A Mechanistic Insight into Sources of Error of Visual Working Memory in Multiple Sclerosis

Ali Motahharynia,[1,2†] Ahmad Pourmohammadi,[1,2,4†] Armin Adibi,[1,2] Vahid Shaygannejad,[1,2,3] Fereshteh Ashtari,[1,2,3] Iman Adibi,[1,2,3*] and Mehdi Sanayei[1,2,4*]

Author affiliations:

1. Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, 81839 83434, Iran
2. Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, 81839 83434, Iran
3. Department of Neurology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, 81746 75731, Iran
4. School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, 19568 36484, Iran

* Corresponding authors:

Mehdi Sanayei
School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, 1956836484, Iran, Tel: +98 212 611 3274
Email: mehdi.sanayei@gmail.com

Iman Adibi
Department of Neurology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, 81746 75731, Iran, Tel: +98 313 620 1286
Email: i.adibi@gmail.com
ORCID: 0000-0003-4409-9404

[†]These authors contributed equally to this work.

Abstract

Working memory (WM) is one of the most affected cognitive domains in multiple sclerosis (MS), which is mainly studied by the previously established binary model for information storage (slot model). Recent observations based on the continuous reproduction paradigms showed that assuming dynamic allocation of WM resources (resource model) instead of the binary hypothesis will give more accurate predictions in WM assessment. However, they have not been employed in the field of MS despite their influence in uncovering novel mechanistic insights into the WM organization. Here, by utilizing two continuous reproduction paradigms, we investigated WM dysfunction in MS. Also, by applying a computational model, the underlying structure of WM dysfunction was further explored.

A total of 121 patients with MS (61 relapsing-remitting and 60 secondary progressive) and 73 healthy controls were enrolled in this study. The precision of visual WM was measured using memory-guided localization (MGL) and n-back paradigms. The classifying performance of these paradigms in distinguishing different groups was assessed using receiver operating characteristic analysis. Moreover, the sources of error in information recall were evaluated by computational modeling on n-back results.

Our findings determined an overall decrease in recall precision and increased recall variability in MS. While the classifying performance of MGL was better in distinguishing MS subtypes, n-back paradigms were more accurate in discriminating healthy control from relapsing-remitting MS. The applied model showed that decreased signal-to-noise ratio and misbinding error were responsible for WM deficits in these patients.

In conclusion, our results determined the magnitude of WM deficit and demonstrated misbinding error as an important component of WM dysfunction in MS. The dissociable functions of these paradigms in classifying MS subtypes provided evidence about the underlying mechanisms of WM deficits in progressive states of the disease.

Keywords: Binding; Multiple sclerosis; Resource model; Swap error; Working memory

Introduction

Multiple sclerosis (MS) is a debilitating inflammatory disorder characterized by demyelinating central nervous system (CNS) plaques creating a progressive neurodegenerative state with heterogeneous clinical characteristics.[1,2] Impairment in cognitive function is a common clinical manifestation of MS, which detrimentally affects different aspects of patients' daily life, from decreased physical performance and productivity to unemployment.[2,3] One of the frequently affected domains of cognition in MS is working memory (WM),[4] which due to its essential role in several cognitive processes,[5,6] is one of the main areas of MS research.[4,7–10]

Multiple neuropsychological cognitive paradigms, such as paced auditory serial addition test (PASAT), n-back, and delayed-match to sample, were developed to investigate different aspects of WM deficit in MS.[4,9–13] The basis of these change detection paradigms is the slot model of WM.[14] In this quantized model, WM is considered as a short-term storage for a limited number of items,[14–16] in which the information is stored in a binary format. This assumption created an all-or-none condition in which only the stored items in these limited slots will be remembered.[14] Nonetheless, recent observations from analog recall paradigms assessing the precision of WM determined the dynamic allocation of WM resources.[17–20] Each stored item in this framework possesses a fraction of WM storage in which the allocated space changes dynamically between them.[14,17,20] This concept is the foundation for the resource-based model of WM.

The analog nature of inputs in resource-based model paradigms makes it possible to investigate the resolution and variability of stored memory.[21,22] Also, measuring the sources of error using analog cued (connected features) recall paradigms further helped in uncovering the underlying structure of visual WM system.[21,23–26] An analog cued recall task is a paradigm in which subjects need to simultaneously recall an item with its binding feature in a continuous space, hence requiring encoding the information of connected features in addition to their distinct value (e.g., object, location, and object-location binding information).[21,23,24] According to the study of Bays et al., there are three different sources of error for recalling information in visual WM tasks with connected features.[18] They are identified as i: the Gaussian variability in response around the target value (imprecision due to noisiness), ii: Gaussian variability in response around the non-target value, i.e., mistakenly reporting feature of other presented items (misbinding or swap error), and iii: random responses (uniform error).[18,23]

Studies based on analog-cued recall paradigms unraveled new insights into the sources of recall error in neurodegenerative disorders. It was determined that random guessing and misbinding contribute to the impairment of visual WM in Parkinson's and Alzheimer's diseases, respectively.[23,26] However, regardless of the huge impact of these study designs on discovery of novel mechanistic insights into the organization of visual WM system, it has not gotten enough attention in the field of MS research.

The current study followed our previous study in which the quantity of MS-related visual WM was assessed using the slot model.[12] Here we aimed to evaluate recall precision (quality) using the resource model. In this regard, we developed two analog recall paradigms, a memory-guided localization (MGL) and two n-back tasks (1-back and 3-back). Primarily using the simplistic design of MGL, the recall error and precision of visual WM in MS were assessed. Similarly, recall error and precision were evaluated using the two designed n-back paradigms, i.e., the low memory load, 1-back, and high memory load, 3-back conditions, respectively. Moreover, the classifying performance of these paradigms in distinguishing different groups was assessed. Finally, using the probabilistic model of Bays, Catalao & Husain,[18] the sources of recall error in n-back paradigms were further investigated. In both MGL and n-back (low and high memory load conditions) paradigms, recall error and precision were impaired in MS. The dissociable function of these paradigms in classifying MS subtypes (relapsing-remitting and secondary progressive) gave some clues about the underlying structure of WM deficit in progressive states of the disease. Investigation into the sources of error in a high memory load condition revealed that recall fidelity and misbinding, and not random guessing, contribute to visual WM dysfunction in MS.

Results

In the memory-guided localization (MGL) paradigm, 45 patients (19 relapsing-remitting MS, RRMS, and 26 secondary progressive MS, SPMS) and 24 healthy controls participated. The mean data of five participants (one healthy control, three RRMS, and one SPMS) were excluded from further analysis as mentioned in the Method. In the 3-back paradigm, from a total of 76 patients (42 RRMS and 34 SPMS) and 49 healthy control who participated, three healthy control, three RRMS, and two SPMS were excluded. The demographic and clinical data of participants were summarized in Tables 1A and 1B.

Recall Error in Multiple Sclerosis

In the MGL paradigm, using a mixed-model analysis of variance (ANOVA), we evaluated recall error as a function of distance. Recall error was significantly different between groups [$F(2,61) = 14.57$, $P < 10^{-5}$] and distance ([$F(2,61) = 85.03$, $P < 10^{-23}$], Fig. 2A). A significant interaction was also observed between groups and distance [$F(4,61) = 7.24$, $P < 10^{-4}$]. Tukey post hoc test determined that recall error was significantly higher in SPMS [1.86° ± 0.92° visual degree] compared to healthy control [0.97 ± 0.26, $P < 10^{-4}$] and RRMS [1.09 ± 0.27, $P < 10^{-3}$]. No significant difference was detected between RRMS and healthy control [$P = 0.83$]. Similarly, recall error as a function of delay interval was also evaluated. Recall error was significantly different between delay intervals ([$F(4,61) = 18.89$, $P < 10^{-12}$], Fig. 2D). No significant interaction was observed between groups and delay intervals [$F(8,61) = 0.69$, $P = 0.70$].

While reaction time (RT) was significantly different between groups [$F(2,61) = 26.44$, $P < 10^{-8}$] and distance ([$F(2,61) = 25.94$, $P < 10^{-9}$], Fig. 2C), it was not significantly different between delay intervals ([$F(4,61) = 0.97$, $P = 0.43$], Fig. 2F). No significant interaction was observed between groups and distance [$F(4,61) = 2.06$, $P = 0.09$] or groups and delay intervals [$F(8,61) = 0.86$, $P = 0.55$]. The statistical results of RT were summarized in Supplementary Table 1A.

Recall error was evaluated for n-back paradigm using the same method. In the 'high memory load' condition (i.e., 3-back paradigm), recall error was significantly different between groups [$F(2,114) = 28.18$, $P < 10^{-9}$] and bar order ([$F(2,114) = 48.74$, $P < 10^{-17}$], Fig. 2G). No significant interaction was observed between them [$F(4,114) = 1.21$, $P = 0.31$]. Tukey post hoc test showed that recall error was significantly higher in RRMS [0.68 ± 0.20 radian] and SPMS [0.76 ± 0.17] compared to healthy control [0.48 ± 0.16, $P < 10^{-5}$, $P < 10^{-8}$, respectively]. However, no significant difference was detected between RRMS and SPMS groups [$P = 0.14$]. After adjusting for gender, age, and education (they were significantly different between groups), the groups' effect on recall error remained significant (Supplementary Table 2A). Moreover, while RT was also significantly different between groups [$F(2,114) = 12.95$, $P < 10^{-5}$] and bar order ([$F(2,114) = 5.92$, $P < 0.004$], Fig. 2I), no significant interaction was observed between groups and bar order [$F(4,114) = 0.16$, $P = 0.96$]. The statistical results of RT were summarized in Supplementary Table 1B.

In the 'low memory load' condition (i.e., 1-back paradigm), recall error was significantly different between groups ([$F(2,114) = 36.85$, $P < 10^{-12}$], Fig. 2G). Tukey post hoc test showed that recall error was significantly higher in RRMS [0.33 ± 0.12 radian] and SPMS [0.43 ± 0.12] compared to

healthy control [0.22 ± 0.08, $P < 10^{-4}$, $P < 10^{-8}$, respectively]. We also observed a significant difference between RRMS and SPMS [$p < 10^{-3}$]. After adjusting for gender, age, and education, the group effect on recall error remained significant (Supplementary Table 2B). Correspondingly, RT differed significantly between groups ([$F(2,114) = 12.59$, $P < 10^{-4}$], Fig. 2I). The statistical results of RT were summarized in Supplementary Table 1B.

Recall Precision in Multiple Sclerosis

In the MGL paradigm, recall precision was also significantly different between groups [$F(2,61) = 13.74$, $P < 10^{-4}$] and distance ([$F(2,61) = 23.39$, $P < 10^{-8}$], Fig. 2B). No significant interaction was observed between groups and distance [$F(4,61) = 0.91$, $P = 0.46$]. Post hoc analysis determined recall precision was significantly lower in SPMS [0.87 ± 0.52 /º] than in both RRMS [1.32 ± 0.51, $P < 0.039$] and healthy control [1.70 ± 0.60, $P < 10^{-5}$]. Recall precision was not significantly different between RRMS and healthy control [$P = 0.08$]. We also determined the effects of delay intervals on recall precision. In this analysis, groups [$F(2,61) = 14.57$, $P < 10^{-5}$], delay intervals [$F(4,61) = 7.30$, $P < 10^{-4}$], and their interaction [$F(8,61) = 2.44$. $P < 0.02$] had significant effects on recall precision (Fig. 2E). Post hoc analysis showed the same pattern of result [SPMS = 0.84 ± 0.47 /º; RRMS = 1.39 ± 0.46; healthy = 1.63 ± 0.61; SPMS vs. RRMS: $P < 0.005$; SPMS vs. healthy: $P < 10^{-5}$, RRMS vs. healthy: $P = 0.34$]. Similarly, in the 'high memory load' condition, recall precision was significantly different between groups ([$F(2,114) = 25.23$, $P < 10^{-9}$], Fig. 2H) and bar order [$F(2,114) = 20.70$, $P < 10^{-8}$]. However, no significant interaction was observed between them [$F(4,114) = 1.84$, $P = 0.12$]. Besides, while in post hoc analysis, recall precision was significantly higher in healthy control [2.39 ± 0.78 /radian] than in both RRMS [1.68 ± 0.50, $P < 10^{-6}$] and SPMS patients [1.52 ± 0.31, $P < 10^{-7}$], no difference was observed between RRMS and SPMS participants [$P = 0.49$]. After adjusting for gender, age, and education, the effect of groups on recall precision remained significant (Supplementary Table 2A).

Accordingly, the 'low memory load' condition showed the same pattern. Recall precision significantly differed between groups ([$F(2,114) = 25.48$, $P < 10^{-9}$], Fig. 2H). Post hoc analysis determined that recall precision was significantly higher in healthy control [6.10 ± 2.41 /radian] than in RRMS [4.16 ± 1.98, $P < 10^{-4}$] and SPMS [2.95 ± 1.05, $P < 10^{-8}$]. Moreover, there was a significant difference between RRMS and SPMS [$P < 0.031$], which after adjusting for gender, age, and education, remained significant (Supplementary Table 2B).

Recall Variability and the Sources of Recall Error in Multiple Sclerosis

To further investigate the underlying mechanisms of WM impairment in MS, the variability of recall error and sources of this error were assessed. In this regard, the recorded data from n-back paradigms was fitted to a probabilistic model developed by Bays et al.[18] In line with the results of recall error and precision, for the 'high memory load' condition, recall variability was significantly different between groups [$F(2,114) = 26.79$, $P < 10^{-9}$] and bar order ([$F(2,114) = 14.95$, $P < 10^{-6}$], Fig. 3A). At the same time, they had no significant interaction [$F(4,114) = 1.19$, $P = 0.31$]. Recall variability was lower in healthy control [0.51 ± 0.12] than both RRMS [0.69 ± 0.20, $P < 10^{-5}$] and SPMS [0.76 ± 0.15, $P < 10^{-8}$]. There was no difference between RRMS and SPMS in recall variability [$P = 0.16$]. In the 'low memory load' condition, recall variability was affected by groups ([$F(2,114) = 33.07$, $P < 10^{-11}$], Fig. 3A). Again, recall variability was lower in healthy control [0.25 ± 0.10] than both RRMS [0.38 ± 0.14, $P < 10^{-5}$] and SPMS [0.47 ± 0.12, $P < 10^{-8}$]. There was a significant difference between RRMS and SPMS in recall variability [$P < 0.02$]. After adjusting for gender, age, and education, the groups' effect on recall variability remained significant in both high and low memory load conditions (Supplementary Table 2A and 2B).

According to the study by Bays et al.,[18] there are three sources of error for recalling the feature value in visual WM paradigms with connected features (color and orientation in our designed n-back paradigm). The sources of these errors were defined as Gaussian variability in response value around the i: target (recall fidelity, Fig. 3B) and ii: non-target values (misbinding error, Fig. 3C) and iii: random responses (uniform error, Fig. 3D). In the 'high memory load' condition, recall fidelity was significantly different between groups [$F(2,114) = 11.04$, $P < 10^{-4}$] and bar order ([$F(2,114) = 9.10$, $P = < 10^{-4}$], Fig. 3B). No significant interaction was observed between groups and bar order [$F(4,114) = 0.43$, $P = 0.78$]. Recall fidelity was higher in healthy control [0.88 ± 0.11] than both RRMS [0.79 ± 0.12, $P < 0.003$] and SPMS [0.76 ± 0.14, $P < 10^{-4}$]. There was no difference in recall fidelity between RRMS and SPMS groups [$P = 0.54$]. After adjusting for gender, age, and years of education, the effect of groups on recall fidelity remained significant (Supplementary Table 2A). Moreover, to further evaluate the effect of orientation on results, the nearest neighbor analysis was performed. Removing the effect of the binding process allowed us to assess the isolated effect of orientation. The findings from the nearest neighbor analysis showed

a similar pattern of results (Supplementary Fig.1). The isolated effect of orientation was significant between groups [$F(2,114) = 29.26$, $P < 10^{-10}$] among different bar order [$F(2,114) = 7.07$, $P = < 10^{-2}$]. No significant interaction was observed between group and bar orders [$F(4,114) = 0.58$, $P = 0.67$]. After adjusting for gender, age, and education, the results of nearest neighbor analysis remained significant (Supplementary Table 2A).

In the 'low memory load' condition, groups significantly affected recall fidelity ([$F(2,114) = 3.11$, $P < 0.049$], Fig. 3B). While recall fidelity of healthy control [0.98 ± 0.03] was significantly higher than SPMS patients [0.94 ± 0.09, $P < 0.04$], after adjusting for gender, age, and years of education, this effect became insignificant (Supplementary Table 2B). In addition, recall fidelity was not significantly differed in healthy vs. RRMS [0.96 ± 0.08, $P = 0.48$] and RRMS vs. SPMS comparisons [$P = 0.37$].

In line with the above findings, binding process was also affected in MS. In the 'high memory load' condition, misbinding error was significantly different between groups [$F(2,114) = 7.11$, $P < 0.002$] and bar order ([$F(2,114) = 31.05$, $P = < 10^{-11}$], Fig. 3C). No significant interaction was observed between groups and bar order [$F(4,114) = 1.45$, $P = 0.22$]. The misbinding error was lower in healthy control [0.07 ± 0.06] than in both RRMS [0.11 ± 0.09, $P < 0.05$] and SPMS patients [0.14 ± 0.09, $P < 0.002$]. There was no difference in misbinding error between RRMS and SPMS [$P = 0.41$]. After adjusting for gender, age, and years of education, the groups' effect on misbinding error remained significant (Supplementary Table 2A). Moreover, while in the 'high memory load' condition, the uniform error, i.e., random guessing, was different between groups ([$F(2,114) = 5.50$, $P < 0.006$], Fig. 3D), no such differences were observed between bar orders [$F(2,114) = 0.81$, $P = 0.45$) or the interaction between them [$F(4,114) = 0.18$, $P = 0.95$]. Post hoc analysis revealed that uniform error was lower in healthy control [0.05 ± 0.08] than both RRMS [0.09 ± 0.08, $P < 0.03$] and SPMS [0.10 ± 0.08, $P < 0.02$]. Moreover, there was no difference in uniform error between RRMS and SPMS groups [$p = 0.95$]. Additionally, although after adjusting for gender, the effect of groups on uniform error remained significant, adding age and years of education made this effect insignificant (Supplementary Table 2A). The result of a uniform error in the 'low memory load' condition is mathematically the same as recall fidelity (uniform error = 1 - recall fidelity, since there was no misbinding error in the 1-back condition).

Dissociable Function of MGL and n-back Paradigms

The classifying ability of MGL and n-back paradigms in differentiating healthy control from MS patients was assessed by receiver operating characteristic (ROC) analysis (Fig. 4A-C). The accuracy of MGL, 3-back, and 1-back paradigms in differentiating MS patients from healthy participants was 80% (Fig. 4A), 83.4% (Fig. 4B), and 86.2% (Fig. 4C), respectively. A closer look at Fig. 2A, 2D, and 2G suggested that these paradigms differentiate MS and healthy control with distinct patterns. Hence, we separately applied ROC analysis to healthy control vs. RRMS, healthy control vs. SPMS, and RRMS vs. SPMS for MGL (Fig. 4D), 3-back (Fig. 4E), and 1-back (Fig. 4F) paradigms. While the MGL paradigm had good accuracy in differentiating SPMS from healthy control [90.1%] and SPMS from RRMS [84.7%], it had poor ability in distinguishing healthy control from RRMS ([64.1%], Fig. 4D). Accordingly, although the 3-back paradigm also had good accuracy in differentiating healthy control from SPMS [88.4%] and better results (compared to MGL) in discriminating healthy control from RRMS [79.3%], it did a poor job in discriminating MS subtypes ([62%], Fig. 4E). Complementary to the above findings, the 1-back paradigm showed an in-between pattern of results. The 1-back paradigm accurately discriminates healthy control from SPMS [94.4%], while it also had good performance in differentiating healthy control from RRMS [79.6%]. However, compared to MGL, it had a weaker ability to discriminate MS subtypes ([72.3%], Figure 4F).

Discussion

In the present study, we investigated the visual working memory (WM) deficits in MS using two continuous reproduction paradigms, memory-guided localization (MGL) and n-back. Our results align with the previous reports regarding WM deficit in MS.[4,10–13] Complementary to these reports, which envisaged a binary model for storing information (slot-based model), we assessed precision and variability in WM using analog reproduction paradigms (resource-based model). Also, by utilizing the unique design of our n-back paradigm, which could assess the sources of error in information recall, we introduced new mechanistic insight into the visual WM dysfunction in MS. Although the results from both MGL and n-back paradigms showed an overall decrease in recall precision and increased recall error in MS, the post hoc analysis demonstrated inconsistent results. In the MGL paradigm, while SPMS patients performed worse than other groups, no significant

difference was observed between healthy control and RRMS. This result contrasted with the 3-back paradigm (high memory load condition) in which MS subtypes (RRMS vs. SPMS) were not significantly different, and healthy control performed better than RRMS and SPMS. The situation also varied for the 1-back paradigm in which all three groups performed with different levels of precision. ROC analysis further confirmed these results, which determined a dissociation between the classifying performance of MGL and n-back paradigms. Moreover, although the low memory load condition was better than the high memory load condition in distinguishing MS subtypes, its classifying performance was not as well as the MGL paradigm. It seems that these paradigms evaluated distinct aspects of WM dysfunction in MS.

The dissociable function of our paradigms could arise from using different types of stimuli (location in MGL vs. orientation in n-back) in which spatial WM was assessed in the MGL paradigm. As the spatial WM process was associated with the function of the hippocampus,[27] the observed difference could indicate more hippocampal disruption in SPMS. This finding is in line with previous studies that showed more hippocampal regional loss[28] and increased hippocampus neuroinflammatory activity in SPMS compared to RRMS,[29] suggesting that examination of the spatial WM could be a specific marker for disorganization of the WM system in SPMS.

Another explanation is the long delay intervals in MGL, which assessed the maintenance of information. Therefore, the observed difference could be due to additional impairment of SPMS patients in keeping that information. This is in line with our previous study, which showed that change detection paradigms with long delay intervals were promising in differentiating multiple sclerosis subtypes.[12] At the same time, one may debate that this difference was related to the longer stimulus presentation time in the MGL paradigm (1000ms vs. 500ms in -n-back). However, since the stimulus presentation time was adequate in n-back and the stimuli were presented sequentially, it did not seem that the inadequate time for encoding information was responsible for this difference.[21,22] Moreover, the different patterns of eye movement could also affect the results. However, due to the centrality and small size of stimuli in n-back (2.57°), which did not require eye movement, and the similar patterns of results for distance and delay conditions in the MGL paradigm in which the effect of distance was removed in delay condition analysis (each distance was uniformly distributed for each delay interval), it seems less likely that the distinct patterns of eye movement were responsible for this difference.

Additionally, one may argue that the observed dissociation could be due to the extra binding process needed in the 3-back paradigm. However, the results from the low memory load condition and our findings from the 3-back nearest-neighbor analysis, which provided a proxy to assess the isolated effect of orientation, demonstrated a similar pattern of dissociation in the absence of a binding effect. Based on these findings, we concluded that the binding process was not responsible for the observed dissociation. Yet, since the evaluated binding process was an intra-term association (i.e., conjunctive binding), we could not be assured that the same results would be reached for an inter-term association (i.e., relational binding).[30] This issue becomes more interesting when we realize that the relational binding function is mainly centered on the hippocampus,[30–32] the structure we presumed was responsible for the observed dissociation in the MGL paradigm.

Finally, due to the diffuse pattern of involved brain areas in MS and evidence demonstrating that brain networks accounted for different WM processes, it is reasonable to assume that distinct WM-related networks, instead of a single region, were responsible for the observed dissociable patterns.[1,4,25,31,33]

To further assess the underlying structure of visual WM dysfunction, we applied a probabilistic model to uncover the sources of recall error in the n-back paradigm. Our finding from the 3-back paradigm showed that in addition to error due to noisiness, the inability to bind objects' information properly was affected in MS. This failure to bind information was seen in various neurological disorders, including different types of Alzheimer's disease,[23,26,34,35] epileptic patients with temporal lobe lobectomy,[36] and voltage-gated potassium channel complex antibody (VGKC-Ab) limbic encephalitis.[37] Based on the assessed type of binding process, these studies proposed impairment of hippocampal and medial temporal lobe regions in relational and occipital-parietal regions in conjunctive-based paradigms.[23,26,34–37] Accordingly, our finding was in line with the studies that showed common patterns of involvement between brain areas associated with conjunctive processing and WM-affected regions in MS (i.e., superior and inferior parietal lobule).[4,38] Additionally, based on the evidence showing the involvement of hippocampal regions in MS,[28,39] we also expect to see relational binding impairment; however, the current study design did not allow us to evaluate this condition. Eventually, the insignificant results from the low memory load condition suggested more impairment in visual WM under high memory load situations, which was not unexpected.

Despite all these findings, our study had some limitations. The current study only assessed WM dysfunction using behavioral paradigms. Further structural and functional evaluations should be performed to confirm our suggested brain areas associated with conjunctive binding and spatial WM in MS. Simultaneous assessment of brain networks using fMRI and EEG or volumetric studies alongside behavioral paradigms could address this issue. Furthermore, although we hypothesized that relational binding could be a specific marker for the progressive state of the disease based on the more disruption of hippocampal-related areas in SPMS, the current study design did not allow us to evaluate this assumption. Future studies aiming to assess the source of WM deficit regarding relational binding could address this issue. Finally, considering the aim of this study, i.e., developing a practical apparatus for WM assessment in clinical settings, we did not use an eye tracker. Although, due to the reasons mentioned earlier, it was unlikely that impaired eye movement influenced the results, considering the possibility of eye movement dysfunction in MS,[12] this effect could not be excluded. Hence, further investigation with an eye tracker is necessary to confirm this issue.

In summary, using different analog recall paradigms, we demonstrated that recall precision and variability were impaired in multiple sclerosis. We provided some evidence regarding the progressive state of the disease by evaluating the underlying mechanisms related to the dissociable behavior of these paradigms. Furthermore, by applying a computational model capable of evaluating the sources of WM dysfunction, we elucidated that decreasing signal-to-noise ratio and patients' difficulty in associating distinct features together were responsible for WM deficit in MS. In conclusion, this study provided a sensitive measure for assessing WM impairment and gave new insight into the organization of WM dysfunction in MS.

Materials and methods

1. Participants

A total of 121 patients with confirmed MS (61 RRMS, and 60 SPMS), based on the 2017 McDonald criteria,[40] and 73 healthy control volunteers participated in this study. Participants were recruited from the Comprehensive Multiple Sclerosis Clinic, Kashani Hospital, Isfahan, between February 2021 and January 2022. The patients' ages were from 18 to 55 years old, were diagnosed

between 1-18 years before entering the study, and had an expanded disability status scale (EDSS) score of 0-6.5. They had no history of clinical relapse or corticosteroid therapy within 2 months before entering the study. The control group ages were 21-59 years old and did not have a family history of MS in their first-degree relatives. Participants had no history of major neurologic or psychiatric disorders or drug and alcohol abuse. They had normal or corrected-to-normal visual acuity, color vision, and normal performance in the Nine-Hole Peg Test (9-HPT, < 45 seconds). Written informed consent was taken from all participants before the start of the study. This study followed the latest update of the Declaration of Helsinki[41] and was approved by the Iranian national committee of ethics in biomedical research (Approval ID: IR.MUI.MED.REC.1400.441).

2. Visual working memory paradigms

Visual WM was assessed using two analog recall paradigms, memory-guided localization (MGL) and n-back. Stimuli were presented on a 15″ cathode ray tube (CRT, 75Hz refresh rate) monitor at a distancing view of 48cm. The paradigms were run in a dimly lit room on a computer with a Linux operating system and MATLAB software (MATLAB 2019a, The MathWorks, Inc., Natick, Massachusetts, USA) with Psychtoolbox 3 extension.[42,43]

2.1. Memory-guided localization

Each trial was initiated by presenting a central fixation point (diameter of 0.51°) for 2 seconds (s), followed by the presentation of a target (a filled green circle with a diameter of 1.29°) for 1s. The target randomly appeared at different eccentricities (3.22°, 6.44°, or 9.66°) on each trial. In each block, targets were presented with equal probability at each eccentricity in random order (pseudo-random selection). While encouraging participants to maintain fixation on the central fixation point, participants were asked to memorize the location of the target circle for a delay period of 0.5, 1, 2, 4, or 8s (chosen pseudo-randomly). After the delay period, the fixation point changed from a circle to a cross, indicating the end of the delay period. Participants were asked to locate the target's position using the computer mouse and confirm their response by pressing the left button on the mouse. Subsequently, visual feedback was presented, showing them the correct position of the target and their response (Fig. 1A). Participants completed six blocks of 30 trials. They also completed a 10-trial training block before the start of the study. Recall error, Euclidian

distance between the target's location and subject response in visual degree, and reaction RT were recorded for further assessment.

2.2. n-back

Two designs of n-back paradigms, i.e., the low memory load condition (1-back) and high memory load condition (3-back), were developed to evaluate the visual WM deficit in MS. In the 'high memory load' condition, each trial started with a small central fixation point (0.26°) for 2s, followed by a sequence of three distinguishable colored bars (red, green, and blue) at the center of the screen in a pseudo-random order. Each bar (2.57° by 0.19°) was presented for 500ms, followed by a 500ms blank interval. The minimum angular difference between the consecutively presented bars was 10 degrees. Participants were asked to memorize both the orientation and color of the presented bars. After the bars were presented, a single bar, the "probe bar", with the color of one of the presented bars was displayed. Participants were asked to adjust the orientation of the probe bar, presented vertically, to match to the orientation of the bar with the same color (target bar). To do that, they used a computer mouse and confirmed their response by clicking the right button. They received visual feedback, which showed the correct orientation of the target bar, their response, and the difference between them in angular degree (Fig. 1B). High memory load condition consisted of six blocks, each with 30 trials. The 'low memory load' condition had the same structure as the high memory load condition except for presenting one bar instead of three (1-back instead of 3-back, Fig. 1C). After the high memory load condition, subjects participated in 30 trials of the low memory load condition. Due to the 1-back design of low memory load condition, the misbinding error was not present, so fewer trials were needed.[21] Before starting the paradigm, they also participated in a 10-trial training block with the same structure as the low memory load condition. The orientation of presented bars, subject's response, recall error (angular difference between the target value and response), and RT were recorded.

3. Statistical analysis

Statistical analyses were conducted using IBM SPSS Statistics for Mac, version 26 (IBM Corp., Armonk, N.Y., USA). The values were reported as mean ± standard deviation (SD). Data with extreme outliers (values greater than 3rd quartile + 3×interquartile range or less than 1st quartile –

411 3×interquartile range) in MGL and 3-back were excluded from further analysis. The level of
412 significance was set at p-value < 0.05.
413 Clinical and demographic profiles of the participants were compared using one-way ANOVA or
414 Kruskal-Wallis H test (three groups comparison), independent sample t-test or Mann-Whitney U
415 test (two groups comparison), and Chi-squared test (gender comparison). The post hoc Tukey's
416 and Dunn's multiple comparison tests were performed following the significant results of ANOVA
417 and Kruskal-Wallis H test. Also, Bonferroni correction for multiple tests was performed following
418 Dunn's post hoc analysis, and the adjusted p-value was reported.
419 For the MGL paradigm, recall error, recall precision (defined as the reciprocal of the standard
420 deviation of recall error), and RT were compared between groups (Healthy, RRMS, and SPMS)
421 among different conditions (distance or delay, mixed model ANOVA, between- and within-
422 subjects comparisons). For the n-back paradigm, since the data was in a circular space, based on
423 the method proposed by Fisher (1993),[44] we used analog report MATLAB toolbox of Bays Lab
424 (https://www.paulbays.com/toolbox/) to calculate circular mean of recall error and recall precision
425 (defined as reciprocal of the circular standard deviation of recall error). For both high and low
426 memory load conditions (3-back and 1-back), recall error, recall precision, and RT were compared
427 between groups with respect to the order of presented bars (mixed model and one-way ANOVA).
428 To further investigate the sources of error and uncover the involved mechanisms in visual WM
429 impairment, the Mixture Model (https://www.paulbays.com/toolbox/), a probabilistic model
430 developed before,[18,20] was utilized. The Mixture Model considers three possible sources for
431 information recall. They are defined as the Gaussian variability in reporting the target and non-
432 target values and random responses.[18,20] In our study, they referred to reporting the orientation of
433 the target bar (recall fidelity), misreporting the orientation of the other two non-target bars instead
434 of the target bar (misbinding error), and random guessing (uniform error). By utilizing the Mixture
435 Model, the probabilities of target, non-target, and random responses and the concentration score
436 of the Von Mises distribution were calculated for each subject. The sources of error were evaluated
437 by comparing the probability of target, non-target, and random responses between groups among
438 different bar orders (mixed model ANOVA). Also, recall variability, defined as the circular
439 standard deviation of the concentration score of von Mises distribution, was assessed using the
440 same method. Moreover, to further evaluate the effect of the binding process and the isolated effect
441 of orientation, nearest neighbor analysis was performed.[37] In this set of analyses, we removed the

binding effect by defining recall error as the difference between the subject response and the nearest presented bar in each trial. The isolated effect of orientation was assessed between groups among different bar orders (mixed model ANOVA). Finally, due to the 1-back design of the low memory load condition, the misbinding error was not present; hence only the recall fidelity, uniform error, and recall variability were compared.

For each comparison, hierarchical regression analyses were performed to evaluate the possible confounding effect of demographic variables (significantly different between groups) on results. Finally, the dissociable function of n-back and MGL paradigms in distinguishing healthy control from MS patients, healthy control from MS subtypes, and MS subtypes from each other was assessed by performing ROC analysis. The AUC was considered as the classifying accuracy.

Data availability

Anonymized data will be available upon request from corresponding authors. The corresponding authors will consider the request against the data-sharing policy in the protocol and ethical approval of the study.

Code availability

The source codes for the analog report MATLAB toolbox are available at https://www.paulbays.com/toolbox/.

Acknowledgment

We sincerely thank Farinaz Tabibian, Leila Sadat Razian, