Effect of cognitive rehabilitation on neuropsychological and semiecological testing and on daily cognitive functioning in multiple sclerosis: the REACTIV randomized controlled study.

Lamargue D, PhD,[1,2] Koubiyr I,[1,2] Deloire M, PhD,[3] Saubusse A, MSc,[3] Charre-Morin J, MSc, [3] Moroso A, MD,[3] Coupé P, PhD,[4] Brochet B, MD,[1,2,3] Ruet A, MD, PhD.[1,2,3]

1 Univ. Bordeaux, F-33000 Bordeaux, France

2 Inserm U1215 - Neurocentre Magendie, F-33000 Bordeaux, France

3 CHU de Bordeaux, F-33000 Bordeaux, France

4 Laboratoire Bordelais de Recherche en Informatique, UMR CNRS 5800, PICTURA, F-33405 Talence, France

Corresponding Author:

Pr. Bruno Brochet, CHU Pellegrin, Place Amélie Raba Léon 33000 Bordeaux, France

Phone +33556795679, bruno.brochet@chu-bordeaux.fr

AUTHORS' CONTRIBUTION

Word count abstract: 204

Word count: 4214

Number of tables: 5

Number of figure: 2

E-tables: 3

Conflicts of interest

On behalf of all authors, the corresponding author states that there are no conflicts of interest in relation to the study.

Dr. Brochet reports grants from ANR, Fondation ARSEP and Merck during the conduct of the study; personal fees and non-financial support from Biogen, grants from Merck, personal fees and non-financial support from Novartis, personal fees and non-financial support from Roche, grant and non-financial support from Genzyme, non-financial support from Teva, grants from Bayer and non-financial support from Medday, non-financial support from Actelion, non-financial support from Celgene outside the submitted work; .

Dr. RUET reports grants from Agence Nationale Recherche, grants from Fondation ARSEP, grants from MERCK, during the conduct of the study; grants, personal fees and non-financial support from genzyme, grants and other from teva, grants from bayer, personal fees and non-financial support from medday, grants, non-financial support and other from ROCHE, personal fees, non-financial support and other from biogen, personal fees, non-financial support and other from NOVARTIS, non-financial support from merck, outside the submitted work; .All other authors have nothing to disclose.

KEYWORDS: MULTIPLE SCLEROSIS, COGNITIVE REHABILITATION, COGNITION, CLINICAL TRIAL, INFORMATION PROCESSING SPEED, VIRTUAL REALITY

ABSTRACT (204 words)

Background: Specific cognitive rehabilitation (SCR) has been suggested for multiple sclerosis (MS). A randomized controlled trial (RCT) evaluating the therapeutic effects of SCR is necessary.

Objective: To demonstrate the superiority of a SCR program (REACTIV) over nonspecific intervention (NSI) for neuropsychological (NP) assessment, virtual reality (VR) cognitive testing and daily cognitive functioning.

Methods: A single-blind RCT compared SCR and NSI in patients with MS with cognitive complaint. Both programs included 50 individual sessions, 3 times a week for 17 weeks in a real-world setting. The primary end-point was NP assessment. Secondary end-points included semiecological VR tasks (Urban Daily Cog®) and daily cognitive functioning assessment. Maintenance of the effects at 8 months was studied.

Results: Of the 35 patients, 18 completed the SCR, and 17 completed the NSI. Several NP and semiecological scores improved significantly more after SCR than after NSI. More NP scores improved significantly after SCR than after NSI. SCR improved daily cognitive functioning. Most improvements were maintained at 8 months.

Conclusion:

SCR performed in a real-world setting is superior to NSI for improving performance in specific cognitive domains and information processing speed, and for improving cognitive functioning, as evaluated by ecological tools close to daily life and a daily cognitive functioning questionnaire.

1. Introduction

Treating cognitive impairment (CI) in multiple sclerosis (MS), the leading cause of disability due to nontraumatic neurological disease in young adults, is an important challenge. The contribution of CI to disability in MS has been increasingly recognized, and CI has been shown to decrease health-related quality of life (HR-QOL), even in the early stages of the disease.[1] CI negatively impacts daily activities such as driving,[2] vocational status,[1] absenteeism,[3] and instrumental activities[4] in persons living with MS (PwMS). No medication has proven to have a consistent symptomatic effect on CI in MS, and disease-modifying therapies only have a small impact on CI progression.[5]

CI in MS is dominated by a slowdown in information processing speed (IPS), as well as by disturbances of more specific cognitive functions such as attention, episodic memory (EM), working memory (WM) and executive function (EF).[6] If a relatively circumscribed alteration in IPS linked to a specific process deficit can occur, changes in IPS can alter other cognitive processes and usually reflect cognitive functioning and efficiency. The alteration of IPS has consequences for WM, attention, EF and EM.[7] IPS impairment predicts subsequent disability,[8] vocational status and changes in quality of life.[1]

Cognitive rehabilitation (CR) is the most promising approach for treating MS-related CI, as concluded by recent reviews and meta-analyses, despite important methodological shortcomings.[9-11] However, methodological limitations in early studies have led to disappointing results[12], and well-designed studies are still scarce.[11] As noted recently, many studies lack an randomized controlled design that includes passive or active control conditions, primary neuropsychological end-points identified a priori, evidence of the sustainability of CR and the inclusion of near and far transfer

outcomes.[11] Tertiary outcomes of QOL, metacognition, or other patient-reported outcomes (PROs) are rarely used.[11]

Learning and memory represent the most common cognitive domain that has been targeted by CR in MS, and data concerning CR for IPS and other domains are less robust.[9] Recently, however, several well-conducted randomized controlled trials (RCTs) have shown significant improvement in some tests of EM,[13,14] WM,[15-17] EF, [5-17] and IPS.[14-18]

There is no consensus on the rehabilitation methods for CI in MS. Different strategies have been proposed, such as group-based rehabilitation, computer-assisted rehabilitation, individual rehabilitation with a therapist and mixed methods.

Computer-assisted rehabilitation has gained interest because of its feasibility without important professional resources. A recent meta-analysis[19] investigated the efficacy of computer-based cognitive rehabilitation on neuropsychological (NP) performance in PwMS and identified 9 studies with a control, but only one of these studies was randomized.[15] This meta-analysis showed a significant effect only on memory.[19] A recent multicenter randomized controlled study of computer-assisted rehabilitation in a large sample size showed a significant, although modest, effect in several domains.[20] Interestingly, this study showed the feasibility of cognitive rehabilitation in community settings.

Many studies have not focused on a specific domain. However, some studies have proposed strategy-oriented rehabilitation. One study used computer-assisted rehabilitation to improve IPS in an RCT in 21 MS patients and showed improvement in IPS through NP testing.[18] One RCT showed the superiority of specific cognitive rehabilitation (SCR) (n=22) over nonspecific psychological intervention (n=19) using

conversation without cognitive stimulation or other activities.[14] Another RCT using strategy-oriented NP rehabilitation in 102 patients, including computer-based attention, working memory retraining, psychoeducation, strategy learning and psychological support, did not show improved cognitive performance but did show reduced perceived cognitive deficits. However, the intensity of the program, one 60-minute session each week for 13 consecutive weeks, was relatively low.

Several important questions remain to be answered, however, including transfer to other domains, transfer to daily life functioning and generalizability. The clinical meaningfulness of the results has been assessed in some studies,[13,16-18] but the results are conflicting. An effect of the subjective report on overall functioning was reported in one study on memory rehabilitation,[13] and a positive effect was reported in the measure of IPS in daily life in one study[18]. However, other studies did not show any effect on the self-assessment of cognition using a visual-analog scale (VAS)[16] or HR-QOL.[17] Currently, the assessment of the effect of rehabilitation on daily functioning is mainly achieved using HRQOL questionnaires[17] or specific questionnaires of daily living activities.[18] A true ecological evaluation assessing functional activity in everyday life has been proposed but has not been used thus far for assessing CR.[22] Cognitive evaluation using a virtual-reality environment (VRE) has recently been proposed to assess cognitive functioning in a more ecological way[23] and could be an interesting tool for the evaluation of CR by collecting responses to random events in an environment similar to that of everyday life and by measuring these responses accurately with reaction times. Last, generalizability is defined as the maintenance of the effectiveness of the treatment with different practitioners, patients or settings, in particular the transfer from research setting to real-world practice,[24] and it has been poorly studied in MS.[11]

We raised the question of whether specific training using a large variety of exercises of increasing complexity that use computerized and pencil and paper mediums, rehabilitation games, and metacognition activities, applied in the same conditions as in clinical practice, focusing on IPS, attention, EF, and WM, could i) improve these cognitive processes more than nonspecific training at the same frequency and duration and ii) have beneficial effects on general cognitive functioning and daily cognitive activities in PwMS with mild to moderate CI. According to the stages of development and evaluation for behavioral treatments proposed by Rounsaville and colleagues,[24] we designed an RCT (stage 2) to answer this question with an SCR that is feasible in clinical practice and with a large NP assessment completed during a semiecological evaluation to measure the effect of SCR on cognitive functioning in a VRE and a specific questionnaire measuring daily cognitive activities.

2. Methods

2.1. Study design

REACTIV was a single-center, single-blinded RCT comparing the efficacy of a specific cognitive intervention (REACTIV) and a nonspecific cognitive intervention (NSI) in PwMS.

