Journal of the International Neuropsychological Society (2006), **12**, 549–558.
Copyright © 2006 INS. Published by Cambridge University Press. Printed in the USA.
DOI: 10.1017/S1355617706060723

Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS)

RALPH H.B. BENEDICT,[1] DIANE COOKFAIR,[2] REBECCA GAVETT,[1] MARK GUNTHER,[2] FREDERICK MUNSCHAUER,[2] NEETA GARG,[2] AND BIANCA WEINSTOCK-GUTTMAN[1]

[1]State University of New York (SUNY) at Buffalo School of Medicine, Department of Neurology, Division of Departmental and Behavioral Neurosciences, Buffalo, New York, USA
[2]Jacobs Neurological Institute, Buffalo General Hospital, Buffalo, New York, USA

(RECEIVED January 23, 2006; FINAL REVISION March 23, 2006; ACCEPTED March 23, 2006)

Abstract

Cognitive impairment occurs in roughly 50% of patients with multiple sclerosis (MS). It is well known that processing speed and episodic memory deficits are the most common neuropsychological (NP) sequelae in this illness. Consensus has emerged about the specific tests that prove most helpful for routine monitoring of MS associated cognitive impairment. The purpose of this study was to examine the validity of the Minimal Assessment of Cognitive Function in MS (MACFIMS), a recommended battery based on the findings of an international conference held in 2001. We tested 291 MS patients and 56 healthy controls. Frequencies of impairment paralleled those reported in previous work for both individual cognitive domains and general impairment. All tests were impaired in the MS group, and distinguished relapsing-remitting (RR) from secondary progressive (SP) course. Principle components analysis showed a distinct episodic memory component. Most of the MACFIMS tests discriminated disabled from employed patients. However, in regression models accounting for all NP tests, those emphasizing verbal memory and executive function were most predictive of vocational status. We conclude that the MACFIMS is a valid approach to routine NP assessment of MS patients. Future work is planned to determine its psychometric properties in a longitudinal study. (*JINS*, 2006, *12*, 549–558.)

Keywords: Multiple sclerosis, Processing speed, Memory, Neuropsychological testing, Validity, Employment

INTRODUCTION

The impact of multiple sclerosis (MS) on neuropsychological (NP) function is well recognized, in that nearly half of all MS patients show deficiency on NP testing (Heaton, 1985; Rao et al., 1991a). Research also indicates that cognitive impairment gives rise to vocational disability (Benedict et al., 2005a; Rao et al., 1991b), social impairment (Benedict et al., 2001b; Rao et al., 1991b), and distress in persons caring for MS patients (Knight et al., 1997). Yet NP testing is not routinely employed in MS clinics for a variety of reasons. NP impairment is not a diagnostic criterion for MS diagnosis (McDonald et al., 2001). Board certified or well-trained clinicians with MS expertise are not available to many MS Centers. Most importantly, NP testing is often costly and time consuming. Combined, these factors impede the implementation of routine NP testing for MS patients.

In 2001, we reasoned that routine NP testing of MS patients might be more feasible if it were conducted in a more standardized and cost-effective manner. In collaboration with the Consortium of MS Centers, we convened a panel of neuropsychologists and psychologists with expertise in MS from the United States, Canada, United Kingdom, and Australia. The objectives of the meeting were to (a) propose a minimal NP examination for clinical monitoring of MS patients and research and (b) propose strategies for improving NP assessment of MS patients. Literature on MS-related cognitive dysfunction was reviewed with an eye toward both psychometric factors (reliability, discriminative validity) and practical considerations (e.g., alternate forms, brevity of administration). The meeting culminated in a publication (Benedict et al., 2002) summarizing the panel discussion and a 90-minute NP battery called

Correspondence and reprint requests to: Ralph Benedict, PhD, Buffalo General Hospital, D-6, 100 High Street, Buffalo, NY 14203. E-mail: benedict@buffalo.edu

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

the Minimal Assessment of Cognitive Function in MS (MACFIMS). Included were seven tests covering the five cognitive domains most commonly impaired in MS patients (Table 1). Tests emphasizing information processing speed, working memory, and unaided recall of recently presented information were judged to be most sensitive to the cerebral pathology of MS (Archibald & Fisk, 2000; Beatty, 1996; DeLuca et al., 1998; Demaree et al., 1999; Fischer, 1988; Rao et al., 1989a; Rao et al., 1989b). Reviewed research also showed that tests in other domains are also, if less frequently, affected in MS patients, especially those involving generative fluency (Beatty & Monson, 1994), spatial processing (Vleugels et al., 2000), and conceptual reasoning (Arnett et al., 1997; Beatty & Monson, 1996). While an encouraging first step, the authors acknowledged that the MACFIMS battery, merely a collection of recommended tests, had yet to be validated as a single battery in a large, diverse sample of MS patients.

In the course of preparing this manuscript, we searched the literature for studies that had included any of the individual NP tests comprising the MACFIMS battery to evaluate MS patients. Included in Table 2 are those studies comparing MS patients and healthy controls and reporting data that permit calculation of effect size using the well known Cohen *d* statistic (Cohen, 1988). Table 2 shows that of 28 between-group comparisons only 2 have effect sizes less than $d = .5$. Most show medium or large effects. Ten comparisons show a between-group difference of 1.0 SD or greater. Such strong effects are not surprising considering that the MACFIMS tests were chosen based largely on evidence for reliability and discriminative validity.

Here, for the first time, we examined the concurrent and construct validity of the entire MACFIMS battery in a large sample of 291 prospectively accrued MS patients. We predicted *a priori* that each of the individual tests that comprise the MACFIMS battery would discriminate MS patients from healthy controls, and that effect sizes and frequencies of impairment would parallel previous findings in the literature. We then explored the construct validity of these tests using principle components analysis. Because episodic memory was distinguished from processing speed in previous literature reviews (Benedict et al., 2002), we predicted that such an analysis would show a clearly discernible episodic memory component. Finally, we predicted that the MACFIMS tests would show good ecological validity regarding their associations with vocational disability status.

