Journal of the International Neuropsychological Society (2021), **27**, 124–135

doi:10.1017/S1355617720000776

A Multidimensional Assessment of Metacognition Across Domains in Multiple Sclerosis

Audrey Mazancieux[1,*], Chris J.A. Moulin[1,2], Olivier Casez[3] and Céline Souchay[1]

[1]LPNC CNRS 5105, Université Grenoble Alpes, Grenoble, France
[2]Institut Universitaire de France, Paris, France
[3]Department of Neurology, Université Grenoble Alpes, Grenoble, France

(Received February 21, 2020; Final revision June 29, 2020; Accepted July 2, 2020; First Published Online August 10, 2020)

Abstract

Objective: In neurological diseases, metacognitive judgements have been widely used in order to assess the degree of disease awareness. However, as yet little research of this type has focused on multiple sclerosis (MS). **Method:** We here focused on an investigation of item-by-item metacognitive predictions (using feeling-of-knowing judgements) in episodic and semantic memory and global metacognitive predictions in standard neuropsychological tests pertinent to MS (processing speed and verbal fluency). Twenty-seven relapsing–remitting MS (RR-MS) patients and 27 comparison participants took part. **Results:** We found that RR-MS patients were as accurate as the group of comparison participants on our episodic and semantic item-by-item judgements. However, for the global predictions, we found that the MS group initially overestimated their performance ($d_s = .64$), but only on a task on which performance was also impaired ($d_s = .89$; processing speed). We suggest that MS patients, under certain conditions, show inaccurate metacognitive knowledge. However, postdictions and item-by-item predictions indicate that online metacognitive processes are no different from participants without MS. **Conclusion:** We conclude that there is no monitoring deficit in RR-MS and as such these patients should benefit from adaptive strategies and symptom education.

Keywords: Multiple sclerosis, Metacognition, Self-awareness, Metamemory, Anosognosia

In the context of pathology, awareness is critical for patient care. Being aware of cognitive or physical impairments is crucial for both the efficacy of cognitive rehabilitation programmes (Prigatano, 1999) and the understanding of the impact of cognitive disabilities on activities of daily living (McGlynn & Schacter, 1989). The focus of this paper is multiple sclerosis (MS). Whilst a considerable number of studies have examined the question of disease awareness in MS (for a review, see Mazancieux, Souchay, Casez & Moulin, 2019), most research has considered self-report and questionnaire measures. In this article, we invoke the metacognition framework to consider disease awareness in MS. Metacognition broadly refers to the knowledge of, the monitoring of (self-evaluation), and the control of (strategy implementation) cognitive activity (Nelson & Narens, 1990). It allows the evaluation of awareness in asking patients to make metacognitive judgements. These judgements refer to a self-assessment of performance on a particular cognitive task.

Although metacognition has been widely evaluated in different neurological and psychiatric diseases (e.g., Pannu & Kaszniak, 2005), there are surprisingly few studies focusing on the evaluation of metacognition in MS despite the high incidence of this pathology. MS is an autoimmune inflammatory disease characterised by lesions which can appear across the whole central nervous system. These lesions produce a neural and neuronal demyelination which compromises the conduction of information (Trapp & Nave, 2008). In addition to physical disabilities, cognitive impairment is also frequent in MS with prevalence rates ranging from 43% to 70% (Chiaravalloti & DeLuca, 2008). Although cognitive symptoms vary in MS, a common profile emerges where the majority of these symptoms are related to an executive functioning deficit as a potential consequence of processing speed impairments (Drew, Tippett, Starkey, & Isler, 2008). As has been shown in other pathologies (see Souchay, 2007 for a review in Alzheimer's disease), in traumatic brain injury (Ciurli et al., 2010), or in healthy ageing (e.g., Souchay & Isingrini, 2004), impaired performance of executive function

*Correspondence and reprint requests to: Audrey Mazancieux: 1251 avenue Centrale, St Martin d'Hères, 38040 Grenoble France. Email: audrey.mazancieux@univ-grenoble-alpes.fr

tests are associated with metacognitive difficulties. For instance, Souchay, Isingrini, Clarys, Taconnat, & Eustache (2004) have found that performance on the Wisconsin Sorting Card Test (i.e., a measure of executive function particularly used in MS) positively correlates with metamemory measures. This leads to the expectation that MS patients might exhibit metacognitive impairment. Perhaps more importantly, from a clinical viewpoint, assessing the level of symptom awareness is crucial in order to help patients to use pertinent strategies when dealing with their symptoms. Apart from executive function-related symptoms (including planning, flexibility, inhibition, working memory; Rabbitt, 2004) and processing speed, impairments in long-term memory in verbal and visual modalities have been reported (e.g., Calabrese, 2006; Ruet, 2015).

The majority of studies investigating awareness in MS have focused on a comparison of self-evaluations of cognitive functioning (mainly by questionnaires) with more objective neuropsychological evaluations (e.g., Maor, Olmer, & Mozes, 2001; Randolph, Arnett, & Freske, 2004; Roberg, Bruce, Lovelace, & Lynch, 2012). In a recent review of the scant literature on metacognition in MS (Mazancieux et al., 2019), we suggested a non-linear relationship between the subjective evaluation of cognitive impairment and a more objective evaluation (i.e., neuropsychological assessment). Patients with a slight decline in their cognitive abilities tend to underestimate their abilities, whereas patients with more cognitive impairment tend to overestimate them. This failure in self-evaluation is also associated with emotional disturbances and fatigue which are prevalent in MS patients (Kesselring & Klement, 2001). For instance, it has been shown that depression is associated with metacognitive inaccuracy in MS (Kinsinger, Lattie, & Mohr, 2010). When cognitive impairment awareness is measured using a correlation between subjective patient evaluation and neuropsychological tests, these correlations are weaker in depressed MS populations suggesting lower symptom awareness (in the context of memory tests; Phillips & Stuifbergen, 2006). Beliefs about cognitive functioning [referred to here as metacognitive knowledge (Flavell, 1979), and which are easily operationalised in questionnaire studies] have been the most evaluated metacognitive construct in MS. However, since methodological issues arise from the comparison between subjective evaluation via questionnaire and neuropsychological evaluation, a more reliable way to measure metacognitive knowledge is the use of global predictions, as used in the current paper.

In the global prediction paradigm, participants are asked to predict their performance in a cognitive task. For instance on an episodic memory task, they predict the number of items they think they will be able to recall from a list. By comparing the prediction to the actual performance (i.e., the number of recalled items), it is possible to have an idea of the awareness of the cognitive function. When the prediction is made before the task, it allows an estimation of or metacognitive knowledge (generalised beliefs about the task which might include lay understandings of ageing or the disease process; Hertzog, 1992). Measured after the task, 'postdictions' evaluate metacognitive experience, especially monitoring processes referring to the update of self-evaluation derived from online monitoring of the ongoing task (Connor, Dunlosky & Hertzog, 1997).

A more fine-grained analysis of monitoring is achieved by asking participants to make item-by-item judgements. In MS, Beatty and Monson (1991) asked patients and non-MS participants to perform item-by-item feeling-of-knowing (FOK) judgements, where participants have to predict their future ability to recognise an item that they have failed to recall. Two versions of the task exist. In the episodic FOK (eFOK) task, participants first learn paired-words and then have to recall the target from a presented cue. If they are unable to recall the target word, they report their likelihood of recognising it in a list of words. This judgement is the FOK. In the semantic FOK (sFOK) task, participants answer general knowledge questions. As in the eFOK task, they have to say if they think they will be able to recognise the answer if they are not able to recall it. From these FOKs, it is possible to examine metacognitive biases (the over- and underestimation of performance) and metacognitive sensitivity (the discrimination between correct and incorrect recognition). In the only study in MS, patients exhibited poor metacognitive sensitivity in the eFOK task (Beatty & Monson, 1991). However, this study is inconclusive for several reasons. First, the authors did not distinguish between different forms of MS. Second, alternative – more reliable – measures of metacognitive sensitivity have subsequently been developed. Third, the examination of metacognition was limited to memory tasks. The current study aimed to address these shortcomings.