2.2. Standard protocol, approval, registration and consent

Each subject gave written informed consent. The ClinicalTrials.gov identifier was NCT01207856, and approval from a local ethics committee (CPP, Bordeaux, France) was obtained.

2.3. Subjects

PwMS complaining of discomfort in their daily lives due to cognitive problems during routine outpatient visits were selected. The inclusion criteria were as follows: MS according to the McDonald criteria[25] with any phenotype, age 18-55 years, disease duration >6 months and ≤15 years, right-handedness, and having a driver's license. To be eligible, PwMS also had to fulfill a cognitive criterion of mild CI (at least 3 scores <1 standard deviation (SD) on tests measuring IPS, attention, WM and EF). Included patients could have a normal score at one or several other tests.

The exclusion criteria were as follows: previous history of other neurological or psychiatric disorders; visual, oculomotor, auditory or motor impairments precluding the ability to perform computerized tasks; addictive behavior; MS attack and/or corticosteroid pulse therapy in the two months preceding the screening; severe cognitive deficits or dementia (Mini-Mental Status Examination <27);[26] moderate to severe visuo-spatial incapacity (raw score<28 on the copy trial of the Rey-Osterrieth Complex Figure Test); [27] and moderate to severe depression (Beck Depression Inventory (BDI) >27).[28]

Healthy subjects (HS) matched for age, gender and education were enrolled as controls. The exclusion criteria included a previous history of other neurological or psychiatric disorders, cognitive complaints, and moderate to severe depression.

2.4. Randomization and blinding

By using a computerized random number generator, groups were assigned via 1:1 randomization. For the purpose of the study, a rehabilitation supervisor, a professional experienced in CR and neuropsychology, supervised all sessions performed by speech therapists to ensure that they complied with the rehabilitation program for each PwMS. Unblinded personnel, aware of group assignments,

included the statistician in charge of the randomization, the rehabilitation supervisor and the therapists in charge of the patients. All other study personnel were blinded, in particular the evaluating neuropsychologist, the evaluating neurologist and the statistician in charge of the analysis. There was no communication between the therapists in charge of the patients, who were in private practices outside the hospital, and the evaluating neuropsychologist and neurologist, who were positioned in the hospital MS Clinic. The same evaluating neuropsychologist rated all the tests and conducted baseline and follow-up assessments whenever possible.

Participants were informed that they could be assigned to an SCR or to a global intervention including cognitive stimulation.

2.5. Interventions

After randomization, PwMS were allocated to either the REACTIV (SCR) group or a global intervention (NSI) group.

Training in both groups had the same frequency, duration and intensity (fifty 45-minute sessions, administered 3 times a week for 4 months). The sessions were supervised by specifically trained speech therapists or neuropsychologists. The two groups differed in the content of the sessions. No intervention was performed in the HS group.

The REACTIV program consisted only of individual sessions.

The REACTIV program was designed for mild to moderate impairment and focused on certain fundamental cognitive processes: IPS, using feedback from reaction times (RTs) in computerized and timed tasks; attention, especially selective, sustained and divided attention; EF, mainly inhibition and flexibility processes, control processes, the allocation and coordination of attentional resources or checking strategies; WM,

especially regarding the central executive and storage capacity; and metacognition. The rehabilitation program was progressive, including a general framework with systematic work on attention, IPS and EF but that was tailored according to the level of specific deficits presented by each patient. The progression was controlled by the validation of consecutive levels of difficulty. REACTIV used a large variety of exercises of increasing complexity to limit familiarization, maintain interest and novelty and stimulate attention, with computerized standardized exercises (TDA and Multiflex software programs (© GERIP)), pencil and paper exercises, and rehabilitation games. Tasks were performed across different modalities (visual or auditory, verbal or nonverbal, oral or written or motor). The program provided time for ecological work focusing on difficulties in daily life and for metacognitive deep thinking.

NSI sessions were devoted to information about the disease, its symptoms and its management; relaxation; physical activity coaching; and global cognitive stimulation, including 10 sessions with a special focus on semantic memory, autobiographical memory and verbal and visual episodic memory.

2.6. Assessments

2.6.1. *Clinical and NP assessments* were performed at baseline, month 4 and month 8 by a neurologist (clinical assessment) and by a neuropsychologist (cognitive testing), both blinded to the treatment allocation.

Neurological status was established using the Expanded Disability Status Scale.[29] PROs were administered at each evaluation. Depressive symptoms were assessed using the BDI,[28] and anxiety was measured by the State-Trait Anxiety Inventory (STAI).[30] Fatigue was assessed using the Modified-Fatigue Impact Scale (MFIS).[31]

The NP battery has been detailed in a previous publication[23] and is detailed in Table 1. The battery included paper-and-pencil tests and computerized subtests of the Test of Attentional Performances (TAP 2.1).[32]

A semiecological evaluation was performed using the Urban DailyCog©.[23] This task, which was conceived in-house and lasted 20 minutes, has been detailed previously.[8] Briefly, the VRE projected on a screen featured an urban environment, placing the subject in a car stopped at a street intersection. Three tasks were performed measuring a simple-alert reaction time (RT), a choice RT (selective attention, distractors to inhibit), and divided attention, including the two previous tasks.[23] The recorded task parameters were RTs and accurate answers (AA).

Daily cognitive activities were assessed by a 12-question PRO questionnaire (Daily Cognitive Activities Questionnaire, DCAQ, e-Table 1), which was designed in-house. Each question was scored from 0 to 5, and the total score was obtained by summing all answers (DCAQ total score 12Q).

HR-QOL was measured by the French-validated version of the Short-Form 36 questionnaire, and two composite scores were calculated: the Physical Composite Score (PCS/SF-36) and the Mental Composite Score (MCS/SF-36).[33]

2.6.2. Magnetic Resonance Imaging

MRI scans were performed at baseline and after 4 and 8 months to compare lesion load and volumetric data at baseline and after follow-up (see Supplementary material).

2.7. Statistics and analyses

The primary end-point was the NP assessment of IPS, attention, EF and WM after 4 months. Secondary end-points included the NP assessment 4 months after

the end of treatment; PROs about depression, anxiety, fatigue, HR-QOL and daily functioning; and VRE assessment (Urban Daily Cog®). The statistical plan of the study used repeated measures analysis of variance for the primary end-point. However, the sample size was limited; the conditions for such an analysis were not fulfilled, as the changing trend was not linear (nonnormal distribution); and the variances lacked homogeneity. Significance between groups was, therefore, assessed by preplanned nonparametric tests (the Wilcoxon test when comparing 2 time points in each group and the Mann-Whitney U test when investigating the change over 4 months between groups). Differences between month 8 and month 4 scores were assessed using the Wilcoxon test for the scores improved at month 4 in the NSI and REACTIV groups.

3. Results

3.1. Demographic, clinical and NP characteristics.

Thirty-five PwMS and 21 HS participated in the study. Figure 1 presents a flow chart of the study.

Table 2 summarizes the baseline demographic, clinical and NP characteristics of the PwMS and HS. MRI characteristics are presented in e-Table 2. None of the demographic, clinical, NP or imaging metrics were different between the two groups, except for one NP score (visual divided attention RT, TAP; $p<0.05$). MRI and clinical parameters of both groups remained stable during the study. All NP scores were different between PwMS and HS, except for a few subtests in the TAP and the recognition score of the CVLT (Table 1). All RTs in the Urban DailyCog® differed significantly between PwMS and HS (Table 2).

PwMS and HS differed on all DCAQ scores (e-Table 1) and on scores across all axes of the SF-36 ($p<0.01$), except for role-emotional/SF-36. The composite scores were significantly lower than those of HS ($p<0.001$) (Table 2). PRO scores (BDI, STAI) differed significantly between PwMS and HS (Table 2). Correlations at baseline between the different assessments are presented as supplementary material (e-Table 3).

3.2. Effect on IPS, attention, EF and WM (primary end-point)

According to Mann-Whitney U test analyses, several scores improved significantly more in the REACTIV group than in the NSI group (data in bold in Table 3). According to the Wilcoxon test results, the majority of IPS test RTs improved significantly in the REACTIV group but not in the NSI group (data in bold in Table 3). Several NP tests of attentional components, WM (storage, updating) and EF (mainly inhibition and flexibility), improved in both groups.

3.3. Effect on other NP measures

According to Mann-Whitney U test analyses, only the CVLT cued immediate recall improved more in the REACTIV group than in the NSI group (Table 3). Some memory scores improved in the REACTIV group (free delayed recall CVLT, $p<0.05$; free and cued immediate recall CVLT, $p<0.01$), while only Rey figure recall improved in the NSI group ($p<0.05$).

3.4. Effect on ecological evaluation

According to Mann-Whitney U test analyses, RTs in two of the three tasks of the Urban DailyCog® improved significantly more in the REACTIV group than in the NSI group, and a trend was noticed in the third task (task 2) ($p=0.06$) (Figure 2). Accurate answers in task 3 improved significantly more in the REACTIV group than in

the NSI group ($p<0.01$) (Table 4). Three RTs improved significantly only in the REACTIV group ($p<0.001$, $p<0.05$ and $p<0.01$, respectively) (Wilcoxon test) (Figure 2). Similar results were observed for the error and omission results in the ecological tasks (data not shown).

3.5. Effect on daily cognitive functioning

A significant improvement was observed only in the REACTIV group for 8 of the 12 questions in the DCAQ and for the DCAQ total score 12Q (Q1-Q12) (Wilcoxon), although the Mann-Whitney U test analyses did not show a significant difference between groups (Table 4).