METHODS

Participants

Studied were 291 patients diagnosed with clinically definite MS (McDonald et al., 2001) by treating or research neurologists. Patients entered the study for one of three reasons: participation in brain imaging research ($n = 77$, 27%), routine monitoring of cognitive function ($n = 106$, 36%), or referral for evaluation of a specified management problem related to suspected cognitive impairment ($n = 108$, 37%). All patients met the following inclusion and exclusion criteria: (a) age 18 or older; (b) fluent in English; (c) able to provide informed consent to all procedures; (d) no neurological disorder other than MS; (e) no psychiatric disorder (American Psychiatric Association, 1994) other than mood, personality, or behavior change following the onset of MS; (f) no other medical condition that might influence cognition; (g) no history of developmental disorder (e.g., ADHD, learning disability); (h) no history of substance or alcohol dependence, or current abuse; (i) no motor or sensory defect (e.g., corrected near vision of at least 20/70) that might interfere with cognitive test performance; and (j) no relapse and/or corticosteroid pulse within four weeks of assessment. All research participants signed informed consent forms approved by institutional review panels prior to participating in the study.

Mean patient age (±SD) was 45.4 ± 8.9 years. On average, patients completed 14.4 ± 2.4 years of education. The

Table 1. Tests included in the Minimal Assessment of Cognitive Function in MS (MACFIMS) battery

Test	Domain	Authors	Most common measures
Controlled Oral Word Association Test	Language	(Benton et al., 1994)	Total # Correct
Judgment of Line Orientation Test	Spatial processing	(Benton et al., 1994)	Total # Correct
California Verbal Learning Test, Second Edition	New learning and memory	(Delis et al., 2000)	Total Learning Trials 1–5 Delayed Recall
Brief Visuospatial Memory Test—Revised	New learning and memory	(Benedict et al., 1996; Benedict, 1997)	Total Learning Trials 1–3 Delayed Recall
Symbol Digit Modalities Test	Processing speed and working memory	(Smith, 1982; Rao et al., 1991b; Rao, 1991)	Total # Correct
Paced Auditory Serial Addition Test	Processing speed and working memory	(Gronwall, 1977; Rao et al., 1991b; Rao, 1991)	3.0 ISI and 2.0 ISI, # Correct
Delis-Kaplan Executive Function System Sorting Test	Executive function	(Delis et al., 2001)	# Correct Sorts Description Score

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

Table 2. Effect sizes derived from previously published studies using the MACFIMS battery tests to discriminate MS patients from healthy controls

Study	MS Sample	Test variable	NC mean	NC SD	MS Mean	MS SD	*d*	*P*
(Archibald & Fisk, 2000); (Fisk & Archibald, 2001)	35 mixed course	CVLT-DR	13.3	2.1	11.9	2.6	0.6	nr
(Beatty & Monson, 1996)	30 mixed course	CCST-CS	17.8	2.4	14.6	3.8	1.0	.001
		CCST-DS	31.9	4.8	26.6	7.8	0.8	.01
		COWAT	41.5	9.9	32.3	14.0	0.8	0.01
(Beatty & Monson, 1994)	30 mixed course	COWAT	41.5	9.9	32.3	14.0	0.8	0.01
(Beatty et al., 1995c; Beatty et al., 1995b)	100 mixed course	COWAT	47.7	14.2	32.1	13.3	1.1	0.001
		JLO	25.3	3.7	23.7	5.2	0.4	ns
		SDMT	59.0	10.2	43.7	15.0	1.2	0.001
		DKEFS-CS	6.2	1.0	5.0	1.6	0.9	0.001
		DKEFS-RS	0.2	0.5	0.6	0.8	0.6	0.01
(Beatty et al., 1989)	42 mixed course	SDMT	59.0	9.1	48.3	11.4	1.0	0.001
		COWAT	42.8	12.7	34.5	11.9	0.7	0.05
(Benedict et al., 2001a)	34 mostly SP	JLO	24.1	4.7	17.7	8.6	1.0	0.01
		CVLT-TL	57.2	17.2	41.3	15.1	1.0	0.01
		BVMT-R-TL	21.1	5.5	14.1	7.9	1.1	0.01
(D'Esposito et al., 1996)	36 mostly RR	COWAT	48.7	15.7	42.2	10.9	0.5	ns
		SDMT	63.1	7.6	53.1	12.4	1.0	.01
(Krupp et al., 1994)	20 mixed course	COWAT	48.0	8.3	39.8	13.6	0.8	0.04
(Rao et al., 1991b)	100 mixed course	COWAT	43.8	11.8	34.7	12.1	0.8	<.001
		JLO	27.2	4.1	25.0	4.8	0.5	<0.01
		PASAT3	48.5	9.6	41.6	11.5	0.7	<.001
		PASAT2	37.1	10.1	30.8	10.5	0.6	<.001
(Ryan et al., 1996)	177 RR	COWAT	41.7	10.3	37.0	11.7	0.4	<.01
(Scarrabelotti & Carroll, 1998)	50 mixed course	CVLT-TL	54.0	14.6	46.2	15.7	0.5	0.02
(Sperling et al., 2001)	28 mixed course	COWAT	34.8	9.8	26.6	6.6	1.0	<0.01
		PASAT3 PC	87.8	13.8	67.9	22.8	1.1	<0.01
(Thornton et al., 2002)	49 mixed course	CVLT-TL	58.7	9.5	50.9	11.9	0.7	<.001
		CVLT-DR	12.4	2.8	10.6	3.7	0.6	<.01

Note. Studies included report mean and SD for MS and control group, allowing calculation of effect size d. Only test forms included in the MACFIMS listed. For example, studies using the written form of the Symbol Digit Modalities Test or the Gronwall version of the PASAT are not included.
Abbreviations: nr = not reported. ns = not significant. MS = multiple sclerosis, NC = normal control, RR = relapsing-remitting course, SP = secondary progressive course after relapsng-remitting course, COWAT = Controlled Oral Word Association Test, JLO = Judgment of Line Orientation Test, CVLT = California Verbal Learning Test, CVLT-TL = CVLT Total Learning, CVLT-DR = CVLT Delayed Recall, BVMTR = Brief Visuospatial Memory Test–Revised, BVMTR-TL = BVMTR Total Learning, PASAT = Paced Auditory Serial Addition Test, PASAT3 = 3.0 inter-stimulus interval of PASAT, PASAT2 = 2.0 inter-stimulus interval of PASAT, PC = Percent Correct, SDMT = Symbol Digit Modalities Test, CCST = California Card Sorting Test, and an early version of the DKEFS = Delis Kaplan Executive Function System Sorting Test, DKEFS-CS = DKEFS Correct Sorts, DKEFS-DS = DKEFS Description Score.

majority were Caucasian (n = 277 or 95%) and women (n = 227 or 78%), consistent with the MS population which is primarily female (Jacobs et al., 1999). Scores derived from the Expanded Disability Status Scale (Kurtzke, 1983), which assess neurologic and physical disability, were available for 186 patients; the mean EDSS for this subset of patients was 3.0 ± 1.8, representing mild and moderate impairment. Most patients (n = 200 or 69%) had relapsing-remitting (RR) rather than secondary-progressive (SP) (n = 78 or 27%), progressive-relapsing (n = 6 or 2%), or primary-progressive (n = 7 or 2%) course.

Patients were compared with 56 healthy controls with a mean age of 43.8 ± 9.5 years and 14.6 ± 2.2 years of education. The majority of the controls were Caucasian (n = 50, or 89%) and women (n = 42 or 75%). The differences between the MS and control groups on these demographic variables did not reach significance by ANOVA or χ^2 test.

MATERIALS AND METHODS

Each patient underwent NP testing in accordance with recent consensus panel recommendations (Benedict et al., 2002). The MACFIMS includes the following tests: Controlled Oral Word Association Test (COWAT; Benton et al., 1994), Judgment of Line Orientation Test (JLO; Benton et al., 1994), California Verbal Learning Test, second edition (CVLT-II; Delis et al., 2000), Brief Visuospatial Memory Test-Revised (BVMT-R; Benedict, 1997), Paced Auditory Serial Addition Test (PASAT; Gronwall, 1977), Symbol Digit Modalities Test (SDMT; Smith, 1982), and the Sorting Test from the Delis-Kaplan Executive Function System (DKEFS; Delis et al., 2001). In accordance with the position paper (Benedict et al., 2002), we employed Rao's adaptations (Rao, 1991) of the PASAT and SDMT.

The COWAT was administered in the standard manner, following the method of Arthur Benton (Benton et al., 1994).

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

In successive one-minute trials, participants generated as many words as possible, beginning with each of three designated letters. The dependent measure was the total number of correct words over the three trials.

The JLO required participants to identify the angle defined by two stimulus lines from among those defined by a visual array of lines covering 180 degrees. Both oral and pointing responses were allowed. The dependent variable was the total number of correct responses over 30 items.

The CVLT-II and BVMT-R are both learning and memory tests with similar formats. Both require discrete exposures to new material that are followed by the participant's unaided recall immediately after presentation. There is a 25-min interval following the final learning trial, after which participants are asked to recall the information again without further exposure to the to-be-learned material. Delayed recall is followed by a yes or no, forced-choice recognition task. For the CVLT-II there were five learning trials. Examiners read 16 words and asked participants to repeat as many words as possible. The entire word list was repeated each time. For the BVMT-R, the stimulus material was a matrix of six visual designs, held before the participant for 10 sec. Participants were asked to render the designs using paper and pencil, taking as much time as needed for reproduction. Each design received a score of 0, 1, or 2 based on accuracy and location scoring criteria. There were three free-recall trials followed by 25-min delayed recall and yes or no recognition trials. In this study, we considered the following measures for each memory test: total recall over all learning trials (Total Learning or TL) and recall after the delay interval (Delayed Recall or DR, i.e., long delayed recall in the case of the CVLT-II).

The PASAT included 60 trials presented at inter-stimulus intervals of 3 (PASAT-3) and 2 (PASAT-2) sec. The dependent measure was the number of correct responses from each of the two trials. For the SDMT, participants were presented with a series of nine symbols, each paired with a single digit in a key at the top of an 8.5 × 11-inch sheet. The remainder of the page presented a pseudo-randomized sequence of symbols. Participants responded by voicing the digit associated with each symbol as quickly as possible. The dependent measure was the number of correct responses in 90 sec.

The DKEFS Sorting Test was employed to assess higher executive function. Participants were presented with six cards each depicting a single word. The cards varied in many ways, allowing conceptual sorting in accordance with at least eight different principles (e.g., card shape, card color, semantic association among words). This was a timed test with 4 min allowed for each of two card sets. We recorded the total number of correct sorts (CS) and calculated the verbal description score (DS), which was based on the abstractness and accuracy of the sort descriptions.

Depression was assessed using the Beck Depression Inventory—Fast Screen for Medical Patients (BDI-FS; Beck et al., 2000). The BDI-FS emphasizes thought content (e.g., negative self-evaluation or guilt) and mood (e.g., dysphoria) and avoids assessment of vegetative signs that can occur in medical illness without depression. This test was recently validated in a MS sample (Benedict et al., 2003).

Finally, we assessed vocational status by asking patients about their present work activities (Benedict et al., 2005a). There were two methods used to classify patients as either employed or disabled. A *conservative* approach to classification required that working patients be gainfully employed with pay, full time, without demotion, reported reprimands, or loss of pay due to MS related problems. Disabled patients were required to be receiving formal disability benefits from either public (e.g., social security disability income) or private (e.g., commercial benefits) sources. The *liberal* approach to classification required that working patients be gainfully employed with pay at least 20 hours per week. Disabled patients for this scheme were required to be receiving either formal disability as described above, or to be unemployed for reasons reported by them or informants to be disease related.

Procedure

The research participants (including all HCs) were contacted by mail or were approached during the course of their usual clinical care at an MS center. The clinical participants were patients undergoing evaluation either for routine monitoring of cognitive functioning, or for specific clinical reasons (e.g., concerns about driving or work capacity, disability evaluation, differential diagnosis of cognitive impairment *vs*. depressive disorder, etc). The tests were administered individually by a psychologist, trained assistant, or a graduate student under the supervision of a board-certified (ABPP-CN) neuropsychologist. The entire test battery, including tests of personality, required 90–120 min of face-to-face testing time.

Statistical Analysis

All statistical analyses were performed using SPSS 13.0. Group differences in continuous variables were assessed using unpaired Student's *t*-test or univariate ANOVA with alpha set at $p < .01$. The chi-square test was used to determine differences in categorical variables. We considered limiting the study to RR and SP patients, because there were few patients with progressive-relapsing and primary-progressive course. However, one of the goals of the study was to test the validity of the MACFIMS in a sample that represents all MS patients receiving clinical treatment. Therefore, normal *versus* MS comparisons included these rarer disease courses.