The present study aims to further examine metacognitive functioning in people with relapsing–remitting MS (RR-MS), the most common form of MS (80% of patients; Rao et al., 1991). The present study proposes a general overview of metacognition in MS with RR-MS patients, since this is the most common form (80% of patients; Rao et al., 1991). First, we decided to measure eFOK and sFOK, a common strategy for exploring metacognition in cognitive impaired groups (e.g., Alzheimer's disease, Souchay, 2007; patients with focal frontal lesions, Schnyer et al., 2004; Korsakoff's syndrome, Shimamura & Squire, 1986; and autism spectrum disorders, Wojcik, Moulin, & Souchay, 2013). A typical profile is of impaired eFOK accuracy when patients exhibit episodic memory impairment. On the contrary, sFOK accuracy is preserved in these studies. There is an overwhelming bias for measuring metacognition through memory tasks in MS (Mazancieux et al., 2019). As such, even though eFOK and sFOK tasks are robust and often used as measures of metacognition, memory function may not be the most pertinent task on which to test the metacognition of people with MS.

We assume that focusing on more relevant functions would allow a more complete picture of awareness in this pathology. From a clinical point of view, we assume that measuring awareness of a cognitive activity is especially

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

Table 1. Means and standard deviation for demographic and clinical data for the MS patients and the healthy control group.

	MS patients	Comparison group	Cohen's *d*
N	27	27	
Age in years	39.48 (9.93)	39.03 (10.80)	0.04
Education in years	14.04 (2.08)	14.56 (2.03)	0.25
EDSS	2.56 (1.93)	n.a	
Disease duration in years	6.96 (3.23)	n.a	

EDSS: the Expanded Disability Status Scale.

relevant when there is a specific impairment in this cognitive activity. Therefore, we also adopted a procedure where participants can make metacognitive judgements about standard neuropsychological tasks where MS patients are often impaired: the Symbol Digit Modalities Test (SDMT) and the conceptual verbal fluency task (Planche, Gibelin, Cregut, Pereira, & Clavelou, 2016; Ruet, 2015).

The SDMT is a processing speed task where first an association of symbols with digits is provided. In the test phase, only the symbols are presented, and participants have to say the digit associated with each symbol as rapidly as possible. As slowing is the main cognitive impairment in MS, patients often exhibit a deficit in this task. In the conceptual fluency task, participants have to generate as many words as possible in a given time from a semantic category (e.g., animals). Similarly, MS patients often show significant impairments (slowing) on this task where self-initiated processes and strategic search in memory are involved. In order to assess awareness of these cognitive abilities, we added metacognitive judgements to these two tasks focusing on global predictions to measure both metacognitive knowledge and metacognitive monitoring.

Our rationale was to have a protocol that mixed very commonly used monitoring tasks (FOKs) and global predictions on tasks that are pertinent for MS. In particular, we proposed metacognitive judgements on neuropsychological tests that are particularly used in this population. The SDMT is one of the most used tests in MS (e.g., Planche et al., 2016; Walker et al., 2016; O'Brien et al., 2007; Basso et al., 2008; Ruet, 2015). Regarding verbal fluency, several studies have found that it is also a good predictor of RR-MS severity (Prakash, Snook, Lewis, Motl, & Kramer, 2008), and a selective impairment of semantic fluency in RR-MS has been shown (despite a preserved phonemic fluency, Santiago, Guardia, Casado, Carmona & Arbizu, 2007). Thus, from a neuropsychological viewpoint, these are tasks where we may expect to find deficits, and as such examining metacognitive awareness in these tasks would be of critical interest, even though these are less typically studied in a metacognitive context.

In sum, there is very little existing research into metacognition in this population, and existing works focus mainly on memory function with varying disease types. This study aims to explore more precisely metacognitive processes in MS that are likely to be impaired due to the neuropsychological profile with executive deficits in this population.

METHOD

Participants

Twenty-seven patients (21 female, 6 male; $M_{age} = 39.48$, $SD_{age} = 9.93$) were included in the study. The diagnosis of MS was established by a neurologist who also informed the patients about the study. Inclusion criteria were to have no recent exacerbation of MS symptoms and no other neurological disease. Participants were excluded if they had a form of MS other than RR-MS or a history of alcohol or drug abuse. Twenty-seven non-MS volunteer participants (21 female, 6 male; $M_{age} = 39.03$, $SD_{age} = 10.80$) also took part in the study as a healthy control group. People in the healthy control group voluntarily chose to participate to the study without being paid for their participation. Information about the study was given in the hospital where patients were tested and in Grenoble Alpes University. This advertisement targeted the general public, but no patient family member was recruited to the control group. Only people with no history of neurological disease, psychiatric disease, or alcohol or drug abuse were included in the control group. Patients and healthy control group participants were matched one by one in terms of gender, age (+/− 5 years), and years of education (+/− 3 years). Demographic and clinical data are summarised in Table 1.

Participants were tested either in the Laboratoire de Psychology et Neurocognition (LPNC) or in the Centre Hospitalier Universitaire in Grenoble. The study was approved by the Ethics Committee for Non-Interventional Research of Grenoble. All data included in this manuscript were obtained in compliance with the Helsinki Declaration.

Material and Procedure

All participants were tested individually in one 60- to 75-min session. The whole procedure included two metacognition tasks: global predictions and the eFOK and sFOK tasks, as well as the completion of two questionnaires. Global prediction and FOK task order were randomly assigned for each participant.

Global Prediction

Participants performed two neuropsychological tasks: a verbal fluency task and the SDMT (oral version). The standard tasks were slightly modified in order to assess and compare metacognitive awareness across tasks. Participants had 45 s (instead of 120 s) to give as many numbers as possible for the SDMT task. Two versions were created in order to have two trials (see Figure 1). For the verbal fluency task, participants again had two trials and had 45 s