3.6. Effect on PROs

A significant improvement in BDI, total MFIS, physical MFIS and cognitive MFIS scores was observed in both groups (Table 5). Changes in NP scores or Urban Daily Cog® scores did not correlate with changes in BDI or STAI scores.

3.7. Change in healthy subjects

Four months after the baseline evaluation, without any intervention, very limited changes were observed in HS (n=22), only in visual scanning scores with and without the target ($p<0.01$ and $p<0.05$) and Stroop task naming and reading times ($p<0.01$). No significant change was observed in Urban DailyCog® scores.

3.8. Month 8 assessment

Among the NP and ecological scores significantly improved at month 4 in the REACTIV group, only two scores changed moderately, but significantly, between months 4 and 8, showing further improvement (decreased simple auditory attention RT and visual scanning with target RT; $p<0.05$ for each). Visual scanning with target RT also improved in the NSI group ($p<0.05$). However, none of the ecological and NP

test scores that had improved more in the REACTIV group than in the NSI group changed significantly during the follow-up after the end of the intervention. The scores on the PROs (including the DCAQ) remained stable between months 4 and 8 in both groups.

4. Discussion

In this study, we showed that REACTIV, an SCR that focused on IPS, EF (mainly inhibition and flexibility), WM and attention for 4 months, was able 1) to improve performance in a larger number of IPS tasks compared to that with an NSI including global cognitive stimulation administered with the same duration and intensity; and 2) to induce superior improvement in several tasks over that seen after nonspecific training, mainly for IPS measured by conventional tests and, more accurately, by computerized tests (RTs in ms). Additionally, in contrast to the control program, the SCR was able 3) to improve ecological tasks performed in a VRE and 4) to improve self-estimated daily cognitive functioning. Interestingly, we found that the principal effects were maintained 4 months after the end of rehabilitation. Finally, the SCR program was performed in a real-world setting (private outpatient practices), which assures its feasibility in daily practice.

The REACTIV rehabilitation program focused on IPS, attention, WM and EF (inhibition and flexibility). Some previous randomized studies performed in MS have targeted memory or various functions, including memory, WM, IPS and EF.[14-17] One study more specifically targeted IPS.[18] In this study, the treatment condition consisted of 10 computerized training sessions administered over a 5-week period using three types of tasks (simple processing speed, divided attention, and selective attention) with three different demands (detection, identification, same/different). The treated group (n=12) performed better than the control group (n=9) on the primary outcome

measures of IPS, the Digit Symbol Coding Subtest from the Wechsler Adult Intelligence Scale-III (WAIS-III) and the Timed Instrumental Activities of Daily Living Test (TIADL), which measures speeded everyday life tasks. No effect was observed on the Letter Comparison and Pattern Comparison scores. No reaction times were used. Another study compared the outcomes of an SCR for IPS, EF and memory to those of an unspecific psychological intervention in 41 PwMS, showing that more tests of IPS and memory were improved in the SCR group than in the control group.[14] Other studies with computer-assisted rehabilitation found an effect in a single test: the SDMT[17] or the PASAT.[16] Scores in a large number of NP tests measuring IPS and attention were improved in the REACTIV group, showing a consistent effect in these domains; a superior effect was shown in the SCR group compared to that in the control group for several RTs in tasks that very precisely measured RTs and the characteristics of IPS (alertness; alertness with warning) and auditory attention.The REACTIV program differed from the abovementioned studies in terms of various important characteristics in the procedures evaluated. First, several studies used computer-assisted rehabilitation,[14,16,17] and one study performed rehabilitation in group sessions.[15] During the REACTIV program, PwMS were trained by speech therapists using a mix of pen-and-pencil tasks and computerized tasks during 3 individual 45-minute sessions each week for 4 months, which was similar to one study measuring intensity[17] but was superior to others and had a longer duration. The increasing difficulty level, the variety of tasks, the type of presentation and the duration of the exercises, which was lengthened gradually according to the progress of the patient, could explain, in part, the effect on cognition, but also on fatigability, during the sessions. One study, using a large and progressive variety of pen-and-pencil exercises during group intervention, showed an effect in several domains.[15]

This content associated with the specific work on metacognition could be important for obtaining a broad effect.

Other methodological aspects should be underlined. The tasks used during exercises and practice were different from tasks used during assessments. The NSI group received an intervention with the same duration and intensity. The only difference was the content of the program. This control program was not a “placebo” intervention or a waiting list group but a global nonspecific cognitive stimulation program. It is likely that this control intervention had some effect because improvement was observed in several NP scores in the NSI group and in some PROs, but interestingly, there was no significant improvement in the NSI group in the DCAQ, a PRO assessing daily cognitive functioning, or in ecological evaluation using the Urban DailyCog®. Moreover, almost no improvement was observed in the HS group, suggesting that the practice effect is unlikely to contribute to the improvement observed in PwMS.

Interestingly, both interventions improved depressive symptoms and fatigue, but changes in BDI, STAI and MFIS scores did not correlate with cognitive improvement in the REACTIV group. These results indicated that the mechanisms for cognitive improvement are different from those of mood and that cognitive improvement is not due to psychological improvement, even though SCR could have an impact on some other symptoms. It was previously observed in one study that reduced perceived cognitive deficits after rehabilitation did not result in significant cognitive improvement in conventional test performance.[21] The effect of SCR appears to be larger and more beneficial than that of NSI, confirming the positive results of a previous trial.[14] A very interesting point is that PwMS in the 2 groups did not have any cerebral modifications in terms of brain or lesion volumes, which could

interfere with the results, a parameter that was barely controlled in the RCT investigating CR over several months.

The evidence of an effect of CR on daily cognitive functioning using VAS, HR-QOL scales, or performance-based measures of functional activities is limited.[13,16-18] In the present study, several lines of evidence suggest that improvement related to the intervention has meaningful functional consequences. The first consequence is that PwMS in the REACTIV group improved on 8 of the 12 questions on the DCAQ and the total DCAQ score 12Q, although no improvement was observed in the NSI group for these measures. All questions in the DCAQ discriminated well between PwMS and HS, and the total DCAQ score 12Q correlated mainly with attention NP scores. A non-specific improvement of the DCAQ, due to the improvement in mood and other emotional dimensions could be discussed. However PROs improved in both groups, although DCAQ improved significantly only in patients in the REACTIV group. The second important argument is related to the results of the ecological evaluation with the Urban DailyCog®. This task, when presented in a VRE, has been shown to be very sensitive to CI in MS but is not strongly correlated with classic NP tasks, suggesting a different way to evaluate and implicate cognitive interactions.[23] The effect of the SCR was significant for the 3 never-trained tasks of the test, and the REACTIV group was more improved than the NSI group in two of the tasks, including the more complex and effortful task requiring a high mental load and several cognitive processes; this result suggested a generalization of the effect. This finding is the first demonstration that an SCR could improve a semiecological evaluation using a VRE in MS. The effect was globally maintained at month 8. Longer follow-up studies will tell us how long such an effect can last.

This study has some limitations. Indeed, the sample size was relatively small, although it was similar to those used in other studies.[14,15,17,18] The recruitment of such trials is quite challenging because participation is time-consuming for PwMS due to the various evaluations and the intervention schedule. However, despite a small sample size, we were able to show consistent results in favor of SCR. One can expect that a larger sample size would have produced more important differences between groups.

The study also has important strengths. The design, an RCT with blind evaluation, was robust. Groups are comparable across clinical and MRI parameters, which remained stable throughout the study. The use of an active global intervention (not a placebo) with the same intensity in the control MS group underlines the specificity of the results. Moreover, the addition of a control group of HS is important to control for practice effects. Ecological and daily cognitive functioning assessments strengthened the clinical meaningfulness of the results. Reporting on improvements or, otherwise, on ecological functioning in everyday life to measure the impact of CR, the possibility of knowledge transfer and generalization in everyday life involving different cognitive interactions is a key issue for continuing to optimize CRs by guiding rehabilitation axes and strategies and by specifying recommendations for current city clinical practice. The results of this study, suggesting some efficacy of specific cognitive rehabilitation on IPS and daily cognitive functioning, need confirmation by further controlled studies.

Acknowledgments

Funding

This study was funded by the ARSEP Foundation (Association pour la Recherche contre la Sclérose en Plaques), by the ANR-10-LABX-57 Translational Research and Advanced Imaging Laboratory (TRAIL), Cluster of Excellence (Université de Bordeaux, France) and by a grant from Merck. These funding partners did not participate in any aspect of the design or execution of the study, including data collection, management, analysis or interpretation. Furthermore, they were not involved in the preparation, review or approval of this manuscript.

Acknowledgments

This work was made possible by a contribution by the AQUISEP network, a clinical network in Nouvelle-Aquitaine (France) devoted to MS care and, in particular, we would like to acknowledge speech therapists participating in this network: Claire Lété (Saint Médard en Jalles), Nelly Blanchot (Bordeaux), Isabelle Chaubron (Tonneins), Bénédicte Darrigrand (Libourne), Elodie Laborie (Arcachon), Laurena Bedouch (La Teste de Buch), Karine Munos-Gaunelle (Bordeaux), Virginie Berland (Bordeaux), Alice Roux (Pessac), Cécile Bellorini (Lormont), and Emmanuelle Sous (Langon).
The following neurologists contributed to the study: J-C Ouallet, A. Laporte, P. Louiset, I. Patry, A. Bredin, P. Gaida, J-M. Faucheux and F. Radji.

Komenvoir society (Toulouse, France), headed by Madeleine Torrell and her associate, especially Alexandre Bouas, realized The Urban Daily Cog with Delphine Lamargue.

Pierre Joly (Biostatistics Team, ISPED, Bordeaux University, France) and Nicolas Paris (Optima society, Floirac, France) gave statistical advice.

Figure Legends:

Figure 1: Flow-chart of the study

Figure 2: Changes in Urban DailyCog® scores from pre- to post-treatment by group (a lower score is better). Outliers were not excluded from the analysis.

REFERENCES:

1. Ruet A, Deloire M, Hamel D et al. Cognitive impairment, health-related quality of life and vocational status at early stages of multiple sclerosis: a 7-year longitudinal study. J Neurol. 2013; 260 : 776-84.
2. Devos H, Ranchet M, Backus D, et al. Determinants of On-Road Driving in Multiple Sclerosis. Arch Phys Med Rehabil. 2017; 98(7):1332-1338.
3. Salter A, Thomas N, Tyry T, Cutter G, Marrie RA. Employment and absenteeism in working-age persons with multiple sclerosis. J Med Econ. 2017; 20(5):493-502.
4. Goverover Y, Genova HM, Hillary FG, DeLuca J. The relationship between neuropsychological measures and the Timed Instrumental Activities of Daily Living task in multiple sclerosis. Mult Scler. 2007; 13(5):636-44.
5. Amato MP, Langdon D, Montalban X, et al. Treatment of cognitive impairment in multiple sclerosis: Position paper. J Neurol. 2013; 260(6): 1452–1468.
6. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. Lancet Neurol. 2008; 7: 1139-51.
7. Costa SL, Genova H, DeLuca J, et al., Information processing speed in multiple sclerosis: Past, present, and future. Mult Scler. 2017; 23: 772–789.
8. Deloire M, Ruet A, Hamel D, Bonnet M, Brochet B. Early cognitive impairment in multiple sclerosis predicts disability outcome several years later. Mult Scler.

2010; 16: 581–587.

9. das Nair R, Martin KJ, Lincoln NB. Memory rehabilitation for people with multiple sclerosis. Cochrane Database Syst Rev. 2016; 3: CD008754.
10. Kalb R, Beier M, Benedict RH, et al. Recommendations for cognitive screening and management in multiple sclerosis care. Mult Scler. 2018; 24(13):1665-1680.
11. Sandroff BM, DeLuca J. Will behavioral treatments for cognitive impairment in multiple sclerosis become standards-of-care? Int J Psychophysiol. 2019; doi: 10.1016/j.ijpsycho.2019.02.010. [Epub ahead of print]
12. Rosti-Otajärvi EM, Hämäläinen PI. Neuropsychological rehabilitation for multiple sclerosis. Cochrane Database Syst Rev. 2014; 2: CD009131.
13. Chiaravalloti ND, Moore NB, Nikelshpur OM, et al. An RCT to treat learning impairment in multiple sclerosis: The MEMREHAB trial. Neurology. 2013; 81(24): 2066–2072.
14. Mattioli F, Stampatori C, Bellomi F et al. A RCT comparing specific intensive cognitive training to aspecific psychological intervention in RRMS: the SMICT study. Front. Neurol. 2014; 5:278.
15. Rilo O, Peña J, Ojeda N, et al. Integrative group based cognitive rehabilitation efficacy in multiple sclerosis: A randomized clinical trial. Disabil Rehabil. 2018; 40(2):208-216.
16. Amato MP, Goretti B, Viterbo RG, et al. Computer assisted rehabilitation of attention in patients with multiple sclerosis: Results of a randomized, double blind trial. Mult Scler. 2014; 20(1): 91–98.
17. Campbell J, Langdon D, Cercignani M, et al. A randomised controlled trial of efficacy of cognitive rehabilitation in multiple sclerosis: A cognitive,

behavioural, and MRI study. Neural Plast. 2016; 2016: 4292585.

18. Chiaravalloti ND, Goverover Y, Costa SL, DeLuca J.A Pilot Study Examining Speed of Processing Training (SPT) to Improve Processing Speed in Persons With Multiple Sclerosis. Front Neurol. 2018; 9:685.
19. Dardiotis E, Nousia A, Siokas V, Tsouris Z, Andravizou A, Mentis AA, Florou D, Messinis L, Nasios G. Efficacy of computer-based cognitive training in neuropsychological performance of patients with multiple sclerosis: A systematic review and meta-analysis. Mult Scler Relat Disord. 2018; 20:58-66.
20. Stuifbergen AK, Becker H, Perez F, Morrison J, Brown A, Kullberg V, Zhang W. Computer-assisted cognitive rehabilitation in persons with multiple sclerosis: Results of a multi-site randomized controlled trial with six month follow-up. Disabil Health J. 2018; 11(3):427-434.
21. Mäntynen A, Rosti-Otajärvi E, Koivisto K, Lilja A, Huhtala H, Hämäläinen P. Neuropsychological rehabilitation does not improve cognitive performance but reduces perceived cognitive deficits in patients with multiple sclerosis: a randomised, controlled, multi-centre trial. Mult Scler. 2014; 20(1): 99-107.
22. Goverover Y, DeLuca J. Assessing everyday life functional activity using actual reality in persons with MS. Rehabil Psychol. 2018; 63(2):276-285.
23. Lamargue-Hamel D, Deloire M, Saubusse A et al. Cognitive evaluation by tasks in a virtual reality environment in multiple sclerosis. J Neurol Sci. 2015; 359: 94–99.
24. Rounsaville, B.J., Carroll, K.M., Onken, L.S.,. A stage model of behavioral therapies research: getting started and moving on from stage I. Clin. Psychol. Sci. Prac. 2001; 8,133–142.

25. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol. 2011; 69(2):292-302.
26. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician, J. Psychiatr. Res. 1975; 12: 189–198.
27. Rey A. Manuel du Test de copie et de reproduction de mémoire de figures géométriques complexes. Paris : Éditions Centre de psychologie appliquée, 1959.
28. Beck AT, Brown G and Steer RA. Beck Depression Inventory II manual. San Antonio, TX: The Psychological Corporation, 1996.
29. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: An Expanded Disability Status Scale (EDSS). Neurology. 1983; 33: 1444–1452.
30. Spielberger CD, Gorssuch RL, Lushene PR, et al. Manual for the State-Trait Anxiety Inventory. Palo Alto: Consulting Psychologists Press, 1983.
31. Fisk JD, Ritvo PG, Ross L, et al. Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. Clin Infect Dis. 1994; 18(S1): S79–S83.
32. Zimmermann P and Fimm B. TAP 2.1 Tests d’evaluation de l’attention. Wurzelen: Psytest 2009.
33. Leplège A, Ecosse E, Coste J, Pouchot J, Perneger T (2001) Le questionnaire MOS SF-36: manuel de l’utilisateur et guide d’interprétation des scores. Editions Estem, Paris.

Table 1: Neuropsychological assessment

Tests	Parameters
SDMT (IPS)	
Alertness (TAP) (IPS)	RT without warning
	RT with warning
	AA without warning
	AA with warning
Visual scanning (TAP) (IPS/Attention)	RT with a target
	AA with a target
	RT without a target
	AA without a target
Divided attention (TAP) (IPS/Attention)	RT, visual simple condition
	Visual: AA, simple condition
	Auditory: RT, simple condition
	Auditory: AA, simple condition
	Visual: RT, dual task
	Visual: AA, dual task
	Auditory: RT, dual task
	Auditory: AA, dual task
N-back (TAP) (WM)	RT
	AA
Stroop test (IPS/EF)	Color Naming (time)
	Word Reading (time)
	Interference (time)
TMT (EF; flexibility: B-A)	Part A (time)
	Part B (time)
	B-A

Baddeley's Dual task (WM/attention)	Forward span (AA) *mu*
Verbal fluency (EF)	Semantic AA
	Phonemic AA
Reverse span (WM)	Backward span (AA)
CVLT (EM, verbal)	Learning trials, list A (AA)
	Learning trial, list B (AA)
	Immediate recall (AA)
	Immediate cued recall (AA)
	Delayed recall (AA)
	Delayed cued recall (AA)
	Recognition (AA)
Rey complex figure (Visuo-construction and EM, visual)	Copy (AA)
	Copy (time)
Naming task (DO 80) (access to lexical store)	AA
Urban DailyCog®	Task 1 TR, AA. Task 2 TR, AA. Task 3 TR (traffic light), AA. Task 3 TR (boy), AA.

IPS = information processing speed; SDMT= Symbol Digit Modalities Test; TAP= Test of Attentional Performance; RT= Reaction time; AA= Accurate answer; WM= Working memory; EF= Executive function; TMT= Trail Making Test; CVLT= California Verbal-Learning Test; EM= Episodic memory; *mu*: index of the ability to coordinate concurrent box-crossing and digit-span tasks in the dual task of Baddeley. DO 80= test of oral naming ("denomination orale 80").