For descriptive purposes, we examined effect sizes using Cohen's *d* statistic, which is the difference between means divided by the pooled SD. *Z* scores were calculated for each individual NP test based on the normal control group. Impairment for a single test was defined as a *z* score < -1.5. NP

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

impairment generally was defined as a defect on two or more test measures.

Principal Components Analysis (PCA) with varimax rotation was used to assess the construct validity of the MACFIMS battery. Because mean NP test scores and levels of impairment were found to differ by course for all NP tests in the MACFIMS battery, separate PCAs were conducted for RR and SP patients to evaluate whether the underlying components (dimensions) in the MACFIMS battery vary according to disease course. All 11 test variables were included in both models. Components were identified using a criteria of Eigenvalue >1.0. Each component was required to consist of at least two variables with loadings of .5 or more as well as explain at least 15% of the variance in the model. The NP variables most representative of a specific component were considered to be those with loadings ≧ .500 for that variable.

Estimates of the risk of being disabled or unemployed associated with NP impairment were obtained from odds ratios calculated using unconditional logistic regression models. Like the PCA, this analysis did not include the 13 progressive-relapsing and primary-progressive patients. One of the criteria for inclusion in final stepwise regression testing was that NP variables significantly discriminate employed from disabled patients after adjusting for covariates that could also impact on vocational status. Thus, unconditional logistic regression models were first used to assess the association between individual NP tests and employment/disability status after adjusting for age, sex, years of education, depression status (*depressed: BDF-FS Score >3; not depressed BDI-FS* ≤ 3), and disease course (*SP vs. RR*). NP tests that were not significant after adjusting for covariates in the preliminary analyses were not entered into the final models. In order to further limit the number of predictor variables in the final models, we selected only one measure from each test (i.e., COWAT, JLO, CVLT-II-DR, BVMT-R-DR, DKEFS-DS, PASAT-3, SDMT) in the final stepwise model, specifically, the measure that showed the highest frequency of impairment. Vocational status (disabled/employed) served as the dependent variable for each model. Two sets of models were run using two different definitions of vocational status (conservative and liberal, see Methods). In preliminary analyses we noted that only 13% of SP patients were employed, leaving only 6 employed SP patients for analysis. Therefore, we also ran a separate set of models for RR patients only, using the same criteria and following the same steps outlined earlier.

RESULTS

Group Differences and Frequency of Impairment

Table 3 shows the group data for the MACFIMS tests. ANOVA showed significant between-group effects for all tests with effect sizes ranging from 0.49 for JLO to 1.31 for SDMT. Frequencies of impairment in MS patients (Figure 1) were as follows: COWAT 13.1%, JLO 22.3%, CVLT-II-TL 29.6%, CVLT-II-DR 34.4%, BVMT-R-TL 54.3%, BVMT-R-DR 56.0%, PASAT-3 27.4%, PASAT-2 27.5%, SDMT 51.9%, DKEFS-CS 15.8%, DKEFS-DS 25.8%. Based on the standard of impairment on two or more tests, 59.5% of MS patients were found to be impaired.

MS patients also showed greater degrees of depression on the BDI-FS (MS mean = 3.6 ± 3.3, Normal mean = 0.9 ± 1.5; $p < .001$). However, there were no significant correlations between BDI-FS and NP test performance in either group (largest MS r = −.15 for SDMT).

ANOVAs also showed that all 11 NP variables significantly discriminated RR and SP patients at a probability of $p < .01$. In each case, RR patients performed better than SP patients. Effect sizes ranged from 0.42 for JLO to 0.92 for SDMT.

Table 3. Group data comparing MS patients and normal healthy controls

Test	NC Mean	NC *SD*	MS Mean	MS *SD*	*d*	*P*
COWAT	40.7	12.3	34.6	10.7	0.53	<.001
JLO	25.0	3.4	22.7	5.9	0.49	.004
CVLT-II Total Learning	55.7	9.4	48.3	11.6	0.70	<.001
CVLT-II Delayed Recall	12.4	2.5	10.0	3.6	0.79	<.001
BVMT-R Total Learning	27.7	5.2	21.0	7.7	1.04	<.001
BVMT-R Delayed Recall	10.6	1.6	8.2	2.9	1.07	<.001
PASAT 3.0 ISI	47.2	10.6	39.6	14.5	0.61	<.001
PASAT 2.0 ISI	36.7	12.0	29.7	12.3	0.58	<.001
SDMT	62.8	8.8	47.8	14.1	1.31	<.001
DKEFS Sorting CS	11.0	2.1	9.5	2.6	0.64	<.001
DKEFS Sorting Description	42.2	8.4	35.3	11.6	0.69	<.001

Note. Groups matched on all demographics.
Abbreviations: COWAT = Controlled Oral Word Association Test; JLO = Judgment of Line Orientation Test; CVLT-II = California Verbal Learning Test, second edition; BVMT-R = Brief Visuospatial Memory Test—Revised; PASAT = Paced Auditory Serial Addition Test; SDMT = Symbol Digit Modalities Test; DKEFS = Delis-Kaplan Executive Function System.

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

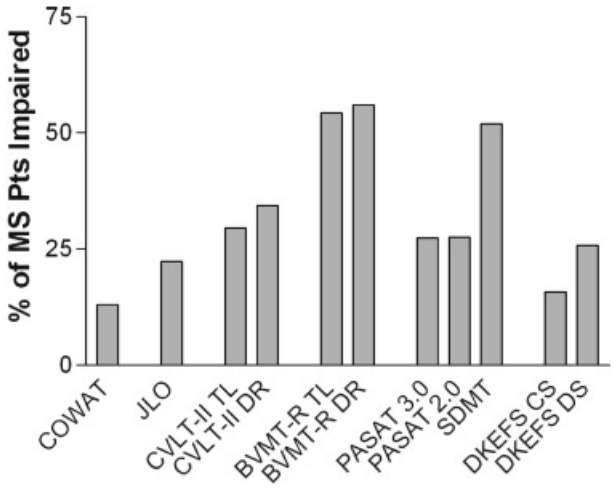


Fig. 1. Frequencies of impairment using a cut-off z score of −1.5 for the healthy comparison group.

