 1566-1571.

16. Huber SJ et al. (1992) Magnetic resonance imaging correlates of executive function impariments in multiple sclerosis. Neuropsychiatry Neuropsychol Behav Neurol **5**: 33-36.

17. Banich MT et al. (2000) Prefrontal regions play a predominant role in imposing an attentional 'set': evidence from fMRI. Cogn Brain Res **10**: 1-9.

18. Poser CM et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol **13**: 227-231.

19. Kurtzke JF. (19983) Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology **33**: 1444-1452.

20. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. (1989) The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol **46**: 1121-1123.

21. Fisk JD et al. (1994) Measuring the functional impact of fatigue: initial validation of the Fatigue Impact Scale. Clin Infect Dis **18** (Suppl 1): S79-S83.

22. Radloff L. (1977) The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Measure **1**: 385-401.

23. McNair DM, Lorr M, Droppleman LF. (1971) EITS Manual for the Profile of Mood States. San Diego, CA: Educational and Industrial Testing Service.

24. Krikorian R, Bartok J, Gay N. (1994) Tower of London procedure: a standard method and developmental data. J Clin Exp Neuropsychol **16**: 840-850.

25. Shallice T. (1982) Specific impairments of planning. Philos Trans R Soc London B. **298**: 199-209.

26. Berg EA. (1948) A simple objective test for measuring flexibility in thinking. J Gen Psychol **39**: 15-22.

27. Heaton RK. (1981) Wisconsin Card Sorting Test manual. Odessa, FL: Psychological Assessment Resources Inc.

28. Milner B. (1963) Effects of different brain lesions on card sorting: the role of the frontal lobes. Arch Neurol **9**: 90-100.

29. Stroop JR. (1935) Studies of interference in serial verbal reactions. J Exper Psychol **18**: 643-662.

30. Golden CJ. (1994) The Stroop Color and Word Test: a manual for clinical and experimental uses. Wood Dale, IL.: Stoelting Company.

31. Metzger FA, Denney DR. Perception of cognitive performance in patients with chronic fatigue syndrome. Ann Behav Med. (In press)

32. Ray C, Phillips L, Weir WR. (1993) Quality of attention in chronic fatigue syndrome: subjective reports of everyday attention and cognitive difficulty, and performance on tasks of focused attention. Br J Clin Psychol **32**: 357-364.

33. Wearden AJ, Appleby L. (1997) Cognitive performance and complaints of cognitive impairment in chronic fatigue syndrome (CFS). Psychol Med. **27**: 81-90.

Table 1

Comparison of Initial Performance for Patients Tested during a Period of High Fatigue (N = 15) and Patients Tested during a Period of Low Fatigue (N = 15)

		High Fatigue	Low Fatigue	t	df	p
Age	M	43.00	46.57	0.92	27	.365
	S.D.	13.08	6.48			
Length of illness	M	11.73	11.64	0.31	27	.976
	S.D.	9.32	5.96			
EDSS: Disability	M	4.68	4.60	0.10	24	.918
	S.D.	2.18	1.81			
CESD: Depression	M	17.13	20.40	0.74	28	.468
	S.D.	10.18	13.84			
FSS: Fatigue	M	5.75	6.17	1.26	27	.218
	S.D.	1.03	0.77			
FIS: Fatigue	M	86.80	82.13	0.36	28	.721
	S.D.	33.06	37.58			
POMS: Fatigue	M	15.67	7.79	3.77	24	.001
	S.D.	5.19	5.42			
TOL: Point Score	M	30.93	31.47	0.43	28	.674
	S.D.	3.15	3.70			
TOL: Initial Planning Time (secs)	M	27.75	23.98	1.29	28	.208
	S.D.	8.03	8.01			

Table 1 (continued)

		High Fatigue	Low Fatigue	t	df	p
Stroop: Word Reading	M	66.00	69.00	0.67	28	.511
	S.D.	15.00	8.93			
Stroop: Color Naming	M	55.40	55.80	0.12	28	.903
	S.D.	9.76	7.93			
Stroop: Color-Word Naming	M	38.27	40.20	0.63	28	.536
	S.D.	9.10	7.74			
Stroop: Interference Score	M	8.25	9.40	0.60	28	.555
	S.D.	5.32	5.24			
PALT: Recall - Related Words	M	14.46	13.80	0.92	26	.366
	S.D.	1.81	1.97			
PALT: Recall - Unrelated Words	M	8.92	7.67	0.83	26	.414
	S.D.	3.73	4.20			
PALT: Recognition - Related Words	M	15.54	15.67	0.67	26	.507
	S.D.	0.52	0.49			
PALT: Recognition - Unrelated Words	M	14.00	14.40	0.53	26	.598
	S.D.	2.55	1.30			
WCST: Concepts	M	5.73	5.08	1.34	22	.196
	S.D.	0.65	1.50			
WCST: Trials to Solution	M	97.00	100.77	0.41	23	.687
	S.D.	21.64	24.38			
WCST: Total Errors	M	19.55	24.15	0.87	22	.393
	S.D.	9.27	15.31			
WCST: Perseverative Errors (%)	M	0.08	0.13	1.68	22	.107
	S.D.	0.05	0.08			

Table 2

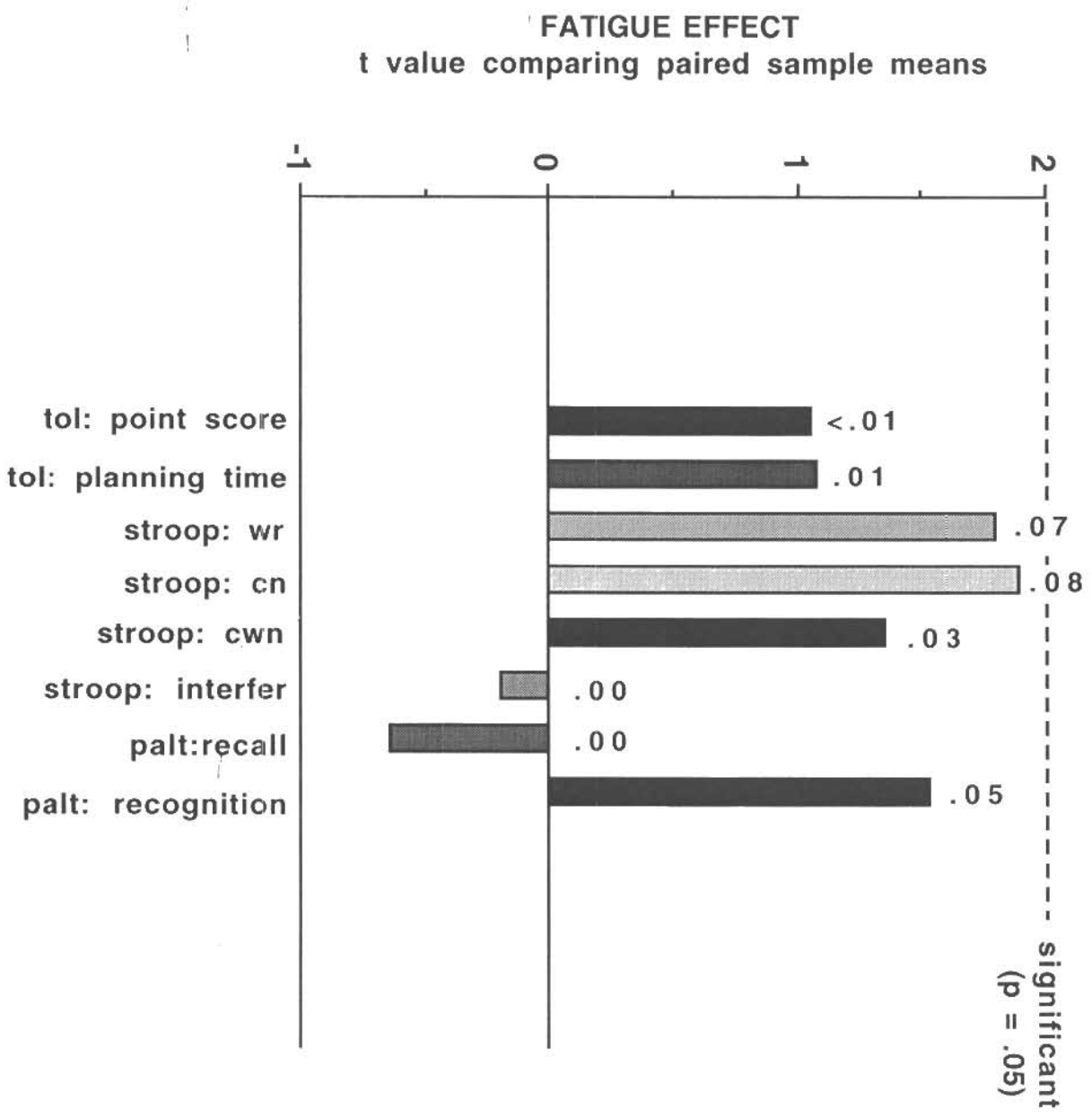
Performance during the First and Second Sessions for Patients Tested in the High-Low Order (N = 15) and Patients Tested in the Low-High Order (N = 15)

Order		High-Low		Low-High	
Session		First	Second	First	Second
POMS: Fatigue	M	15.67	8.42	7.79	15.43
	S.D.	5.19	7.38	5.42	6.00
TOL: Point Score	M	30.93	33.47	31.47	32.53
	S.D.	3.15	2.07	3.70	2.90
TOL: Initial Planning	M	27.75	21.22	23.98	20.77
Time (secs)	S.D.	8.03	9.73	8.01	6.59
Stroop: Word	M	66.00	69.53	69.00	66.80
Reading	S.D.	15.00	17.91	8.93	10.25
Stroop: Color	M	55.40	56.27	55.80	53.60
Naming	S.D.	9.76	11.65	7.93	7.53
Stroop: Color-Word	M	38.27	41.67	40.20	41.13
Naming	S.D.	9.10	9.74	7.74	6.15
Stroop: Interference	M	8.25	10.67	9.40	11.50
Score	S.D.	5.32	5.93	5.24	4.23
PALT: Recall -	M	14.46	13.77	13.80	12.80
Related Words	S.D.	1.81	2.98	1.97	1.78
PALT: Recall -	M	9.91	8.91	7.64	8.14
Unrelated Words	S.D.	2.91	4.90	4.36	4.49
PALT: Recognition -	M	15.54	15.15	15.67	14.67
Related Words	S.D.	0.52	0.69	0.49	0.98
PALT: Recognition -	M	14.00	14.23	14.40	14.00
Unrelated Words	S.D.	2.55	1.83	1.30	1.60

Figure Caption

Figure 1. Fatigue effects on various cognitive measures. The t values based on comparisons of paired sample means for scores during the period of high fatigue and during the period of low fatigue are graphed. The effect sizes (omega square) for fatigue are printed above each plot.

Note: tol = Tower of London; wr = word reading score; cn = color naming score; cwn = color-word naming score; interfer = interference score; palt = Paired Associates Learning Test.


FATIGUE EFFECT
t value comparing paired sample means
-1
0
1
2
tol: point score
tol: planning time
stroop: wr
stroop: cn
stroop: cwn
stroop: interfer
palt:recall
palt: recognition
<.01
.01
.07
.08
.03
.00
.00
.05
significant
(p = .05)