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

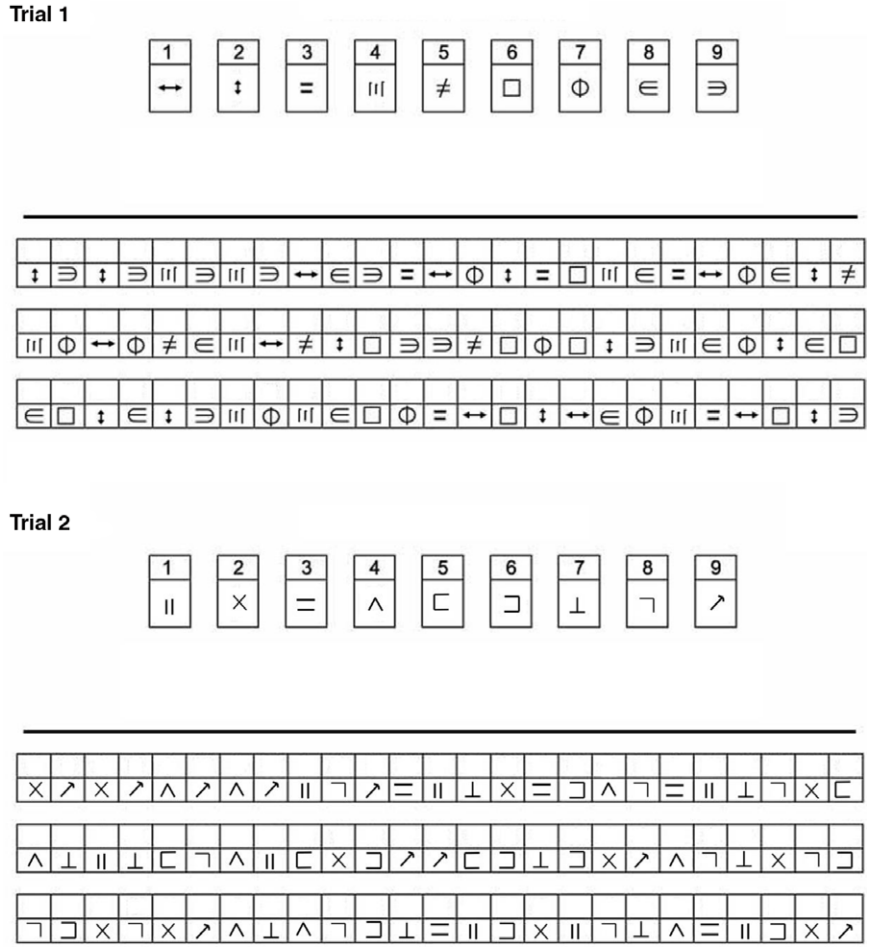

Figure 1. The two trials of our version of the SMDT tasks. Participants have to read aloud digits that correspond to the presented symbols as rapidly as possible. They have 45 sto read as many digits as they can.

(instead of 120 s) to give words either from the category 'animals' or 'fruits and vegetables'. After the task was explained to the participants, they were asked to predict the score they would achieve. For the fluency task, participants were asked 'how many words from the category do you think you will generate in 45s?' For the SDMT task, participants were asked 'how many numbers do you think you will read in 45s?' These predictions were made once before the task was performed (prediction) and for a second time after completion (postdiction). For the postdiction, participants were asked to estimate their prior performance on the same basis (number of items achieved). There were two trials per task, which enables the examination of the ability to integrate feedback from having completed the task into the predictions for a second trial. Therefore, for each task, participants performed an initial prediction of performance, then conducted the task, and following the task, made a postdiction. Then, they had to make a second prediction, complete a different version of the task, and make a second postdiction. Trial order (version 1 and 2 for the SDMT and animal category or fruit and vegetable category for the verbal fluency task) was randomly assigned for each participant.

FOK Tasks

The material used for the eFOK and sFOK tasks was similar to those used by Souchay, Moulin, Clarys, Taconnat and Isingrini (2007). These materials allow some control of difficulty between the episodic memory task and the semantic memory task since the same target word is used in both tasks. Each target has a definition used in the sFOK task and an associative cue used in the eFOK task. All the targets were divided into two lists so that each participant would not have the same target word in both tasks. Half of the participants had the first list for the episodic task and the second list for the semantic task, with the other half having the reverse pattern.

The eFOK task included three stages: encoding, cued recall, and recognition. Participants firstly attempted to learn 40 paired-words with the first word written in uppercase and the second written in lowercase. Each word pair was presented for 5 s. During the recall stage, only the cue (i.e., the word written in lowercase) was presented and the participant was asked to retrieve the associated target word (i.e., the uppercase word) with 15 s to do so. After this time

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

passed, they had to give a FOK judgement, reporting whether they thought they would recognise the correct target amongst a five-word list. As in Souchay and colleagues (2007), the FOK decisions were in a 'yes' or 'no' format. No feedback about the correctness of the recall was given to the participants, and FOK judgements were made for all items. After the recall stage for all cues had been completed, participants performed a five-alternative forced choice recognition task. The 40 cues were presented again and the participants had to find the correct associated target with the presented cue. There was no time limit for this stage.

The sFOK task included recall and recognition phases. First of all, participants attempted recall for 40 general information questions. As in the eFOK task, they had 15 s to respond and then made an FOK judgement win the same manner as the eFOK procedure. After this, they performed a recognition task, where participants were again presented the 40 general information questions with five alternative responses. The two tasks were constructed using E-prime software and were presented to the participants on a 15.6 inch computer screen. Half of the participants started with eFOKs and half with sFOKs.

Emotional and Fatigue Assessment

Both patients and the healthy control group completed two questionnaires at the end of the testing session. The first one was the Beck Depression Inventory II (BDI-II; Beck, Steer, & Brown, 1996) and the second was the Fatigue Impact Scale (FIS; Fisk, Ritvo, Ross, Haase, Marrie, & Schlech, 1994). For participants who reported being too tired by the experimental procedure, questionnaires were sent by e-mail and were completed within 1 week.

Statistical Analyses

Analyses were conducted using R software. Data and analysis scripts are available on Open Science Framework (https://osf.io/fyshb/files/). The main interest in metacognition is the accuracy of the judgements, that is, the comparison between the judgement and the performance. Regarding global predictions, we first focused on the magnitude of predictions as simply the number of items participants predict. Then, we calculated accuracy scores in terms of the relation between predicted and actual performance. This score is non-directional meaning that it allows an estimate of how precise are participants without being influenced by metacognitive bias (underestimation or overestimation of performance; see Moulin, Perfect, & Jones, 2000). These two measures capture different aspects of metacognition: someone can consistently overestimate their performance but yet be relatively accurate with a small discrepancy between their prediction and the score. Because we expect differences in terms of task performance and because these differences might influence accuracy scores, the prediction was transposed into a percentage of performance. More precisely, each prediction was expressed in a proportion of performance for each trial using the following formula: Prediction * 100/Performance (e.g., a participant with task performance of 30 and prediction of 10 would have predicted 1/3 of their performance having therefore a percentage of performance of 33%. A participant with task performance of 20 and prediction of 40 would have predicted 150% of his or her performance). To control for bias, the non-directional difference between this score and performance (that refers to 100% in this context) was computed. Therefore, an accuracy score of 0 suggests that the participant has a perfect accuracy, and an accuracy score of 10 refers to a deviation of 10% from performance. Due to recording issues, one patient did not have prediction and postdiction scores for the fluency task.

To avoid effects of potential outliers which might be found in patients who have by definition a non-normal behaviour, we used linear mixed-effect models computed using 'lmerTest' and 'lme4' packages (Kuznetsova, Brockhoff, & Christensen, 2017). Therefore, we estimated for each model an intercept per participant as a random effect. These effects are not the main focus of this paper, therefore we only reported fixed effects. As there is no consensus regarding the calculation of effect size for mixed-effects models, especially when several variables are included in the model (Westfall, Kenny, & Judd, 2014), we decided to calculate Cohen's d from the t value as is done for regular t tests (Lakens, 2013). When the effect included the between-subject group comparison the d_s value was calculated, and we used the d_z formula in cases where the effect included only within-subject variables (Lakens, 2013).

For the FOK tasks, we focused on both metacognitive bias and metacognitive sensitivity. Metacognitive sensitivity was estimated by two different approaches. First, we calculated the Type 2 d' (Higham, 2011; Nelson, 1984) as follows: Type 2 $d' = z(H2) - z(FA2)$ where z is the inverse of the cumulative normal distribution function. Here, H2 refers to Type 2 hits which are the proportion of reported 'yes' FOKs for correct responses and FA2 refers to Type 2 false alarms which are the proportion of reported 'yes' FOKs for incorrect responses. When H2 and the FA2 rates were equal to either 1 or 0, we used standard corrections (Green & Swets, 1966), using 1/(2NC) instead of a rate of 0 and 1-1/(2NI) instead of a rate of 1 (where NC is the number of correct responses and NI the number of incorrect responses). However, because Type 2 d' is influenced by metacognitive bias (see Fleming & Lau, 2014), we also computed mixed-effects logistic regressions between task performance (correct and incorrect responses) and FOK (yes and no). The difference (i.e., the slope) between yes and no FOK allows the estimation of the capacity to judge future recognition according to task performance. Therefore, the larger the difference is, the higher the discrimination between correct and incorrect responses in the recognition task. Moreover, this mixed-effect model effect allows the estimation of an intercept and a slope for FOK per participant as a random effect controlling for cross-participants variability. Finally, we calculated the percentage of correct answers for the 'yes' FOKs

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

Table 2. Means and standard deviations for predictions, performance, and postdictions (in number of items) according to groups, trials, and tasks.

	SDMT task		Fluency task	
	MS patients	Healthy control group	MS patients	Healthy control group
Trial 1				
Prediction	30.07 (12.81)	30.48 (12.16)	21.70 (8.50)	22.15 (6.67)
Performance	28.78 (7.76)	33.74 (6.62)	21.04 (5.04)	23.04 (3.69)
Postdiction	25.41 (9.00)	30.89 (10.74)	20.65 (7.29)	22.56 (6.64)
Trial 2				
Prediction	21.81 (6.20)	28.41 (11.33)	18.22 (6.25)	19.96 (6.03)
Performance	27.41 (6.39)	33.74 (5.80)	20.59 (5.83)	21.00 (6.97)
Postdiction	24.78 (6.47)	31.44 (9.08)	18.88 (7.09)	20.44 (7.96)

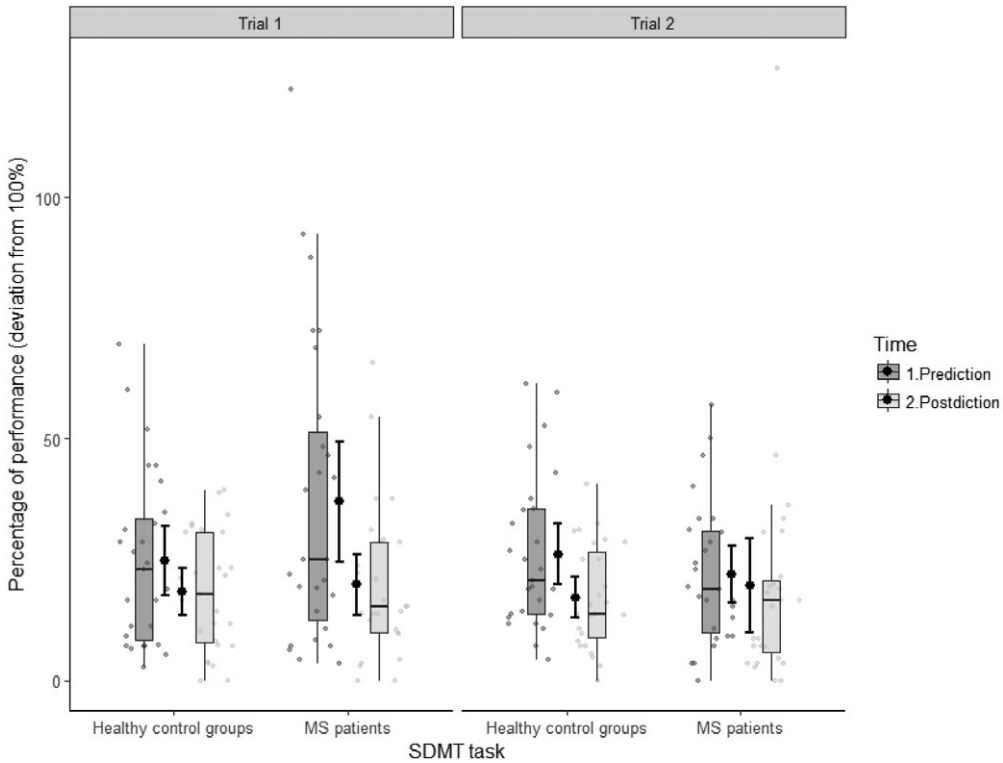

Figure 2. Means and standard errors for predictions and postdiction metacognitive accuracy scores according to groups and trials for the SDMT task.

for each participant in order to estimate bias in metamemory judgements. Other analyses were standard *t* tests.

RESULTS

Global Predictions

Task Performance

Analyses of task performance for the SDMT task showed a main effect of group, $t(52) = 3.27$, $p = .002$, $d_s = 0.89$, with patients having a lower score. There was neither an effect of trial, $t(52) = 1.21$, $p = .232$, nor an interaction between the two factors, $t(52) = -1.21$, $p = .232$. Regarding the fluency task, we found no main effect of group, $t(52) = 0.94$, $p = .352$, no effect of trial, $t(52) = 1.59$, $p = .119$, and no interaction, $t(52) = 1.02$, $p = .313$ (see Table 2).

Magnitude of Predictions

Magnitudes of raw predictions were compared according to group, trial, and judgement type (prediction *vs*. postdiction) for each task (Figures 2 and 3). For the SDMT task, the analyses revealed a main effect of group, $t(52) = 2.04$,

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

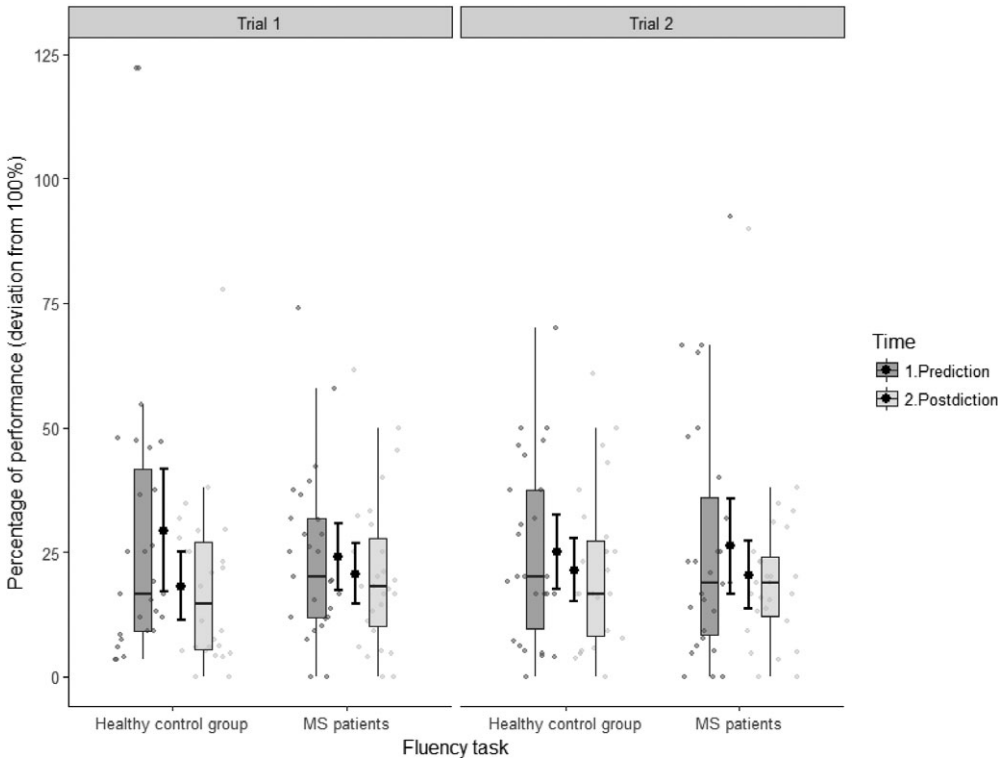

Figure 3. Means and standard errors for predictions and postdiction metacognitive accuracy scores according to groups and trials for the fluency task.