Principle Components Analysis

The *PCA yielded similar results for RR and SP patients (Table 4). Three components were derived for RR patients: (1) processing speed/working memory, (2) memory, and (3) executive function. Each of the three components explained over 15% of the variance in the 11 variables included in the analysis and 70.1% of the variance was explained by the three components combined. COWAT showed only modest loadings on all three components. Similarly, the low communality for this variable (.243) indicates that the three components explained only a very small percentage of COWAT variance. The two BVMT-R variables loaded most heavily onto the memory component, but they also loaded significantly onto the processing speed component. Although the loading for SDMT was only over 0.5 for component 1 (processing speed), it also loaded marginally on the memory component.

The PCA yielded two components for SP patients: a general component encompassing language, processing speed, working memory, executive function, and a more specific memory component. Each explained over 15% of the total variance. Overall, 67.9% of the variance was explained by the two components combined. The only variable that loaded onto both components (i.e., loading >.5) was the SDMT.

Association with Vocational Status

When the more conservative definition of vocational status was used, 5 NP tests (COWAT, CVLT-DR, PASAT-3, SDMT, DKEFS-DS) were significant in the logistic models that included only a single NP test adjusted for covariates (i.e., age, sex, education, course, and depression), with odds ratios ranging from 3.07 for SDMT to 6.02 for COWAT ($p <$ .01). The final model, which included covariates and the five NP variables found to be significant in preliminary analyses, retained CVLT-DR and PASAT-3 (Table 5).

When the liberal definition of vocational status was used, the following measures were found to be significant in preliminary logistic models that included only a single NP test: COWAT, CVLT-II-DR, BVMT-R DR, PASAT-3, SDMT, DKEFS-DS. After adjusting for covariates, odds ratios for these individual NP tests ranged from 2.05 ($p = .03$) for SDMT to 3.25 ($p < .01$) for COWAT. In the final model predicting disability/employment, which included all six of the NP variables found to be significant in the preliminary logistic regression analyses, after controlling for covariates, DKEFS-DS and CVLT-II-DR were retained (Table 5).

These analyses were repeated including only the RR patients. In the final model predicting the conservative disability/employment status for RR patients, there were 3 retained NP tests: DKEFS-DS (odds ratio 3.06, $p = .10$), CVLT-II-DR (odds ratio 2.85, $p = .07$), and PASAT-3 (odds ratio 3.57, $p = .03$). In the liberal model, BVMT-R-DR (odds ratio 2.25, $p < .05$) and DKEFS-DS (odds ratio 2.65, $p < .05$) were retained.

DISCUSSION

The primary purpose of this study was to investigate the validity of the MACFIMS battery in a single, prospectively-accrued sample. All seven NP tests discriminated MS patients from healthy controls and RR from SP patients. Effect sizes ranged from medium to very large, and tests emphasizing processing speed and memory were most sensitive to illness. PCA identified a single episodic memory component in both RR and SP samples, and components corresponding to either processing speed and working memory, or a more general component involving these constructs, as well as spatial and executive functions. For most of the MACFIMS tests, poor cognitive performance was associated with vocational disability; in logistic regression models accounting for all NP tests, tests emphasizing processing speed, verbal memory, and executive function were most predictive of vocational status, even after controlling for the effects of age, education, sex, depression, and disease course.

As predicted, the MACFIMS tests significantly discriminated MS patients and controls and showed a frequency of impairment consistent with the previous MS literature. The effect sizes were roughly parallel with those obtained previously using the same test measures (Table 2). Like the seminal work of Rao and colleagues (Rao et al., 1991a), we found that roughly half of MS patients exhibit cognitive impairment and that impairment is most common on tests that emphasize processing speed, working memory, and recent or episodic memory. Differences in the frequency of impairment can be expected to differ across studies because

*Because of likely correlation between NP measures, oblique rotation was also pursued. This approach resulted in the same frequency of variable overlap seen with varimax rotation. In addition to significant overlap between test scores, the oblique rotation resulted in components that did not meet the *a priori* criteria that a component must explain at least 15% of the variance in the model. In at least one instance, a variable that had loaded over 0.5 on a single component and less than 0.5 on the other components with varimax method now loaded as almost 0.5 on multiple components. Both rotation methods resulted in overlap among test scores, but with the varimax rotation we met the *a priori* criteria described above.

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

Table 4. Results of Principle Components Analysis

	RR (N = 200)					SP (N = 78)		
	C1	C2	C3	h^2		C1	C2	h^2
COWAT	.345	.174	.315	.248	COWAT	.453	.385	.375
JLO	**.589**	.180	.142	.399	JLO	**.604**	.385	.513
CVLT-II Total Learning	.124	**.833**	.292	.795	CVLT Total Learning	.136	**.855**	.750
CVLT-II Delayed Recall	.128	**.902**	.143	.851	CVLT Delayed Recall	.135	**.892**	.814
BVMT-R Total Learning	**.594**	**.627**	.103	.757	BVMT Total Learning	.368	**.746**	.693
BVMT-R Delayed Recall	**.589**	**.628**	.071	.747	BVMT Delayed Recall	.378	**.757**	.716
PASAT 3.0 ISI	**.836**	.159	.262	.793	PASAT 3.0	**.832**	.269	.765
PASAT 2.0 ISI	**.775**	.008	.285	.682	PASAT 2.0	**.815**	.191	.701
SDMT	**.655**	.466	.097	.655	SDMT	**.583**	**.567**	.661
DKEFS Correct Sorts	.266	.237	**.866**	.877	DKEFS Correction Sorts	**.835**	.181	.729
DKEFS Description Score	.195	.118	**.919**	.896	DKEFS Description Score	**.844**	.191	.749
% variance explained	27.6%	24.3%	18.1%	70.1%	% variance explained	37.3%	31.6%	67.9%

Note. Table shows loadings using varimax with Kaiser rotation for each component (C). Communalities (h^2) are the proportion of variance in each variable explained by the derived factors. Final row shows percent of variance explained by all 11variables for factor and the total model.

of variability in testing methods and sample composition. This sample, which includes patients undergoing routine monitoring of cognitive function, patients referred to address specific clinical problems (e.g., disability, cause of established memory disorder), as well as research volunteers, is representative of patients attending a hospital-based MS center. Lower frequencies of impairment would be expected in a population-based or community sample, and higher frequencies might be found in samples of patients seen only for specified clinical purposes.

