Abnormal Neuropsychological Findings Are Not Necessarily a Sign of Cerebral Impairment: A Matched Comparison Between Chronic Fatigue Syndrome and Multiple Sclerosis

*Sieberen P. van der Werf, M.Sc., *Judith B. Prins, M.Sc., †Peter J. H. Jongen, M.D., ‡Jos W. M. van der Meer, M.D., and *Gijs Bleijenberg, Ph.D.

*Department of Medical Psychology, University Hospital Nijmegen; †Multiple Sclerosis Center, Nijmegen; and ‡Department of General Internal Medicine, University Hospital Nijmegen, Nijmegen, The Netherlands

Objective: The aim of this study was to assess the potential impact of effort in comparative studies assessing neurocognitive dysfunction in patients with and without a neurologic diagnosis. Background: It was hypothesized that a subgroup within a group of patients with prominent neurocognitive complaints but without a neurologic diagnosis would have impaired performance on a task originally designed to detect malingering. Method: We compared the neuropsychological performance of a group of 40 patients with a definite diagnosis of multiple sclerosis (MS) with that of 67 patients with chronic fatigue syndrome (CFS). The Amsterdam Short-Term Memory Test, a forced-choice memory task, served as measure to detect submaximal effort. In addition, we administered a regular neuropsychological task generally considered to be sensitive for cognitive deterioration. Results: Compared with the MS group (13%), a larger proportion of the matched CFS group (30%) obtained scores indicative of reduced effort. In contrast, the proportions of patients scoring below the cutoff value on a conventional neuropsychological test did not differ significantly (17% of MS patients and 16% of CFS patients). Conclusions: The results obtained raise the question of to what extent abnormal test findings in the absence of documented neurologic impairment should be interpreted as a sign of cerebral impairment. The suggestion has been made to screen more often for biased results in comparative research studies so as to enhance valid interpretation of neuropsychological findings. NNBN 2000;13:199–203

Chronic fatigue syndrome (CFS) is a medically unexplained but clinically defined condition characterized by long-lasting (≥6 months), severe, disabling fatigue and a combination of at least four of eight specific symptoms (1). One of these eight symptoms is self-reported impairment in concentration and short-term memory. Despite the fact that many patients with CFS report problems with concentration and memory, neuropsychological findings are not consistent (2–9). Recent reviews suggest that these varying results are caused by differences in the diagnostic criteria for CFS, the heterogeneity of the CFS population, small sample sizes, or methodological diversity (10,11).

Imaging studies in CFS have also yielded heterogeneous results, and, until now, no common pattern of cerebral abnormalities has been found (12).

Although several studies have reported relations between psychological factors and the severity and persistence of fatigue in CFS (13–16), little is known about the possible influence of psychological or behavioral factors on neuropsychological performance in CFS. During past clinical and investigative neuropsychological assessments of CFS patients, we sometimes had the impression that the behavior of a small proportion of CFS patients was indicative of psychological distress or poor motivation. We were not able to quantify such behavior or to reliably distinguish between normal and abnormal test behavior, however.

Therefore, the aim of this study was to compare and evaluate the performance of patients with CFS and patients with multiple sclerosis (MS) on a test developed to
detect feigned memory deficits. In the absence of a history of neurologic impairment, abnormal results on this task were interpreted as a sign of submaximal effort or unreliable test performance. Because relations have been found to exist between psychosocial factors and the persistence of fatigue, we hypothesized that psychological factors were likely to influence test performance in a subgroup of CFS patients. Therefore, we expected that a subgroup of CFS patients would have abnormal scores on a detection task. With respect to a conventional neuropsychological task sensitive to neurologic impairment and requiring effort, we hypothesized that equal proportions of the MS and CFS groups would score below normal expectations. Some MS patients would do so because of neurologic impairment, and some CFS patients would do so because of submaximal effort. Hence, we expected that a conventional task would be less likely to differentiate between the two patient groups. Furthermore, we tested whether there was a relation between depressive symptoms and performance on the detection task and conventional test.

METHOD

Subjects

This study was conducted at the Department of Medical Psychology of the University Hospital Nijmegen. Forty nonhospitalized patients with clinically definite MS according to the revised Poser criteria (17) and 144 CFS patients who were medically screened and fulfilled the most recent Centers for Disease Control criteria for CFS (1) received the combination of a detection task and a standard neuropsychological test. We were able to match 67 patients on the basis of gender, educational level, and age. Educational levels were scored on a five-point scale: 1 = primary school only, 3 = high school (approximately 11 years of education), and 5 = university degree (approximately 17 years of education). Subject characteristics of the total CFS sample and both matched patient samples are shown in Table 1. Both patient groups participated in ongoing longitudinal fatigue studies. The assessment of CFS patients was made at baseline, although the MS patients were administered the combination of tests 2 years after undergoing a baseline placebo-controlled treatment study with interferon beta-1a. In the MS group, the Expanded Disability Status Scale (EDSS) of Kurtzke (18) was used to rate the degree of neurologic impairment. An experienced neurologist performed the EDSS ratings.

Instruments

Amsterdam Short-Term Memory Test

The Amsterdam Short-Term Memory Test (ASTMT) (19) is a forced-choice verbal recognition task. The test consists of 30 items, with each item covering three separate pages. On the first page of each item, patients are asked to read aloud five stimulus words from a common semantic category (e.g., clothes, countries, drinks) and to memorize them. Subsequently, this page is turned over, and on the second page, the subject is distracted with a simple arithmetic task (e.g., 23 + 6, 43 + 5, 18 + 7) and is asked to give an oral answer. On the third page, the subject is presented with five words of the same semantic category as those found on the first page. Three words are the same as three on the first page, and two words are new. Subjects are asked to indicate which three words they recognize from the previous trial. The number of words recognized correctly is the total score. The maximum score is 90 points (3 × 30 items).

The validation study (18) showed high average scores (>89) for healthy subjects and mild cerebral head injury patients. The test discriminated perfectly between cerebral head injury patients (all scores above 86) and subjects who were instructed to feign cognitive impairment (score range, 75–85). The authors subsequently used the cutoff value of 86 points in a study investigating postwhiplash complaints in the context of litigation (20). None of the healthy control subjects and none of the patients with closed-head injury and well-documented cognitive impairment scored below the cutoff value. In contrast, a signifi-

<table>
<thead>
<tr>
<th>TABLE 1. Demographics of patient groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>CFS total</td>
</tr>
<tr>
<td>(n = 144)</td>
</tr>
<tr>
<td>Gender M/F (% male)</td>
</tr>
<tr>
<td>26/110 (19.1)</td>
</tr>
<tr>
<td>Mean age (range)</td>
</tr>
<tr>
<td>36.7 (18–60)</td>
</tr>
<tr>
<td>Mean education (range)</td>
</tr>
<tr>
<td>3.2 (1–5)†</td>
</tr>
<tr>
<td>Mean EDSS (range)</td>
</tr>
<tr>
<td>4.1 (1.01–6.5)</td>
</tr>
<tr>
<td>MS (n = 40)</td>
</tr>
<tr>
<td>15/25 (37.5)</td>
</tr>
<tr>
<td>38.5 (22–33)</td>
</tr>
<tr>
<td>3.1 (1–5)‡</td>
</tr>
<tr>
<td>4.1 (1.01–6.5)</td>
</tr>
<tr>
<td>Matched CFS (n = 67)</td>
</tr>
<tr>
<td>25/42 (37.2)</td>
</tr>
<tr>
<td>37.2 (18–54)</td>
</tr>
<tr>
<td>3.2 (2–5)¶</td>
</tr>
<tr>
<td>Matched CFS versus MS (two-tailed p)</td>
</tr>
<tr>
<td>1.00‡</td>
</tr>
<tr>
<td>0.53†</td>
</tr>
<tr>
<td>0.62‡</td>
</tr>
</tbody>
</table>

*Fisher exact test. †Mann-Whitney U test. ‡Range: 1, primary school; 3, approximately 12 years of education; 5, university degree.

CFS, chronic fatigue syndrome; MS, multiple sclerosis; M, male; F, female; EDSS, Expanded Disability Status Scale.
cant proportion of postwhiplash patients had ASTMT scores indicative of submaximal performance, although no differences were found on the regular neuropsychological tests between the closed-head injury group and the postwhiplash group. Furthermore, comparing postwhiplash patients who were assessed for litigation procedure and patients who were not revealed 61% abnormal test scores in the litigation group and 29% in the nonlitigation group.

Symbol Digit Test

The Symbol Digit Test (SDT) of the Dutch version of the Wechsler Adult Intelligence Scale (21) was administered to all subjects. The SDT is considered to be one of the most sensitive subtests of the Wechsler Adult Intelligence Scale for measuring aging and cognitive deterioration (22). Raw SDT scores were converted to T scores according to the most recently published norms (n = 595) in the Netherlands (23). Scores corrected for age and gender were compared with the average SDT T scores of five different educational levels. An SDT T score of 2 SDs below the average SDT score of the appropriate educational group was chosen as a sign of cognitive impairment.

Beck Depression Inventory

The Beck Depression Inventory (BDI) (24) served as a self-report measure of psychological distress. A high level of depressive symptoms was defined as a BDI score above 16.

Data Analysis

The validation studies showed that the data of the ASTMT were considerably skewed toward the maximal score of 90; therefore, nonparametric testing was conducted. Fisher exact tests for dichotomous variables were used to test for proportional differences between groups. Because we hypothesized that the proportion of abnormal ASTMT scores of the matched CFS group would be higher than those of the MS group and normal performance would indicate 0% below the cutoff value, this difference was measured using a one-sided test (p = 0.05). Because no specific directional group differences between the proportions of patients with abnormal SDT and BDI scores were assumed, differences were measured using a two-sided test. The relations between levels of depression and the SDT and ASTMT scores were expressed by Spearman rank correlation coefficients.

RESULTS

ASTMT

Figure 1 displays a cumulative percentage plot of the ASTMT scores of the matched patient group samples. Table 2 shows that compared with the CFS sample, a

![Cumulative percentage plot of Amsterdam Short-Term Memory Test (ASTMT) scores of the chronic fatigue syndrome (CFS) and multiple sclerosis (MS) samples.](image)

**TABLE 2.** Proportions of multiple sclerosis and chronic fatigue syndrome patients scoring below or above cutoff values of the three outcome measures

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>CFS total (n = 144)</th>
<th>MS (n = 40)</th>
<th>Matched CFS (n = 67)</th>
<th>Matched CFS versus MS p values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASTMT &lt;86</td>
<td>29%</td>
<td>13%</td>
<td>30%</td>
<td>0.03*</td>
</tr>
<tr>
<td>SDT &lt; cutoff value</td>
<td>18%</td>
<td>18%</td>
<td>16%</td>
<td>0.89†</td>
</tr>
<tr>
<td>BDI &gt; 16</td>
<td>26%</td>
<td>3%</td>
<td>25%</td>
<td>0.02†</td>
</tr>
</tbody>
</table>

*Fisher exact test, one-sided. †Fisher exact test, two-sided.

ASTMT: Amsterdam Short-Term Memory Test; CFS, chronic fatigue syndrome; MS, multiple sclerosis; SDT, Symbol Digit Test; BDI, Beck Depression Inventory.

significantly smaller proportion of the MS sample obtained scores below the cutoff value of 86 points (13% of MS patients vs. 30% of CFS patients, \( \chi^2 = 4.2, p = 0.03 \)).

SDT

Table 2 also shows that there were no significant differences between the proportions of patients scoring under the proposed cutoff value of the SDT (18% of CFS total patients and 16% of CFS matched patients vs. 18% of MS patients, \( \chi^2 = 0.2, p = 0.89 \)).

Depression

Compared with the MS sample, a significantly larger proportion of the matched CFS group had elevated BDI scores (3% of MS patients vs. 25% of matched CFS patients, \( \chi^2 = 9.4, p = 0.02 \)). In the CFS group, the correlations between BDI scores and ASTMT scores (−0.12,
DISCUSSION

Closed-head injury patients, who were purposely selected because of well-documented brain damage, had no difficulties with the ASTMT, although their performance on a conventional auditory verbal learning task was significantly below that of healthy control subjects (19). Nevertheless, five of our MS patients scored below the proposed cutoff value. These five patients had significantly higher EDSS scores (5.5 vs. 3.9, z = 2.29, p = 0.02) compared with MS patients with normal ASTMT performance, and their ratings (range, 3.5–6.5) indicated moderate to more severe neurologic impairment.

The percentage of CFS patients scoring below the cutoff value was more than twice the size of the percentage of MS patients scoring below that value. Four CFS patients scored below the level of the worst-performing MS patient. Because the ASTMT should be considered as relatively insensitive to mild brain damage or mild cognitive impairment, it would be unlikely that a subtle cognitive dysfunction would cause such discrepancies between MS and CFS patients.

Inspecting the cumulative frequencies of the ASTMT scores showed that a cutoff value of 87 maximized the group differences. Forty percent of the CFS patients and 18% of the MS patients scored below 87 points (x² = 6, p = 0.01). In their previous research, the authors of the ASTMT did suggest that the proposed cutoff value of 86 points was perhaps too stringent and that a cutoff value of 87 points would be more effective. Nevertheless, to minimize the chance that patients would be classified falsely as malingeringers, they maintained the proposed cutoff value (20).

Although equal proportions of the MS sample had abnormal ASTMT and SDT scores, abnormal ASTMT performance did not always coincide with an abnormal SDT score. One would expect that impaired performance on a relatively easy task would also lead to abnormal results on a more sensitive task. Nevertheless, 5% of the MS patients had an abnormal ASTMT score combined with a normal SDT score. This finding could suggest that the SDT cutoff value of 2 SDs below expectation was too stringent. A cutoff value of 1.5 SDs did indeed result in 0% of the MS patients having an abnormal ASTMT score in combination with a normal SDT score.

When a similar comparison was made for CFS patients, 19% of the CFS patients had an abnormal ASTMT score and an SDT score greater than 2 SDs below expectations. Still, 13% of the CFS patients had an abnormal ASTMT score in the presence of an SDT score greater than 1.5 SDs below expectations (0% of MS patients vs. 13% of CFS patients, two-sided p = 0.02). This finding demonstrated that the disparity in the CFS group between performance on a relatively easy task and on a more sensitive measure could not merely be resolved by choosing an appropriate cutoff value.

This disparity in the CFS sample could indicate that the ASTMT is sensitive to a specific cognitive ability. Perhaps some CFS patients are characterized by a specific verbal recognition deficit; however, two previous studies that used the Californian Verbal Learning Test and compared the performance of a group of CFS patients with that of a group of healthy control subjects showed no differences in first trial performance or between the average delayed recognition scores (2, 25). Two review articles on neuropsychological function in CFS (10, 11) did not mention evidence of a specific verbal recognition deficit, and one of these reviews even stated that memory deficits have not been found for word learning lists (11).

The discrepancy between the findings in the current study and the absence of verbal recognition deficits in other research could stem from the fact that patients who complain about memory deficit experience excessive stress when confronted with a memory task or assume beforehand that they will fail on such a task. Alternatively, it is possible that some CFS patients fear that their problems go unnoticed on a relatively easy task and adjust their effort to the difficulty or type of task involved. Because subjective neuropsychological complaints are one of the eight key symptoms of the Centers for Disease Control criteria of CFS, some patients may adapt their behavior to their disease expectations. The unknown etiology of CFS, the “difficult” disease status, and subsequent fear of complaints not taken seriously could enhance such a process.

Although the proportion of elevated depression ratings was significantly higher for the CFS group, there was no significant relation between levels of self-reported depression and the ASTMT and SDT scores. Deluca et al (6) also found no significant relation between neuropsychological performance and psychological ratings on questionnaires for somatization, depression, and distress. Ray et al (26) reported significant correlations between levels of emotional distress and word and color naming on the Stroop Color-Word Interference Test, however. A previous CFS study by our group showed a relation between the extent of physical causal attributions and physical activity reduction. Future research could assess whether performance on neuropsychological tasks is also related to those types of attributions.

Because a low ASTMT score does not provide any information about the motivational basis of that finding, it is not suggested that abnormal neuropsychological test findings in CFS be exclusively attributed to psychological...
factors. Nevertheless, extremely low ASTMT scores indicated that such factors almost certainly played a role in some cases. To warrant valid neuropsychological comparisons between neurologic disorders and disorders of unknown etiology, one should screen for biased results. Comparative group studies can be especially confounded by invalid performance of subgroups. To attribute abnormal neuropsychological findings to cerebral impairment could be as harmful as too readily interpreting those results as a consequence of psychological processes. Implying cerebral deficits can perhaps confirm or strengthen patient belief that a possible neurologic disease is the cause of experienced cognitive problems. Although most instruments and methods for detecting submaximal performance are still far from perfect, application of these methods for research purposes might at least reduce the heterogeneity of neuropsychological test results to some extent.

REFERENCES