TABLE 2: Demographic, clinical and NP characteristics of PwMS and HS at baseline:

		NSI (n= 17)	REACTIV (n=18)	HS (n=21)
Age (years)		38.3±8.2	43.8±5.6	39.7±7.3
Gender (% of Women)		82.4% (14 F/3 M)	66.6% (12 F/6 M)	80.9% (17 F/4 M)
Education (> bac)		47.1%	55.5%	57.1%
Disease duration (years)		6.5±5.5	6.7±3.1	NA
EDSS		2.0 [0-4]	3.0 [1-8]	NA
Clinical phenotype (RR/SP/PP)		15/1/1	14/3/1	NA
BDI		15.6±7.7	16.4±7.4	5.9±5.2***
STAI A (state)		32.9±6.8	36.6±12.6	30.0±8.8
STAI B (trait)		45.8±8.0	47.0±10.7	34.7±9.9***
PCS SF-36		52.2±14.9	50.3±17.2	89.9±7.5***
MCS SF-36		53.0±18.6	50.0±22.8	80.2±12.5***
MFIS (total score)		56.5±13.0	54.2±13.2	15.3±12.6***
Tests	Subtests			
SDMT		52.0±8.8	51.7±10.2	69.9±8.7***
Alertness (TAP)	RT (without warning)	274.2±38.1	334.1±208.9	242.3±28.4***
	RT (with warning)	250.2±26.3	335.7±238.4	237.6±36.7*
Visual scanning (TAP)	RT with a target	3344.2±552.8	3499.5±1241.8	2734.5±558.6**
	AA with a target	35.5±9.4	39.1±7.4	43.4±5.2**
	RT without a target	6089.2±1500.2	6101.2±2076.1	5351.5±1443.3
	AA without a target	49.8±0.4	49.8±0.4	49.9±0.3
Divided attention (TAP)	Visual: RT, simple task	884.9±111.0	888.7±136.7	772.0±86.1**
	Visual: AA, simple task	15.8±1.5	15.4±1.8	16.6±0.9**
	Auditory: RT, simple	572.7±126.4	646.2±151.2	572.9±85.9

	task Auditory: AA, simple task	15.8±0.5	15.4±1.9	15.9±0.4
	Visual: RT, dual task	818.9±91.4	904.4±145.2[a]	741.4±67.3***
	Visual: AA, dual task	15.8±2.2	15.5±1.6	16.4±1.2*
	Auditory: RT, dual task	642.5±149.5	704.3±210.6	584.2±103.0*
	Auditory: AA, dual task	14.7±1.4	14.8±2.9	15.9±0.3**
N-back (TAP)	RT	724.2±146.9	781.2±208.9	563.2±146.7***
	AA	13.4±1.8	12.5±1.8	14.3±0.9**
Stroop test	Color Naming (time)	75.2±16.2	73.1±22.8	54.2±7.2***
	Word Reading (time)	51.4±9.2	55.4±24.4	41.0±6.1***
	Interference (time)	50.2±14.9	59.7±29.9	33.9±19.0***
TMT	Part A (time)	34.1±8.1	44.4±25.9	24.6±6.3***
	Part B (time)	80.8±17.1	89.6±32.8	52.0±14.9***
	B-A	46.8±14.3	45.2±19.7	27.4±11.5***
Baddeley's Dual task	Forward span (AA)	5.5±1.0	5.1±1.0	6.0±0.9*
	Dual task (*mu*)	91.2±33.1	89.6±38.7	118.8±16.8***
Verbal fluency	Semantic AA	29.0±7.7	28.2±6.5	32.2±8.0
	Phonemic AA	21.2±5.0	18.8±5.7	24.6±6.1**
Reverse span	Backward span (AA)	3.8±0.8	3.8±0.8	4.3±1.1
CVLT	Learning trials, list A	62.5±8.6	61.2±7.7	70.1±6.2***
	Learning trials, list B	8.3±2.4	7.6±2.3	9.9±1.4*
	Immediate recall	13.0±2.8	12.1±2.4	15.2±0.9***
	Immediate cued recall	13.5±1.8	12.8±1.9	15.2±0.9***
	Delayed recall	13.6±1.8	12.9±2.5	15.2±1.2***
	Delayed	13.6±1.9	13.2±2.4	15.4±0.8***

	cued recall			
	Recognition	15.5±0.8	15.2±1.5	15.7±0.5
Rey complex figure	copy (AA)	33.5±2.6	33.6±2.4	35.4±0.9**
	time	19.2±6.0	17.9±6.2	24.5±5.0**
Naming task	AA	77.9±1.3	78.1±1.1	79.0±1.3**
Urban DailyCog®	Task 1 RT	636.9±102.8	677.4±258.5	562.5±86.8**
	Task 2 RT	1489.1±215.9	1567.5±316.5	1356.6±227.6*
	Task 3 RT (traffic light)	876.4±152.1	997.2±430.4	718.4±107.2***
	Task 3 RT (boy)	1444.7±241.6	1625.9±415.3	1233.4±246.1**

Data are expressed as the means ± SD except for the EDSS (median). NSI = Nonspecific training group (MS patients); REACTIV= Specific training group (MS patients); HS= Healthy subjects; bac= Baccalaureate; EDSS= Expanded Disability Status Scale; RR= Relapsing-remitting; SP= Secondary progressive; PP= Primary progressive; BDI= Beck Depression Inventory; STAI= State-Trait Anxiety Inventory; MFIS= Modified Fatigue Impact Scale; SDMT= Symbol Digit Modalities Test; TAP= Test of Attentional Performance; RT= Reaction time; AA= Accurate answer; TMT= Trail Making Test; CVLT= California Verbal-Learning Test; *mu*: index of the ability to coordinate concurrent box-crossing and digit-span tasks in the dual task of Baddeley.

*:p<0.05; **: p<0.01; ***: p<0.001 (p value between Persons with MS (REACTIV + NSI groups) and HS using Mann-Whitney U test). [a] = p value <0.05 between the REACTIV and NSI groups.

TABLE 3: Change in NP test scores

Tests	Subtests	Variables	NSI V0	REACTIV V0	NSI V4	REACTIV V4	NSI V8	REACTIV V8	Scores improved at V4 (NSI) (Wilcoxon test)	Scores improved at V4 (REACTIV) (Wilcoxon test)	Scores more improved at V4 (REACTIV/ NSI (Mann-Whitney U test)
Symbol Digit Modalities Test		**AA (in 90s)**	**52 ±8.8**	**51.7 ± 10.2**	**57.2 ±9.1**	**57.8 ±10.2**	**59.4 ±10**	**58.7 ±10.3**	**p<0.01**	**p<0.01**	**ns**
Alertness (TAP)	Phasic Alertness without a warning	**RT**	**274.2 ±38.1**	**334.1 ±208.9**	**273.3 ±38.6**	**250.1 ±45.1**	**267.2 ±38.4**	**251.1 ±38.9**	**ns**	**p<0.001**	**p<0.01**
		AA	40 ±0	40 ±0	40 ±0	40 ±0	40 ±0	40 ±0	NA	NA	NA
	Phasic alertness with a warning	**RT**	**250.2 ±26.3**	**335.7 ±238.4**	**261.9 ±41.2**	**248.2 ±45.5**	**250 ±31**	**262.5 ±75.1**	**ns**	**p<0.01**	**p<0.01**
		AA	40 ±0	40 ±0	40 ±0	40 ±0	40 ±0	40 ±0	NA	NA	NA
Visual scanning (TAP)	Visual scanning with a target	**RT**	**3344.2 ±552.8**	**3499.5 ±1241.8**	**3023.5 ±786**	**3325.3 ±1291.4**	**2789.2 ±628.7**	**2860.2 ±906.8**	**ns**	**ns**	**ns**

		AA	35.5 ±9.4	39.1 ±7.4	40.3 ±7.8	43.6 ±4.4	43.1 ±4.4	40.8 ±6	p<0.05	p<0.05	ns
	Visual scanning without a target	**RT**	**6089.2 ±1500.2**	**6101.2 ±2076.1**	**5723.4 ±1837.3**	**6460.1 ±2882.9**	**5456.7 ±1467.1**	**5596.6 ±1995.7**	**ns**	**ns**	**ns**
		AA	49.8 ±0.4	49.8 ±0.4	50 ±0	49.8 ±0.4	49.8 ±0.4	49.8 ±0.5	p<0.05	ns	ns
Divided Attention(TAP)	Visual attention (*simple task condition*)	**RT**	**884.9 ±111**	**888.7 ±136.7**	**856.9 ±120.7**	**803.6 ±97.3**	**829.2 ±98.1**	**809.3 ±145.2**	**ns**	**p<0.01**	**ns**
		AA	15.8 ±1.5	15.4 ±1.8	16.2 ±1.4	15.9 ±1.3	15.8 ±1.4	15.7 ±2.2	ns	ns	ns
	Auditory attention (*simple task condition*)	**RT**	**572.7±126.4**	**646.2 ±151.2**	**560.1 ±119.8**	**525.4 ±90.6**	**561.5 ±124.4**	**558.4 ±64.2**	**ns**	**p<0.01**	**p<0.05**
		AA	15.8 ±0.5	15.4 ±1.9	15.5 ±1	16 ±0	15.8 ±0.4	15.8 ±0.4	p<0.05	ns	p<0.05
	Visual attention (*dual-task condition*)	**RT**	**818.9 ±91.4**	**904.4 ±145.2**	**800.5 ±91.8**	**824.8 ±115.5**	**811.2 ±79.7**	**782.9 ±116.7**	**ns**	**p<0.05**	**ns**
		AA	15.8 ±2.2	15.5 ±1.6	16.1 ±1.2	16.3 ±0.8	15.8 ±1.1	16 ±1	ns	p<0.05	ns