In evaluating the PCAs, it is important to bear in mind that the MACFIMS tests were chosen to represent a minimal record of cognitive function in MS, and not to assess cognitive functioning in a comprehensive manner (Benedict et al., 2002). Therefore, these results should not be viewed as a representation of cognitive functioning in MS. Also, the MACFIMS includes only one test of language, the COWAT, which equally emphasizes lexical search and response speed. Indeed, this test is also commonly referred to as a test of executive ability. However, we did expect *a priori* that a single component representing episodic memory would be seen in the data, a hypothesis borne out by the PCA models for both RR and SP patients.

Fewer components emerged for the SP sample, which may be related, in part, to the low number of participants. It is also possible that as the disease progresses there is less variation in the NP presentation of MS. Although cognitive impairment can be found early in the disease, most studies

Table 5. Logistic regression models for MACFIMS tests associated with increased risk of vocational disability in all MS patients

	Model predicting disabled/employed status, conservatively defined (n = 102)			Model predicting disabled/employed status, liberally defined (n = 162)		
	Odds Ratio	95% Confidence Interval	p	Odds Ratio	95% Confidence Interval	p
Depression†	2.30	0.95–5.54	0.065	2.70	1.45–5.04	.002
Disease Course††	5.79	1.88–17.8	0.002	5.00	2.2–11.4	<.001
COWAT			ns			ns
CVLT-II-DR	3.20	1.19–8.63	0.022	2.82	1.41–5.66	.003
BVMT-R-DR			ns			ns
PASAT 3.0	4.34	1.63–11.6	0.003			ns
SDMT			ns			ns
DKEFS DS			ns	2.27	1.03–5.01	.042

†BDI-FS Score >3 = depressed. †† RR *versus* SP.
For both models, NP variables (impaired *vs*. not impaired), depression status, course and demographics entered first. Backward stepwise logistic regression then used to establish the best final model. In order to increase the ratio of sample size to variables, thereby increasing power, all covariates with $p \leq .15$ retained, Variables with $p > 0.15$ eliminated from the final model.

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

show that patients with progressive disease are more neuropsychologically affected than RR patients (Comi et al., 1995; Heaton, 1985; Kraus et al., 2005). In a more demented state, impairment becomes more common across all NP tests, which makes identification of underlying components difficult.

It is interesting that the SDMT had a split loading in the PCA for SP patients, and nearly so in the RR analysis. This test is regarded to be very strongly correlated with brain atrophy in recent MS studies (Benedict et al., 2004; Christodoulou et al., 2003) perhaps because, like the PASAT (Lockwood et al., 2004), performance depends on rapid communication between widely dispersed cortical regions *via* long white matter tracts. However, in our previous work, the SDMT was correlated not only with general measures of brain atrophy but also regional temporal lobe atrophy (Benedict et al., 2005b). The explanation for this may lie in the fact that SDMT performance is also enhanced by remembering new associations between symbols and numbers. This multimodal aspect of the SDMT may account for its very high sensitivity in the present study.

The ecological validity of NP testing is increasingly emphasized in the MS literature (Goverover et al., 2006; Higginson et al., 2000; Beatty et al., 1995a). Evaluations are often requested to determine whether or not a patient has a legitimate claim for disability benefits. In this study, logistic regression analysis determined the extent to which these MACFIMS tests are associated with a dichotomous distinction of disabled *versus* employed. This is a difficult psychosocial construct to measure. Our method was based on patient and informant reported status ascertained at the time of NP testing. Whereas we have no reason to suspect that these reports were biased, our measurements do not account for why a person is disabled. To account for this, we eliminated patients for whom the decision to stay out of the work force is probably multifactorial (e.g., homemaker). Nevertheless, there were strong associations between NP testing and vocational outcomes, even after controlling for the influence of demographics, depression, and disease course. While depression and disease course also were significantly associated with employment status, the associations between NP testing and vocational outcomes remained even after accounting for the effects of these variables. As would be expected, verbal memory and reasoning tests demonstrated the strongest associations with vocational status. While the SDMT was the NP test showing the most frequent impairment in this sample, the PASAT-3 was the processing speed measure showing closest alignment with vocational disability. One explanation for this may be that the PASAT also involves numerical processing, a skill required in many occupations.

In addition to our concerns about the measurement of vocational disability in this study, there are other methodological limitations that merit comment. The study was cross-sectional in design, and whereas we can say something about the validity of these NP tests at a single point in time, their predictive validity, in a strict sense, is unknown. Quantitative assessment of neurological status using the EDSS was missing in about one-third of the sample, which means we could not control for this factor in the logistic regression analysis. However, we were able to control for disease course (RR *vs*. SP), which has been shown to be associated with EDSS (Jacobs et al., 1999). While the patients were screened for gross deficiency in upper extremity motor control, we did not administer the copy trial of the BVMT-R, and it is conceivable that mild motor dysfunction may have contributed to performance on this test. The effect size for this particular measure was large, and may have been inflated by drawing problems due purely to manual or motor dysfunction. In future work, we plan to study the contribution of motor control to BVMT-R performance using standardized tests of manual speed and dexterity. One of the advantages of these MACFIMS tests lies with their potential for using manifold variables from a single procedure. This is especially true of the CVLT-II and the DKEFS. Because the number of variables was intentionally limited to control for type 1 error, future studies might examine these tests in greater detail.

The above notwithstanding, we find support for the validity of the MACFIMS battery in MS patients. These tests demonstrate construct validity, discriminate MS patients from healthy controls and RR from SP patients, and are associated with greater risk for vocational disability. Future work will investigate the test-retest reliability of these tests in a large sample, and the longitudinal course of NP dysfunction in MS using these measures.

ACKNOWLEDGMENTS

The author acknowledges the support of the clinical and administrative staff from Jacobs Neurological Institute and Buffalo Neuroimaging Analysis Center. Some of the research was supported, in part, by an unrestricted educational grant from Biogen Idec.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*. (4th ed.) Washington, DC: American Psychiatric Association.

Archibald, C.J. & Fisk, J.D. (2000). Information processing efficiency in patients with multiple sclerosis. *Journal of Clinical & Experimental Neuropsychology*, *22*, 686–701.

Arnett, P.A., Rao, S.M., Bemardill, L., Luchetta, T., Frost, J., Meidenbauer, K., & Heise, R. (1997). Executive functions in multiple sclerosis: An analysis of temporal ordering, semantic encoding, and planning abilities. *Neuropsychology*, *11*, 535–544.