$p = .047$, $d_s = 0.56$, with patients overall predicting completing fewer items compared to the healthy control group. There was a main effect of trial, $t(156) = 3.29$, $p = .001$, $d_z = 0.45$, with a higher prediction (i.e., more items) for the first trial compared to the second trial. We also found a significant interaction between group and trial, $t(156) = -2.33$, $p = .021$, $d_s = 0.63$. Irrespective of judgement type, MS patients have lower predictions compared to the healthy control group for the second trial, $t(64.15) = 2.67$, $p = .010$, $d_z = 0.73$, but not for the first trial, $t(64.15) = 1.19$, $p = .240$. Finally, the analyses revealed an interaction between trial and judgement type, $t(156) = 3.25$, $p = .001$, $d_s = 0.44$. Irrespective of groups, participants have a trend for lower postdictions compared to predictions in the first trial, $t(156) = 1.91$, $p = .059$, and have the opposite pattern of results in the second trial, $t(156) = -2.68$, $p = .008$, $d_z = 0.37$. Regarding the fluency task, we found a main effect of trial, $t(145.09) = 3.80$, $p < .001$, $d_z = 0.52$, with a larger prediction for the first trial compared to the second trial. No other effect was significant.

Metacognitive Accuracy

We calculated accuracy scores as outlined above which were compared according to group, trial, and judgement type (prediction *vs*. postdiction) for each task (Figures 2 and 3). For the SDMT task, the analyses revealed a main effect of judgement type, $t(156) = 3.90$, $p < .001$, $d_z = 0.53$, predictions being less accurate than postdictions. No other main effects or interactions were significant but we found a trend for a three-way interaction, $t(156) = 1.92$, $p = .056$, $d_s = 0.52$. Therefore, we compared the interaction between group and trial for prediction on the one hand and postdiction on the second hand. Although we found no effect for postdiction, $t(156) = -0.15$, $p = .884$, predictions showed a significant interaction between group and trial, $t(156) = 2.58$, $p = .011$, $d_s = 0.70$. Critically, patients were less accurate at predicting their performance than the healthy control group for the first trial, $t(156) = 2.37$, $p = .019$, $d_s = 0.64$, which was not the case for the second trial, $t(156) = -0.79$, $p = .433$. Regarding the fluency task[1], the analyses revealed only a main effect of judgement type, $t(155.66) = 2.45$, $p = .015$, $d_z = 0.33$, predictions being less accurate than postdictions.

FOK Tasks

Recall and Recognition

The percentage of correct recall and correct recognition were calculated for each task and each participant. No difference between MS patients and the healthy control group was found for recall either in the episodic memory task or the semantic memory task. The same result was found for the recognition performance (see Table 3).

[1]For this analysis, we excluded one prediction of a comparison participant in the second trial which was extremely inaccurate (deviation of 328%). Running the same analysis but leaving in this participant did not change the pattern of significant results.

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

Table 3. Means and standard deviations for proportion of correct recall and recognition according to group and memory task.

	MS patients	Healthy control group	t(52) value	p-Value
Episodic memory				
Recall	0.37 (0.19)	0.38 (0.20)	0.28	0.784
Recognition	0.85 (0.11)	0.86 (0.10)	0.23	0.816
Semantic memory				
Recall	0.43 (0.17)	0.45 (0.23)	0.29	0.775
Recognition	0.75 (0.11)	0.77 (0.14)	−0.56	0.581

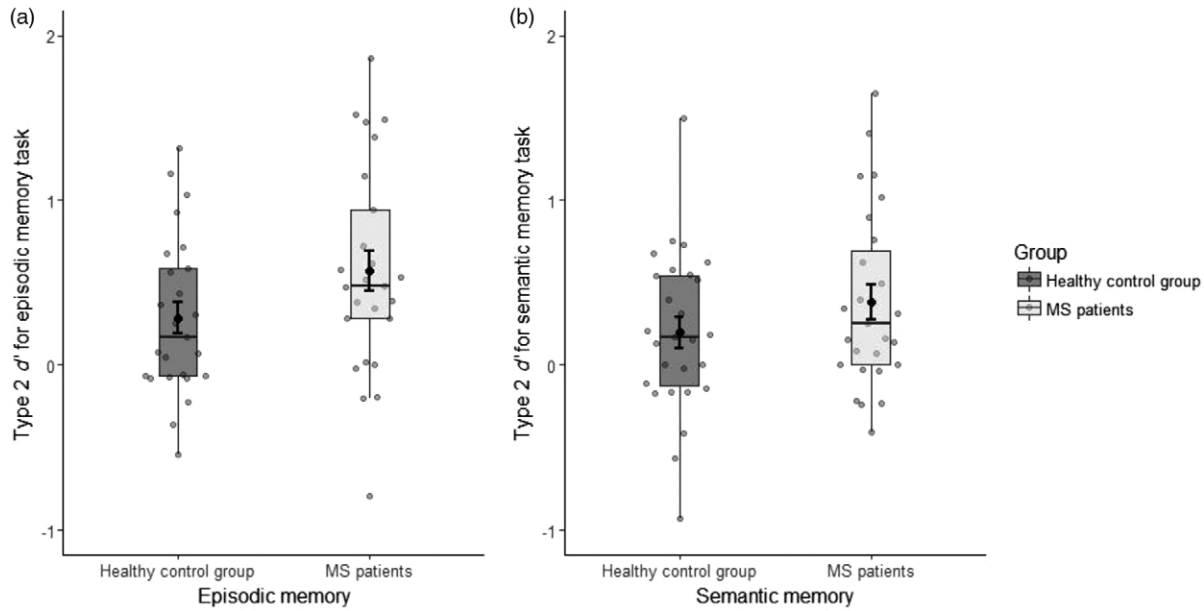

Figure 4. Individual values, means, and standard errors for metacognitive sensitivity measured by Type 2 *d*' according to group for the episodic memory task (A) and semantic memory task (B).

Metacognitive Sensitivity

A Type 2 *d'* was calculated for each participant and each task (Figure 4). For the episodic memory task, four participants had a performance rate of 1 so they were excluded from the following analysis. Overall, participants had a Type 2 *d*' significantly different from 0 for both the sFOK task, $t(52) = 4.15$, $p < .001$, and the eFOK task, $t(52) = 5.35$, $p < .001$. There were no differences between metacognitive sensitivity between MS patients and non-MS participants for both the sFOK task, $t(52) = -1.31$, $p = .195$, and the eFOK task, $t(52) = -1.78$, $p = .082$.

Moreover, we fitted two mixed-effects logistic regressions on sensitivity, with FOKs and group as fixed effects (Figure 5). We estimated an intercept and a slope for FOKs by participants as random effects. For the episodic memory task, the model showed a significant relationship between task accuracy and FOKs (estimate = 0.48, $Z = 2.85$, $p = .004$) revealing that participants were able to predict correctly their memory performance. This relationship was not different according to group (estimate = −0.51, $Z = -1.66$, $p = .097$), MS patients being as accurate as healthy control group participants. For the semantic memory task, the model only showed a trend between task performance and FOKs (estimate = 0.33, $Z = 1.83$, $p = .067$). This relationship was not different according to group (estimate = −0.45, $Z = -1.32$, $p = .188$), MS patients being as accurate as non-MS participants.

Metacognitive Bias

Metacognitive bias was estimated by calculating the percentage correct responses in the recognition task for the 'yes' FOKs for each participants and each task (Figure 6). For the episodic memory task, the analysis showed no effect of group, $t(52) = -0.56$, $p = .578$, as well as for the semantic memory task, $t(52) = 0.51$, $p = .614$. MS patients and the healthy control group have therefore the same tendency to report 'yes' FOK for correct responses in the recognition task.

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

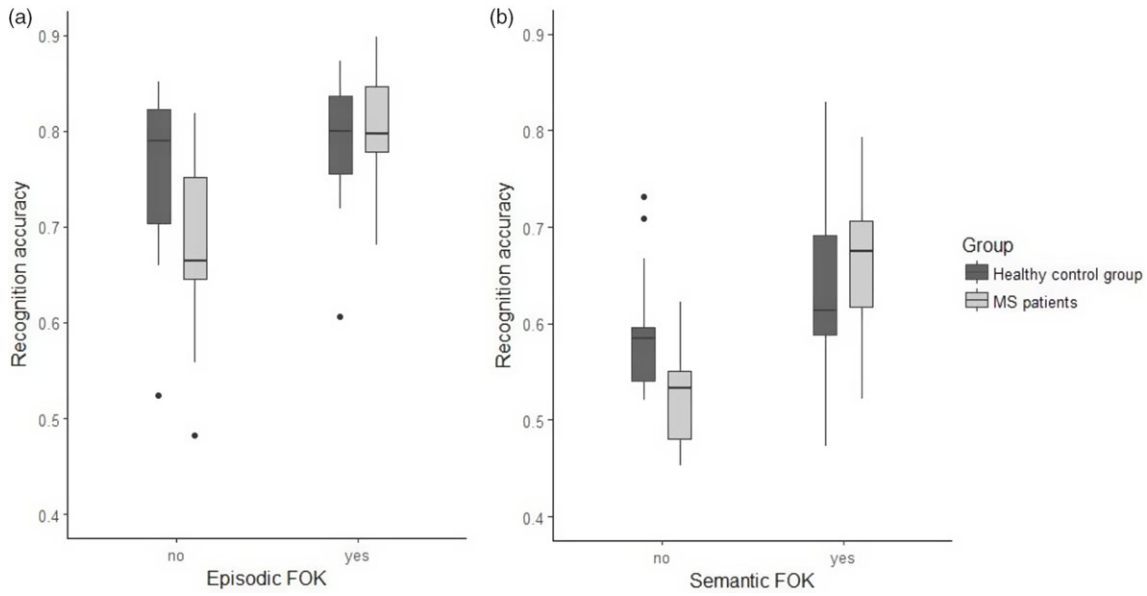

Figure 5. Boxplots for the mixed logistic regressions between task accuracy in the recognition tasks and confidence in MS patients and healthy control group participants for the episodic memory task (A) and semantic memory task (B).

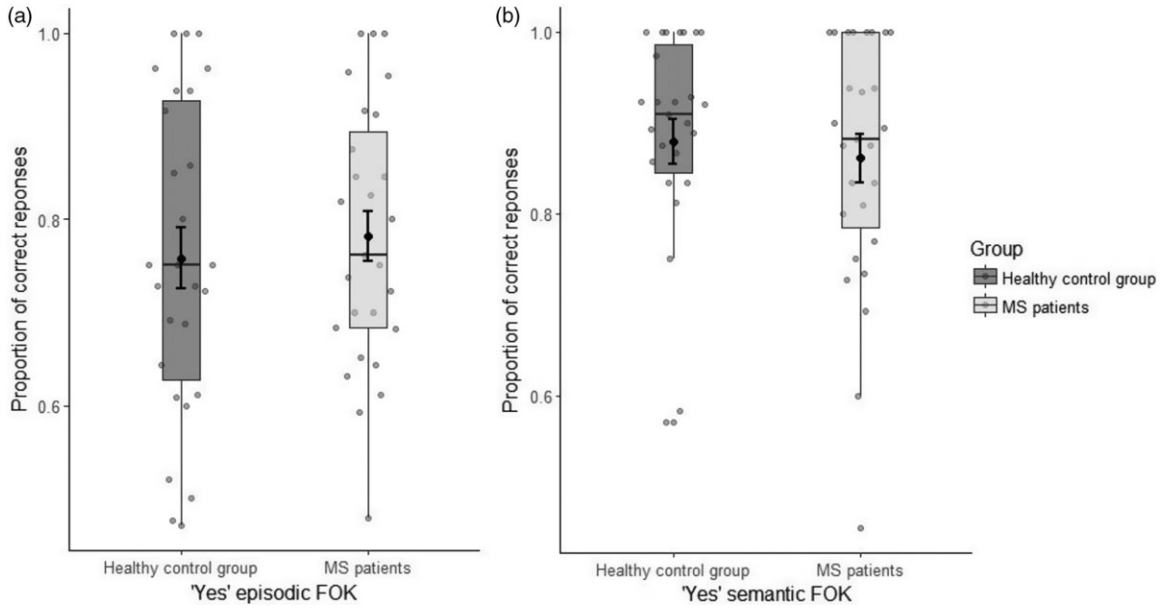

Figure 6. Individual values, means, and standard errors for metacognitive bias (proportion of correct responses for 'yes' FOK) according to group for the episodic and semantic memory tasks.

Relationship Between Metacognition and Others Variables

We compared scores to the FIS and the BDI between MS patients and healthy control group participants. MS patients had a higher score than the healthy control group on the BDI ($M_{patients} = 13.63$, $SD_{patients} = 9.63$; $M_{comparison} = 6.70$, $SD_{comparison} = 5.02$), $t(52) = 3.32$, $p = .002$, $d_s = 0.90$, and on the FIS ($M_{patients} = 75.33$, $SD_{patients} = 27.57$; $M_{comparison} = 54.22$, $SD_{comparison} = 33.22$), $t(52) = 2.54$, $p = .014$, $d_s = 0.69$. To investigate the relationship between metacognitive sensitivity and emotional and fatigue variables, we performed correlational analyses with patients. No Type 2 *d'* values correlated with the BDI scale, the FIS scale, or the Expanded Disability Status Scale. Recall did not correlate with any of these individual difference variables either. Finally, as the first prediction for the SDMT was impaired in patients, we explored the relationship between this score and depression, fatigue, and disease duration. No correlation reached significance.

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

DISCUSSION

The current study proposes a multidimensional assessment of metacognition in RR-MS patients. We used global predictions and item-by-item predictions to measure both metacognitive knowledge and metacognitive monitoring. The novelty of this experiment was to measure metacognition on tasks which are relevant in MS (the SDMT and the verbal fluency) as well as typical metacognitive tasks (FOKs). Patients showed only significant impairment for the SDMT task which is consistent with the fact that processing speed is one of the main cognitive impairments in MS (Planche et al., 2016).

Regarding global predictions, predictions before the tasks were less accurate than postdictions in both groups and for the two tasks therefore replicating previous results in memory (e.g., Moulin, Perfect, & Jones, 2000). For the fluency task, there was no group difference in terms of performance: MS patients predicted their performance at the same magnitude as our healthy control group and were therefore as accurate. For the SDMT task, MS patients had a lower task performance. However, patients overall predicted the same number of items as our healthy control group therefore being less accurate. This was the case for the prediction of the first trial only. Thus, MS patients were able to have accurate predictions when having experienced the task (i.e., for postdictions and second-trial predictions). Across all tasks, MS patients can update their self-evaluation even though they have dysfunctional beliefs at first.

Patients' metacognitive knowledge was inaccurate, as gauged by the initial global predictions, for tasks before completing the task. As proposed in Mazancieux et al. (2019), such predictions are more associated with mood variables (depression, anxiety, etc.), fatigue, and self-esteem than with executive functions (that are more involved in monitoring processes). Although our sample of patients were more depressed and reported more fatigue than healthy control group participants, these variables were not correlated with the accuracy on the first prediction for the SDMT task. As we have previously suggested (Mazancieux et al. 2019), depression and fatigue in MS could lead to an underestimation of performance. However, in the present study, patients on average predict the same number of items as healthy control group participants. Therefore, we looked at the signed difference between this initial prediction and performance. The number of overestimators (16 patients) and underestimators (11 patients) was almost the same; however, there were no difference between these groups in terms of depression, $t(25) = -0.04$, $p = .971$, fatigue, $t(25) = 1.37$, $p = .181$, or disease duration, $t(25) = -0.20$, $p = .844$. There was thus no systematic under or overestimation in the MS group and no relation to other measures.

The processes involved in under- and overestimation are not the same. The underestimation of performance might be associated with low self-beliefs and concern about function. On the contrary, overestimation of performance can occur when patients have more cognitive impairment and therefore do not have enough cognitive resources to perform accurate predictions and to update their metacognitive knowledge when faced with changes in their function. In previous studies in MS, overestimation was also associated with more cognitive impairments (Carone et al., 2005; Rosti-Otajärvi et al., 2014; but see Smith & Arnett, 2010). Moreover, these impairments were more related to tasks measuring executive functioning; however, we did not measure these abilities, so we do not know if our two subsets of patients differ with this respect. Therefore, we suggest that future research should focus on executive function differences in MS and the relationship with a potential metacognitive deficit. In conclusion, in our sample of MS patients, we have a pattern which is consistent not with over or underestimation but a lack of accuracy in estimating an upcoming task for which they have not experienced. They are less accurate than the healthy control group in this regard only on a task where they are impaired (SDMT). If anything, future research could consider beliefs prior to conducting tasks, but in all other regards we did not find deficits in MS patients in metacognitive awareness per se with global measures, even when there is a significant deficit in performance. Once they have had the opportunity to experience a task, people with MS make an appropriate evaluation of their performance. Metacognition and disease awareness are complex multidimensional constructs, and it is clear that mood and knowledge impinge on people's evaluations. We proposed a multidimensional consideration of metacognition in a previous review (Mazancieux et al., 2019) but less is known about how these factors relate in MS than in other pathologies such as Alzheimer's disease (Mograbi & Morris, 2014).

Likewise, regarding FOKs, MS patients have the same metacognitive sensitivity as the healthy control group participants which does not reproduce previous findings (Beatty & Monson, 1991). The main difference between our work and the previous study is that we exclusively focused on RR-MS. Primary progressive (PP-MS) and secondary progressive (SP-MS) are the forms of MS with the most cognitive impairment (e.g., Planche et al., 2016). In an awareness interview, Sherman, Rapport, & Ryan (2008) showed that 51.5% of SP-MS patients have an unawareness of deficit compared to only 14.7% for RR-MS patients. Similarly, in Beatty and Monson's (1991), groups with impairment in episodic memory monitoring included at least half of PP-MS and SP-MS. It is therefore very likely that their patients are both more impaired and heterogenous than our sample (note that they had a lower score than controls on a verbal fluency test which was not the case in the present study). It remains a priority to consider disease type, severity, and duration to produce a full picture of metacognitive function in MS.

The present study suggests that RR-MS patients with slight cognitive impairments can adequately update their evaluations, therefore showing intact metacognitive monitoring. In our sample of MS patients, cognitive impairment results in lower performance on our version of the SDMT task only. If there is any evidence of metacognitive impairment, it is in inaccurate self-knowledge on a task where the MS group

https://doi.org/10.1017/S1355617720000776 Published online by Cambridge University Press

showed impairment, namely processing speed. On this initial prediction, consistent with the large variability in MS, half of the patients overestimated their performance whereas the other half underestimated it. The difference between under and over estimation was not captured by depression, fatigue, or disease duration measures in this study, perhaps due to our sample size. On a clinical note, it suggests that these patients are likely able to have adaptive strategies in daily living activities and will benefit from cognitive rehabilitation techniques more efficiently (Prigatano, 1999). A priority is now to verify this pattern in relatively homogenous groups of MS patients as used here but with more pronounced cognitive impairment. This would allow observing whether monitoring dysfunction occurs with more cognitive impairment. In particular, a focus on executive function would be interesting, as this is the domain typically affected in MS. In other populations, we see metacognitive impairment linked to the domain that is impaired (e.g., episodic memory and episodic memory monitoring in Alzheimer's disease; Souchay, 2007). In MS, if monitoring difficulties were linked uniquely to executive function, this might suggest the metacognitive deficit is a result of cognitive impairment rather than a direct symptom of MS which generalises to other domains. It will also be of clinical and theoretical relevance to take the metacognitive approach into domains which are perhaps more sensitive to the cognitive changes in MS, such as autobiographical memory (e.g., Ernst et al., 2013).

ACKNOWLEDGEMENTS

No specific financial support was provided for this research.

CONFLICT OF INTEREST

The authors have nothing to disclose.

References

Basso, M.R., Shields, I.S., Lowery, N., Ghormley, C., Combs, D., Arnett, P.A., & Johnson, J. (2008). Self-reported executive dysfunction, neuropsychological impairment, and functional outcomes in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *30*(8), 920–930.

Beatty, W.W., & Monson, N. (1991). Metamemory in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *13*(2), 309–327.

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). Beck depression inventory-II. *San Antonio*, *78*(2), 490–498.

Calabrese, P. (2006). Neuropsychology of multiple sclerosis. *Journal of Neurology*, *253*, 10–15.

Carone, D.A., Benedict, R.H.B., Munschauer III, F.E., Fishman, I., & Weinstock-Guttman, B. (2005). Interpreting patient/informant discrepancies of reported cognitive symptoms in MS. *Journal of the International Neuropsychological Society*, *11*(5), 574–583.

Chiaravalloti, N.D., & DeLuca, J. (2008). Cognitive impairment in multiple sclerosis. *The Lancet Neurology*, *7*(12), 1139–1151.

Connor, L.T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, *12*(1), 50–71.

Drew, M., Tippett, L.J., Starkey, N.J., & Isler, R.B. (2008). Executive dysfunction and cognitive impairment in a large community-based sample with Multiple Sclerosis from New Zealand: a descriptive study. *Archives of Clinical Neuropsychology*, *23*(1), 1–19.

Ernst, A., Blanc, F., Voltzenlogel, V., De Seze, J., Chauvin, B., & Manning, L. (2013). Autobiographical memory in multiple sclerosis patients: assessment and cognitive facilitation. *Neuropsychological Rehabilitation*, *23*(2), 161–181.

Fisk, J.D., Ritvo, P.G., Ross, L., Haase, D.A., Marrie, T.J., & Schlech, W.F. (1994). Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. *Clinical Infectious Diseases*, *18*, 79–83.

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911.

Fleming, S.M., & Lau, H.C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, 443.

Green, D.M., & Swets, J.A. (1966). *Signal Detection Theory and Psychophysics* (Vol. *1*). New York: Wiley.

Hertzog, C. (1992). Improving memory: the possible roles of metamemory. In *Memory Improvement* (pp. 61–78). Springer, New York.

Higham, P.A. (2011). Accuracy discrimination and type-2 signal detection theory: clarifications, extensions, and an analysis of bias. In *Constructions of Remembering and Metacognition* (pp. 109–127). Palgrave Macmillan, London.

Kesselring, J., & Klement, U. (2001). Cognitive and affective disturbances in multiple sclerosis. *Journal of Neurology*, *248*(3), 180–183.

Kinsinger, S.W., Lattie, E., & Mohr, D.C. (2010). Relationship between depression, fatigue, subjective cognitive impairment, and objective neuropsychological functioning in patients with multiple sclerosis. *Neuropsychology*, *24*(5), 573–580.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13).

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863.

Maor, Y., Olmer, L., & Mozes, B. (2001). The relation between objective and subjective impairment in cognitive function among multiple sclerosis patients-the role of depression. *Multiple Sclerosis Journal*, *7*(2), 131–135.

Mazancieux, A., Souchay, C., Casez, O., Moulin, C.J.A. (2019). Metacognition and self-awareness in multiple sclerosis. *Cortex*, *111*, 238–255.

McGlynn, S.M., & Schacter, D.L. (1989). Unawareness of deficits in neuropsychological syndromes. *Journal of Clinical and Experimental Neuropsychology*, *11*(2), 143–205.

Mograbi, D.C., & Morris, R.G. (2014). On the relation among mood, apathy, and anosognosia in Alzheimer's disease. *Journal of the International Neuropsychological Society*, *20*(1), 2–7.

Moulin, C.J., Perfect, T.J., & Jones, R.W. (2000). Global predictions of memory in Alzheimer's disease: evidence for preserved metamemory monitoring. *Aging, Neuropsychology, and Cognition*, *7*(4), 230–244.

Nelson, T. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*(1), 109–133.

Nelson, T.O., & Narens, L. (1990). The psychology of learning and motivation. In *Metamemory: A Theoretical Framework and New Findings*.

O'Brien, A., Gaudino-Goering, E., Shawaryn, M., Komaroff, E., Moore, N.B., & DeLuca, J. (2007). Relationship of the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ) to functional, emotional, and neuropsychological outcomes. *Archives of Clinical Neuropsychology*, *22*(8), 933–948.

Pannu, J.K., & Kaszniak, A.W. (2005). Metamemory experiments in neurological populations: a review. *Neuropsychology Review*, *15*(3), 105–130.

Phillips, L.J., & Stuifbergen, A.K. (2006). Predicting continued employment in persons with multiple sclerosis. *Journal of Rehabilitation*, *72*(1), 35–43.

Planche, V., Gibelin, M., Cregut, D., Pereira, B., & Clavelou, P. (2016). Cognitive impairment in a population-based study of patients with multiple sclerosis: differences between late relapsing–remitting, secondary progressive and primary progressive multiple sclerosis. *European Journal of Neurology*, *23*(2), 282–289.

Prakash, R.S., Snook, E.M., Lewis, J.M., Motl, R.W., & Kramer, A.F. (2008). Cognitive impairments in relapsing-remitting multiple sclerosis: a meta-analysis. *Multiple Sclerosis Journal*, *14*(9), 1250–1261.

Prigatano, G.P. (1999). *Principles of neuropsychological rehabilitation*. New York: Oxford University Press.

Rabbitt, P. (Ed.). (2004). *Methodology of frontal and executive function*. London: Psychology Press.

Randolph, J.J., Arnett, P.A., & Freske, P. (2004). Metamemory in multiple sclerosis: exploring affective and executive contributors. *Archives of Clinical Neuropsychology*, *19*(2), 259–279.

Rao, S.M., Leo, G.J., Bernardin, L., & Unverzagt, F. (1991). Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology*, *41*(5), 685–691.

Roberg, B.L., Bruce, J.M., Lovelace, C.T., & Lynch, S. (2012). How patients with multiple sclerosis perceive cognitive slowing. *The Clinical Neuropsychologist*, *26*(8), 1278–1295.

Rosti-Otajärvi, E., Ruutiainen, J., Huhtala, H., & Hämäläinen, P. (2014). Relationship between subjective and objective cognitive performance in multiple sclerosis. *Acta Neurologica Scandinavica*, *130*(5), 319–327.

Ruet, A. (2015). Cognitive Impairment in Multiple Sclerosis. In *Neuropsychiatric Symptoms of Inflammatory Demyelinating Diseases* (pp. 227–247). Cham: Springer.

Santiago, O., Guardia, J., Casado, V., Carmona, O., & Arbizu, T. (2007). Specificity of frontal dysfunctions in relapsing–remitting multiple sclerosis. *Archives of Clinical Neuropsychology*, *22*(5), 623–629.

Schnyer, D.M., Verfaellie, M., Alexander, M.P., LaFleche, G., Nicholls, L., & Kaszniak, A.W. (2004). A role for right medial prefrontal cortex in accurate feeling-of-knowing judgments: Evidence from patients with lesions to frontal cortex. *Neuropsychologia*, *42*(7), 957–966.

Sherman, T.E., Rapport, L.J., & Ryan, K.A. (2008). Awareness of deficit in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *30*(3), 301–311.

Shimamura, A.P., & Squire, L.R. (1986). Memory and metamemory: a study of the feeling-of-knowing phenomenon in amnesic patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(3), 452–460.

Smith, M.M., & Arnett, P.A. (2010). Awareness of executive functioning deficits in multiple sclerosis: self versus informant ratings of impairment. *Journal of Clinical and Experimental Neuropsychology*, *32*(7), 780–787.

Souchay, C. (2007). Metamemory in Alzheimer's disease. *Cortex*, *43*(7), 987–1003.

Souchay, C., & Isingrini, M. (2004). Age related differences in metacognitive control: role of executive functioning. *Brain and Cognition*, *56*(1), 89–99.

Souchay, C., Isingrini, M., Clarys, D., Taconnat, L., & Eustache, F. (2004). Executive functioning and judgment-of-learning versus feeling-of-knowing in older adults. *Experimental Aging Research*, *30*(1), 47–62.

Souchay, C., Moulin, C.J., Clarys, D., Taconnat, L., & Isingrini, M. (2007). Diminished episodic memory awareness in older adults: evidence from feeling-of-knowing and recollection. *Consciousness and Cognition*, *16*(4), 769–784.

Trapp, B.D., & Nave, K.A. (2008). Multiple sclerosis: an immune or neurodegenerative disorder? *Annual Review of Neuroscience*, *31*, 247–269.

Walker, L.A., Osman, L., Berard, J.A., Rees, L.M., Freedman, M.S., MacLean, H., & Cousineau, D. (2016). Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS): Canadian contribution to the international validation project. *Journal of the Neurological Sciences*, *362*, 147–152.

Westfall, J., Kenny, D.A., & Judd, C.M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045.

Wojcik, D.Z., Moulin, C.J., & Souchay, C. (2013). Metamemory in children with autism: exploring "feeling-of-knowing" in episodic and semantic memory. *Neuropsychology*, *27*(1), 19–27.