	Auditory attention (*dual-task condition)*	**RT**	**642.5 ±149.5**	**704.3 ±210.5**	**583.6 ±98.6**	**574.4 ±102.3**	**588.4 ±101.4**	**596.9 ±106.5**	**p<0.05**	**p<0.05**	**ns**
		AA	14.7 ±1.4	14.8 ±2.9	15.3 ±1.5	15.7 ±0.6	15.3 ±1.7	15.7 ±0.8	ns	ns	ns
N-back (TAP)	N-Back	**RT**	**724.2 ±146.9**	**781.2 ±208.9**	**753.1 ±209.8**	**703.3 ±120.8**	**698.5 ±214.7**	**648.3 ±103.5**	**ns**	**ns**	**ns**
		AA	13.4 ±1.8	12.5 ±1.8	13.2 ±1.6	13.7 ±1	13.6 ±1.4	13.7 ±1.4	ns	p<0.05	ns
Test de Stroop	Stroop color naming	**Time**	**75.2 ±16.2**	**73.1 ±22.8**	**66.5 ±9.9**	**61.6 ±9.4**	**64.1 ±11.2**	**60.8 ±9.3**	**p<0.05**	**p<0.01**	**ns**
	Stroop word reading	**Time**	**51.4 ±9.2**	**55.4 ±24.4**	**48.5 ±7.5**	**50.3 ±14.4**	**48.4 ±8.4**	**46.8 ±7.3**	**ns**	**ns**	**ns**
	Stroop interference	**Time**	**50.2 ±14.9**	**59.7 ±29.9**	**38.2 ±15.8**	**44.6 ±26.3**	**40.2 ±16.5**	**38.4 ±14.2**	**p<0.05**	**p<0.05**	**ns**
Trail-Making Test	Part A	**Time**	**34.1 ±8.1**	**44.4 ±25.9**	**30.2 ±9.9**	**34.9 ±11.5**	**28.3 ±10.6**	**31 ±8.3**	**ns**	**p<0.05**	**ns**
	Part B	**Time**	**80.8 ±17.1**	**89.6 ±32.8**	**63.5 ±16.5**	**69.6 ±19.2**	**57.2 ±17.9**	**67.1 ±24.4**	**p<0.01**	**p<0.01**	**ns**
	B-A	**time**	**46.8 ±14.3**	**45.2 ±19.7**	**33.2 ±10**	**34.7 ±17.3**	**28.9 ±15.4**	**36.1 ±19.7**	**p<0.01**	**ns**	**ns**

Forward span (of Baddeley dual task)	Forward span	AA	5.5 ±1	5.1 ±1	5.5 ±1.1	5.8 ±0.9	5.4 ±1.1	5.7 ±1	ns	p<0.05	ns
Baddeley dual task	Dual-task	Dual Indice mu	91.2 ±33.1	89.6 ±38.7	91.1 ±14.4	88.6 ±9.9	90.0 ±9.9	92.3 ±14	ns	ns	ns
Fluency	Semantic	**AA (in 120s)**	**29 ±7.7**	**28.2 ±6.5**	**30.8 ±7.8**	**29.6 ±7.6**	**31.5 ±7.6**	**29.6 ±6.5**	**ns**	**ns**	**ns**
	Phonemic	**AA (in 120s)**	**21.2 ±5**	**18.8 ±5.7**	**21.2 ±4.3**	**20.6 ±5.8**	**22.1 ±4.5**	**21.3 ±6.1**	**ns**	**ns**	**ns**
Empan envers	Backward Span	AA	3.8 ±0.8	3.8 ±0.8	3.7 ±0.9	4.1 ±1.1	4.2 ±1	4.7 ±1	ns	ns	ns
California Verbal Learning Test	Learning trials, list A	AA	62.5 ±8.6	61.2 ±7.7	65.7 ±10.2	63.7 ±6	66.1 ±7.8	67.8 ±6.4	ns	ns	ns
	Learning trials, list B	AA	8.3 ±2.4	7.6 ±2.3	8.3 ±2.3	7.9 ±2.2	8 ±2.9	8 ±2.4	ns	ns	ns
	Immediate recall	AA	13 ±2.8	12.1 ±2.4	13.5 ±2.4	13.7 ±2	14.1 ±2	13.5 ±2.4	ns	p<0.01	ns
	Immediate cued recall	AA	13.5 ±1.8	12.8 ±1.9	13.8 ±1.8	14.3 ±1.7	14.2 ±1.9	14.6 ±1.9	ns	p<0.01	p<0.05
	Delayed recall	AA	13.6 ±1.8	12.9 ±2.5	14.1 ±1.8	14.2 ±1.7	14.4 ±1.8	14.4 ±2	ns	p<0.05	ns
	Delayed cued recall	AA	13.6 ±1.9	13.2 ±2.4	14.2 ±1.8	14 ±1.9	14.6 ±1.5	14.8 ±1.6	ns	ns	ns

	Recognition	AA	15.5 ±0.8	15.2 ±1.5	15.7 ±0.6	15.4 ±1	15.8 ±0.4	15.7 ±0.7	ns	ns	ns
Figure complexe de Rey	Copy	AA	33.5 ±2.6	33.6 ±2.4	33.9 ±2	34.5 ±1.5	33.7 ±1.4	34.7 ±1.1	ns	ns	ns
	Copy	**Time**	**195.9 ±90.2**	**184.2 ±76.5**	**162.7 ±73**	**192.2 ±67.1**	**158.9 ±69.5**	**173 ±55.2**	**p<0.05**	**ns**	**p<0.01**
		AA	19.2 ±6	17.9 ±6.2	33.9 ±2	34.5 ±1.5	33.7 ±1.4	34.7 ±1.1	p<0.05	ns	ns
DO 80	Naming Test	AA	77.9 ±1.3	78.1 ±1.1	78.1 ±1.9	78.2 ±1.5	78.7 ±1.3	79 ±1.3	ns	ns	Ns
California Verbal Learning Test	Learning trials, list A	AA	62.5 ±8.6	61.2 ±7.7	65.7 ±10.2	63.7 ±6	66.1 ±7.8	67.8 ±6.4	ns	ns	ns
	Learning trials, list B	AA	8.3 ±2.4	7.6 ±2.3	8.3 ±2.3	7.9 ±2.2	8 ±2.9	8 ±2.4	ns	ns	ns
	Immediate recall	AA	13 ±2.8	12.1 ±2.4	13.5 ±2.4	13.7 ±2	14.1 ±2	13.5 ±2.4	ns	p<0.01	ns
	Immediate cued recall	AA	13.5 ±1.8	12.8 ±1.9	13.8 ±1.8	14.3 ±1.7	14.2 ±1.9	14.6 ±1.9	ns	p<0.01	p<0.05
	Delayed recall	AA	13.6 ±1.8	12.9 ±2.5	14.1 ±1.8	14.2 ±1.7	14.4 ±1.8	14.4 ±2	ns	p<0.05	ns
	Delayed cued recall	AA	13.6 ±1.9	13.2 ±2.4	14.2 ±1.8	14 ±1.9	14.6 ±1.5	14.8 ±1.6	ns	ns	ns

	Recognition	AA	15.5 ±0.8	15.2 ±1.5	15.7 ±0.6	15.4 ±1	15.8 ±0.4	15.7 ±0.7	ns	ns	ns

Data are expressed as the means ± SD. NSI = Nonspecific training group (MS patients; n= 17); REACTIV= Specific training group (MS patients, n= 18). AA= Accurate Answers; RT= Reaction Time. V0 = baseline; V4= month 4; V8 = month 8.

The data corresponding to the quantitative responses implicating Information Processing Speed (RT, time and AA within a given time) are highlighted in bold.

Table 4: Change in Daily Cognitive Activities Questionnaire:

QUESTION DCAQ	scores in NSI group at baseline Mean ± SD	Mean ± SD of tests scores in REACTIV group baseline	Mean ± SD of tests scores in NSI group v4	Mean ± SD of tests scores in REACTIV group V4	NP tests scores improved in NSI group (Wilcoxon test) (N=17)	NP tests scores improved in REACTIV group (Wilcoxon test) (N=18)
1	27.9±15.0	25.0±21.7	29.7 ±20.9	41.7 ±17.1	0.705	0.029
2	45.6±28.3	40.3±35.5	46.9 ±30.1	58.3 ±28.4	0.527	0.016
3	47.1±27.8	40.3±24.5	57.8 ±29.9	50.0 ±22.7	0.058	0.197
4	47.1±24.8	45.8±30.0	46.9 ±20.2	55.6 ±20.2	0.967	0.202
5	30.9±25.8	25.0±24.3	39.1 ±27.3	48.6 ±31.5	0.161	0.10
6	39.7±29.4	40.3±34.4	45.3 ±29.2	54.2 ±28.8	0.468	0.077
7	54.4±29.6	54.2±36.6	64.1 ±18.2	65.3 ±17.4	0.229	0.219
8	57.4±23.0	51.4±34.8	65.6 ±20.2	69.4 ±23.6	0.202	0.027
9	55.9±27.3	44.4±30.4	59.4 ±22.1	65.3 ±25.9	0.351	0.022
10	44.1±27.3	37.5±26.1	46.9 ±22.1	50 ±24.3	0.566	0.029

11	30.9±20.8	38.9±23.0	37.5 ±22.4	58.3 ±22.7	0.206	0.005
12	39.7±21.8	36.1±26.0	51.6 ±28.1	54.2 ±27.5	0.059	0.01
Total score, 12 Q	43.4±16.4	40.3±23.9	49.2 ±16.9	55.9 ±14.1	0.091	0.005
Total score, 17 Q			54.4 ±19.7	57.9 ±10.3	0.027	0.001

NSI group= Nonspecific training group; REACTIV group= Specific training group; DCAQ= Daily Cognitive Activities Questionnaire.

Table 5: Change in Patient-Related Outcomes

Tests	Subtests	scores in NSI group at baseline Mean ± SD	Mean ± SD of tests scores in REACTIV group baseline	Mean ± SD of tests scores in NSI group v4	Mean ± SD of tests scores in REACTIV group V4	NP tests scores improved in NSI group (Wilcoxon test) (N=17)	NP tests scores improved in REACTIV group (Wilcoxon test) (N=18)
BDI		15.6 ±7.7	16.4 ±7.4	9.5 ±6.7	10.5 ±7.3	0.03	0.019
STAI	STAI A	32.9 ±6.8	36.6 ±12.6	32.2 ±8.4	36.9 ±16.1	0.875	0.983
	STAI B	45.7 ±8	47 ±10.7	39.4 ±9.2	42.5 ±12.1	0.057	0.147
MFIS	Total MFIS	56.5 ±13	54.2 ±13.2	41.7 ±17.1	42.5 ±10.1	0.009	0.008
	Physical fatigue	26.4 ±6.5	25.1 ±5.9	21.3 ±8.4	21.2 ±6.9	0.017	0.03
	Cognitive fatigue	25.2 ±7.4	24.4 ±7.7	17.5 ±9.9	17.2 ±7.9	0.006	0.007
	Social score	4.9 ±2.5	4.7 ±2.5	2.9 ±2.2	4.1 ±2	0.019	0.273
SF36	Physical composite	52.2 ±14.9	50.3 ±17.2	55.8 ±20.5	58.1 ±16.7	0.379	0.064
	Mental composite	53 ±18.6	50.0 ±22.8	57.8 ±20.2	59.9 ±17.7	0.796	0.145

NSI group= Nonspecific training group; REACTIV group= Specific training group; BDI= Beck Depression Inventory; STAI= State-Trait Anxiety Inventory; MFIS= Modified Fatigue Impact Scale.

SUPPLEMENTARY MATERIAL

e-Table 1: Daily Cognitive Activities Questionnaire (DCAQ) scores at baseline in PwMS and HS

Is it difficult for you to do the following?	NSI	REACTIV	HS
1: To maintain your attention for a long time? Do you experience fatigability?	27.9±15.0	25.0±21.6	64.3±16.9***
2: To read a book without going back or follow a full movie?	45.6±28.3	40.3±35.5	85.7±12.7***
3: To pay attention to all the information in your environment and to react rapidly (while driving, for example)?	47.1±27.8	40.3±24.5	71.4±22.8***
4: To resume an activity when you have been interrupted?	47.1±24.8	45.8±30.0	79.8±18.7***
5: To dial a phone number that someone has just given you orally?	30.9±25.9	25.0±24.3	65.5±16.7***
6: To follow a conversation involving several people?	39.7±29.4	40.3±34.4	76.2±14.7***
7: To organize what you have to do (shopping list, vacation plans, meeting arrangements, plans, etc.)?	54.4±29.6	54.2±36.6	84.5±14.7***
8: To manage your official papers?	57.4±23.0	51.4±34.8	79.8±20.3**
9: To check what you did, if necessary?	55.9±27.3	44.4±30.4	77.4±20.8***
10: To formulate your ideas?	44.2±27.3	37.5±26.1	75.0±17.7***
11: To remember something that someone just told you or told you earlier in the day, or an event that occurred in the last few hours or in the last few days?	30.9±20.8	38.9±23.0	71.4±12.0***
12: To remember older events (last month, last year, etc.)?	39.7±21.8	36.1±26.0	65.5±18.5***
Total score, 12 Q	43.4±16.4	40.3±23.9	74.7±10.7***

PwMS= Patients with multiple sclerosis; NSI= Nonspecific training group; REACTIV= Specific training group; HS= Healthy subjects. **: $p<0.01$; ***: $p<0.001$ (p value between PwMS and HS (Mann-Whitney)).

Magnetic resonance imaging (MRI)

Magnetic resonance imaging (MRI) scans were acquired at baseline and after 4 and 8 months on a 3T Achieva TX system (Philips Healthcare, Best, The Netherlands) with an 8-channel phased array head coil. The morphological protocol consisted of 3D T1 weighted MR images acquired using magnetization prepared rapid gradient echo (MPRAGE) imaging (TR= 8.2 ms, TE= 3.5 ms, TI= 982 ms, α=7°, FOV= 256 mm, voxel size=1 mm^3, 180 slices) and 2D multislice FLAIR images (TR= 11,000 ms, TE= 140 ms, TI= 2800 ms, FOV= 230 mm, 45 axial slices, 3 mm thick).

Postprocessing and image analysis were performed as follows. Lesions were segmented by the lesion growth algorithm[21] as implemented in the Lesion Segmentation Tool (LST) version 2.0.15 (http://www.applied-statistics.de/lst.html) in Statistical Parametric Mapping (SPM12). This process results in a lesion probability map that is thresholded to 50% to obtain a binary map of the lesions. These maps were manually corrected by two blinded experts. Using these maps, a lesion-filling algorithm[22] was applied to the T1-weighted images before brain tissue segmentation.

The segmentation procedure for the volumetric analysis of brain structures on T1-weighted images using volBrain (http://volbrain.upv.es) has been described in detail previously.[1] Briefly, after denoising[2] and inhomogeneity correction,[3] images were affine-registered into the Montreal Neurological Institute (MNI) space using Advanced Neuroimaging Tools (ANTs)[4,] and the total brain volume was estimated. The automatic segmentation of gray and white matter using a patch-based multitemplate method has been described.[7] Every mask was then blindly checked and manually

corrected, if needed. To control for variations in head size, each structure's volume was assessed as a fraction of the total intracranial volume (TIV).

MRI metrics between the two groups (NSIG and REACTIV) at baseline were compared using a general linear model (GLM), including age, gender and level of education as covariables.

MRI metrics within each group between different time points were compared using the Wilcoxon test.

RESULTS:

MRI characteristics at baseline and at the two follow-ups are summarized in e-Table 2.

The normalized brain fraction and WM fraction differed significantly between the HS and each MS group (GLM). The median T2 and normalized brain volumes and the GM and WM fractions did not differ between each time point within each group (Wilcoxon test).

REFERENCES:

1. Schmidt P, Gaser C, Arsic M, et al. An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. Neuroimage. 2012;59(4):3774-3783. doi:10.1016/j. Neuroimage.2011.11.032.
2. Prados F, Cardoso MJ, Kanber B, et al. A multi-time-point modality-agnostic patch-based method for lesion filling in multiple sclerosis. Neuroimage. 2016;139:376-384. doi:10.1016/j.neuroimage.2016.06.053.
3. Manjón J V, Coupé P. volBrain: An Online MRI Brain Volumetry System. Front Neuroinform. 2016;10. doi:10.3389/fninf.2016.00030.

4. Manjón J, Coupé P, Martí-Bonmatí L, Collins L. Adaptive non-local means denoising of MR images with spatially varying noise levels. J Magn Reson Imaging. 2010;31(1):192-203. http://www.hal.inserm.fr/inserm-00454564. Accessed September 20, 2017.
5. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 Bias Correction. doi:10.1109/TMI.2010.2046908.
6. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration. Neuroimage. 2011. doi:10.1016/j.neuroimage.2010.09.025.
7. Coupé P, Manjón J V., Fonov V, Pruessner J, Robles M, Collins DL. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. Neuroimage. 2011;54(2):940-954. doi:10.1016/j.neuroimage.2010.09.018.

e-Table 2: MRI metrics

	NSI			REACTIV			HS
	V0	V4	V8	V0	V4	V8	V0
Median T2-Lesion volume (ml)	5.14 [0.71 – 17.29]	5.3 [0.72 – 17.40]	7.91 [0.72 – 17.43]	3.15 [0.45 – 46.43]	4.85 [0.47 – 26.34]	6.71 [0.47 – 46.47]	NA
Normalized brain fraction (%)	83.10 ± 3.89	83.17 ± 4.21	81.65 ± 3.85	81.10 ± 5.14	81.82 ± 4.17	80.30 ± 5.70	86.94 ± 2.99**†††
Normalized WM fraction (%)	34.16 ± 3.72	34.55 ± 3.31	33.54 ± 4.83	33.18 ± 4.65	33.85 ± 2.97	31.90 ± 4.71	37.08 ± 2.66*††
Normalized GM fraction (%)	48.93 ± 2.30	48.62 ± 2.29	48.11 ± 2.76	47.92 ± 2.61	47.97 ± 2.73	48.40 ± 2.68	49.85 ± 2.69

Data are expressed as the means ± SD. MRI= Magnetic resonance imaging; NSI= Nonspecific training group; REACTIV= Specific training group; HS= Healthy subjects; WM= White matter; GM= Gray matter. Percentage: (structure volume/TIV)*100. TIV: Total intracranial volume. NSIG vs HS: *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$. STG vs HS: ††: $p<0.01$, †††: $p<0.001$.

e-Table 3: Correlation between RTs on the Urban DailyCog® and NP scores

	RT Task 1	RT Task 2	RT Task 3 (traffic light)	RT task 3 (boy)
IPS	Stroop words reading time r=0.5**	SDMT r=0.4* Stroop words reading time r=0.4* TMT A (time) r=0.4** TAP Auditory attention: RT simple cond. R=0.4*	Stroop words reading time r=0.6***	SDMT r=0.4** Stroop words reading time r=0.5** TMT A (time) r=0.5**
Memory	CVLT - List A Learning trials list A. r= 0.5* - List A: Immediate recall r= 0.4*	CVLT - List A Learning trials list A. r= 0.4*		
EF	Verbal fluency (semantic) r= 0.4*			

IPS= Information processing speed; TAP= Test of Attentional Performance; RT= Reaction time; EF= Executive function; TMT= Trail Making Test; CVLT= California Verbal-Learning Test. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$.

Significant correlations between RTs on the Urban DailyCog® and NP scores are presented in e-Table 3. No correlations were observed between the Urban DailyCog® RTs and BDI, MFIS and STAI scores. No correlations were observed

between the BDI score and NP scores or between the fatigue cognitive score of the MFIS and NP scores. Very few correlations were observed between STAI scores or the Physical MFIS scores and NP scores: STAI-A and Stroop color naming time ($r=0.4$; $p<0.05$) and CVLT List A Learning trials list A ($r= 0.4$; $p<0.05$); STAI-B and accurate answers in visual scanning with a target ($r=0.4$; $p<0.05$); MFIS physical score and accurate answers in the visual attention simple task ($r=0.5$, $p<0.01$); accurate answers that cued long-term retrieval CVLT ($r=0.4$; $p<0.05$); and n-back accurate answers ($r=0.4$; $p<0.05$).

The DCAQ total score 12Q correlated with some attentional scores of the TAP, as follows: accurate answers in the visual attention simple condition, visual attention double condition RT and accurate answers in the dual-task, and accurate answers in visual scanning with a target ($r=0.4$ and $p<0.05$ for all correlations).

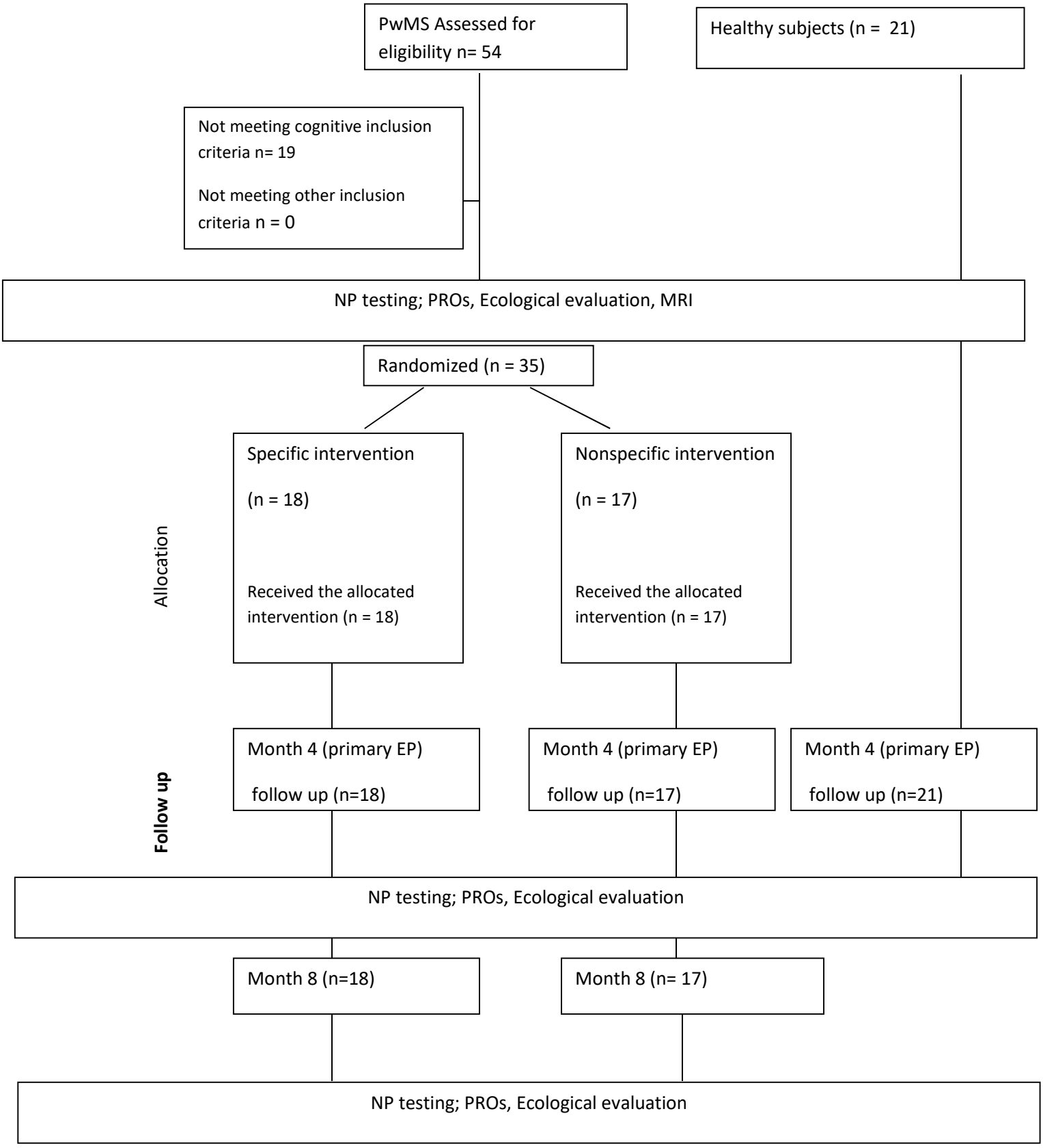

FIGURE 1: FLOW CHART

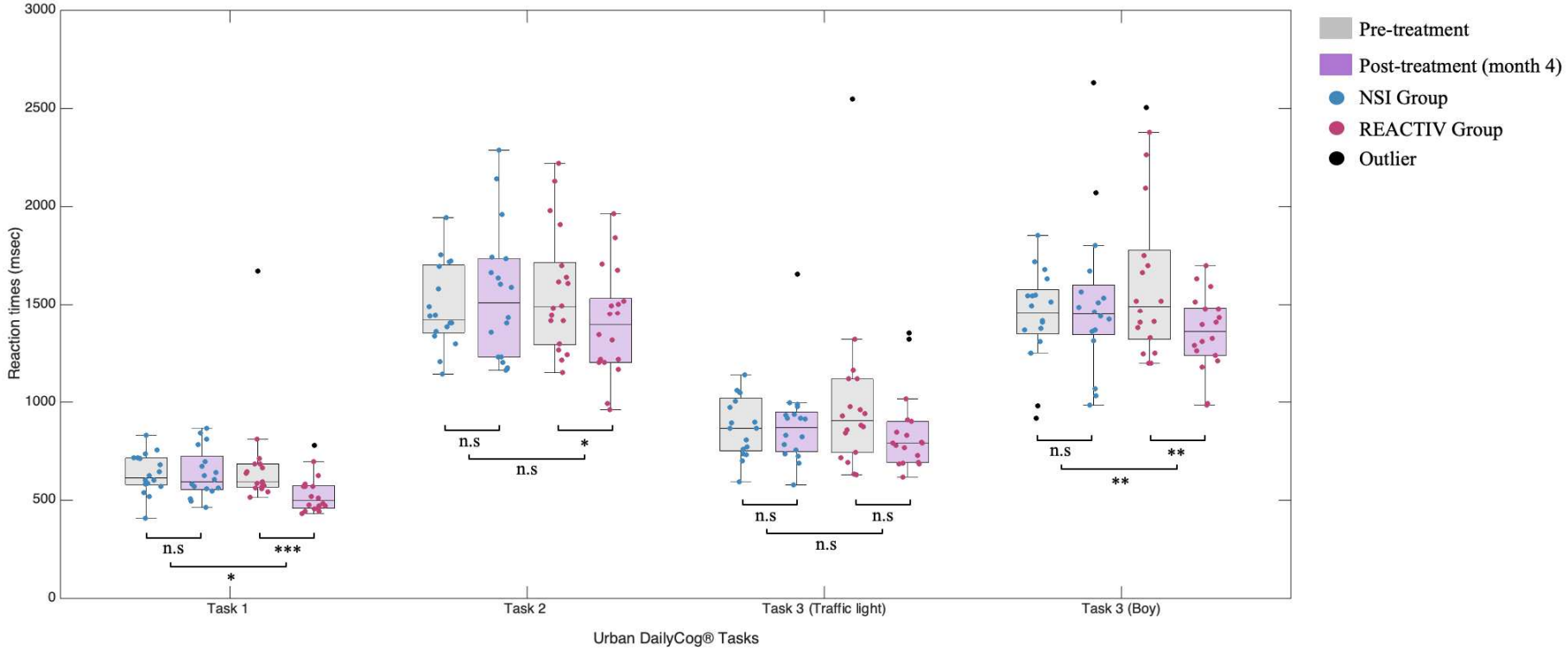

FIGURE 2: Changes in Urban DailyCog® scores from pre- to post-treatment by group (a lower score is better).

Outliers were not excluded from the analysis.