Beatty, W.W. (1996). Memory disturbance in multiple sclerosis: Reconsideration of patterns of performance on the Selective Reminding Test. *Journal of Clinical & Experimental Neuropsychology*, *18*, 56–62.

Beatty, W.W., Blanco, C.R., Wilbanks, S.L., & Paul, R.H. (1995a). Demographic, clinical, and cognitive characteristics of multiple sclerosis patients who continue to work. *Journal of Neurologic Rehabilitation*, *9*, 167–173.

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

Beatty, W.W., Goodkin, D.E., Monson, N., & Beatty, P.A. (1989). Cognitive disturbances in patients with relapsing remitting multiple sclerosis. *Archives of Neurology*, *46*, 1113–1119.

Beatty, W.W., Hames, K.A., Blanco, C.R., Paul, R.H., & Wilbanks, S.L. (1995b). Verbal abstraction deficit in multiple sclerosis. *Neuropsychology*, *9*, 198–205.

Beatty, W.W. & Monson, N. (1996). Problem solving by patients with multiple sclerosis: Comparison of performance on the Wisconsin and California Card Sorting Test. *Journal of the International Neuropsychological Society*, *2*, 134–140.

Beatty, W.W. & Monson, N. (1994). Picture and motor sequencing in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *16*, 165–172.

Beatty, W.W., Paul, R.H., Wilbanks, S.L., Hames, K.A., Blanco, C.R., & Goodkin, D.E. (1995c). Identifying multiple sclerosis patients with mild or global cognitive impairment using the screening examination for cognitive impairment (SEFCI). *Neurology*, *45*, 718–723.

Beck, A.T., Steer, R.A., & Brown, G.K. (2000). *BDI-Fast Screen for Medical Patients: Manual*. San Antonio, TX: Psychological Corporation.

Benedict, R.H.B. (1997). *Brief Visuospatial Memory Test–Revised: Professional Manual*. Odessa, Florida: Psychological Assessment Resources, Inc.

Benedict, R.H.B., Fischer, J.S., Archibald, C.J., Arnett, P.A., Beatty, W.W., Bobholz, J., Chelune, G.J., Fisk, J.D., Langdon, D.W., Caruso, L., Foley, F., LaRocca, N.G., Vowels, L., Weinstein, A., DeLuca, J., Rao, S.M., & Munschauer, F. (2002). Minimal Neuropsychological Assessment of MS Patients: A Consensus Approach. *Clinical Neuropsychologist*, *16*, 381–397.

Benedict, R.H.B., Fishman, I., McClellan, M.M., Bakshi, R., & Weinstock-Guttman, B. (2003). Validity of the Beck Depression Inventory—Fast Screen in multiple sclerosis. *Multiple Sclerosis*, *9*, 393–396.

Benedict, R.H.B., Priore, R.L., Miller, C., Munschauer, F., & Jacobs, L. (2001a). Personality disorder in multiple sclerosis correlates with cognitive impairment. *Journal of Neuropsychiatry & Clinical Neurosciences*, *13*, 70–76.

Benedict, R.H.B., Schretlen, D., Groninger, L., Dobraski, M., & Shpritz, B. (1996). Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. *Psychological Assessment*, *8*, 145–153.

Benedict, R.H.B., Shapiro, A., Priore, R.L., Miller, C., Munschauer, F.E., & Jacobs, L.D. (2001b). Neuropsychological counseling improves social behavior in cognitively-impaired multiple sclerosis patients. *Multiple Sclerosis*, *6*, 391–396.

Benedict, R.H.B., Wahlig, E., Bakshi, R., Fishman, I., Munschauer, F., Zivadinov, R., & Weinstock-Guttman, B. (2005a). Predicting quality of life in multiple sclerosis: Accounting for physical disability, fatigue, cognition, mood disorder, personality, and behavior change. *Journal of the Neurological Sciences*, *231*, 29–34.

Benedict, R.H.B., Weinstock-Guttman, B., Fishman, I., Sharma, J., Tjoa, C.W., & Bakshi, R. (2004). Prediction of neuropsychological impairment in multiple sclerosis: Comparison of conventional magnetic resonance imaging measures of atrophy and lesion burden. *Archives of Neurology*, *61*, 226–230.

Benedict, R.H.B., Zivadinov, R., Carone, D., Weinstock-Guttman, B., Gaines, J., Maggiore, C., Sharma, J., Tomassi, M., & Bakshi, R. (2005b). Regional Lobar Atrophy Predicts Memory Impairment in Multiple Sclerosis. *American Journal of Neuroradiology*, *26*, 1824Y1831.

Benton, A.L., Sivan, A.B., Hamsher, K., Varney, N.R., & Spreen, O. (1994). *Contributions to Neuropsychological Assessment*. (2nd ed.) New York: Oxford University Press.

Christodoulou, C., Krupp, L.B., Liang, Z., Huang, W., Melville, P., Roque, C., Scherl, W.F., Morgan, T.F., MacAllister, W.S., Li, L., Tudorica, L.A., Roche P., & Peyster, R. (2003). Cognitive performance and MR markers of cerebral injury in cognitively impaired MS patients. *Neurology*, *60*, 1793–1798.

Cohen, J. (1988). *Statistical Power for the Behavioral Sciences*. (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Comi, G., Filippi, M., Martinelli, V., Campi, A., Rodegher, M., Alberoni, M., Siranbian, G., & Canal, N. (1995). Brain MRI correlates of cognitive impairment in primary and secondary progressive multiple sclerosis. *Journal of the Neurological Sciences*, *132*, 222–227.

D'Esposito, M., Onishi, K., Thompson, H., Robinson, K., Armstrong, C., & Grossman, M. (1996). Working memory impairments in multiple sclerosis: Evidence from a dual-task paradigm. *Neuropsychology*, *10*, 51–56.

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001). *Delis-Kaplan Executive Function System*. San Antonio, Texas: Psychological Corporation.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (2000). *California Verbal Learning Test Manual*, Adult Version (2nd ed.). San Antonio, TX: Psychological Corporation.

DeLuca, J., Gaudino, E.A., Diamond, B.J., Christodoulou, C., & Engel, R.A. (1998). Acquisition and storage deficits in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *20*, 376–390.

Demaree, H.A., DeLuca, J., Gaudino, E.A., & Diamond, B.J. (1999). Speed of information processing as a key deficit in multiple sclerosis: Implications for rehabilitation. *Journal of Neurology, Neurosurgery & Psychiatry*, *67*, 661–663.

Fischer, J.S. (1988). Using the Wechsler Memory Scale–Revised to detect and characterize memory deficits in multiple sclerosis. *Clinical Neuropsychologist*, *2*, 149–172.

Fisk, J.D. & Archibald, C.J. (2001). Limitations of the Paced Auditory Serial Addition Test as a measure of working memory in patients with multiple sclerosis. *Journal of the International Neuropsychological Society*, *7*, 363–372.

Goverover, Y., Kalmar, J., Gaudino-Goering, E., Shawaryn, M., Moore, N.B., Halper, J. & DeLuca, J. (2006). Relation between subjective and objective measures of everyday life activities in persons with multiple sclerosis. *Archives of Physical Medicine & Rehabilitation*, *in press*.

Gronwall, D.M.A. (1977). Paced auditory serial addition task: A measure of recovery from concussion. *Perceptual and Motor Skills*, *44*, 367–373.

Heaton, R.K. (1985). Neuropsychological findings in relapsing-remitting and chronic-progressive multiple sclerosis. *Journal of Consulting & Clinical Psychology*, *53*, 103–110.

Higginson, C.I., Arnett, P.A., & Voss, W.D. (2000). The ecological validity of clinical tests of memory and attention in multiple sclerosis. *Archives of Clinical Neuropsychology*, *15*, 185–204.

Jacobs, L.D., Wende, K.E., Brownscheidle, C.M., Apatoff, B., Coyle, P.K., Goodman, A., Gottesman, M.H., Granger, C.V., Greenberg, S.J., Herbert, J., Krupp, L., Lava, N.S., Mihai, C., Miller, A.E., Perel, A., Smith, C.R., & Snyder, D.H. (1999). A profile of multiple sclerosis: The New York State Multiple Sclerosis Consortium. *Multiple Sclerosis*, *5*, 369–376.

Knight, R.G., Devereux, R.C., & Godfrey, H.P.D. (1997). Psychosocial consequences of caring for a spouse with multiple scle-

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press

rosis. *Journal of Clinical and Experimental Neuropsychology*, *19*, 7–19.

Kraus, J.A., Schutze, C., Brokate, B., Kroger, B., Schwendemann, G., & Hildebrandt, H. (2005). Discriminant analysis of the cognitive performance profile of MS patients differentiates their clinical course. *Journal of Neurology*, *252*, 808–813.

Krupp, L.B., Sliwinski, M., Masur, D.M., Friedberg, F., & Coyle, P.K. (1994). Cognitive functioning and depression in patients with chronic fatigue syndrome and multiple sclerosis. *Archives of Neurology*, *51*, 705–710.

Kurtzke, J.F. (1983). Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Annals of Neurology*, *13*, 227–231.

Lockwood, A.H., Linn, R.T., Szymanski, H., Coad, M.L., & Wack, D.S. (2004). Mapping the neural systems that mediate the Paced Auditory Serial Addition Task (PASAT). *Journal of the International Neuropsychological Society*, *10*, 26–34.

McDonald, W.I., Compston, A., Edan, G., Goodkin, D.E., Hartung, H., Lublin, F., McFarland, H.F., Paty, D.W., Polman, C.H., Reingold, S.C., Sandberg-Wollheim, M., Sibley, W. A., Thompson, A., van der Noort, S., Weinshenker, B.Y., & Wolinsky, J. S. (2001). Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis. *Annals of Neurology*, *50*, 121–127.

Rao, S.M. (1991). *A Manual for the Brief, Repeatable Battery of Neuropsychological Tests in Multiple Sclerosis*. New York, NY: National Multiple Sclerosis Society.

Rao, S.M., Leo, G.J., Ellington, L., Nauertz, T., Bernardin, L., & Unveragt, F. (1991a). Cognitive dysfunction in multiple sclerosis. II. Impact on employment and social functioning. *Neurology*, *41*, 692–696.

Rao, S.M., Leo, G.J., & St. Aubin-Farbert, P. (1989a). Information processing speed in patients with multiple sclerosis. *Journal of Clinical & Experimental Neuropsychology*, *11*, 471–477.

Rao, S.M., Leo, G.J., Bernardin, L., & Unverzagt, F. (1991b). Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology*, *41*, 685–691.

Rao, S.M., Leo, G.J., Haughton, V.M., Aubin-Faubert, P.S., & Bernardin, L. (1989b). Correlation of magnetic resonance imaging with neuropsychological testing in multiple sclerosis. *Neurology*, *39*, 161–166.

Ryan, L., Clark, C.M., Klonoff, H., Li, D., & Paty, D. (1996). Patterns of cognitive impairment in relapsing-remitting multiple sclerosis and their relationship to neuropathology on magnetic resonance images. *Neuropsychology*, *10*, 176–193.

Scarrabelotti, M. & Carroll, M. (1998). Awareness of remembering achieved through automatic and conscious processes in multiple sclerosis. *Brain & Cognition*, *38*, 183–201.

Smith, A. (1982). *Symbol digit modalities test: Manual*. Los Angeles: Western Psychological Services.

Sperling, R.A., Guttmann, C.R., Hohol, M.J., Warfield, S.K., Jakab, M., Parente, M., Diamond, E.L., Daffner, K.R., Olek, M.J., Orav, E.J., Kikinis, R., Jolesz, F.A., & Weiner, H.L. (2001). Regional magnetic resonance imaging lesion burden and cognitive function in multiple sclerosis: A longitudinal study. *Archives of Neurology*, *58*, 115–121.

Thornton, AE., Raz, N., & Tucke, K.A. (2002). Memory in multiple sclerosis: contextual encoding deficits. *Journal of the International Neuropsychological Society*, *8*, 395–409.

Vleugels, L., Lafosse, C., van Nunen, A., Nachtergaele, S., Ketelaer, P., Charlier, M., & Vandenbussche, E. (2000). Visuoperceptual impairment in multiple sclerosis patients diagnosed with neuropsychological tasks. *Multiple Sclerosis*, *6*, 241–254.

https://doi.org/10.1017/S1355617706060723 Published online by Cambridge University Press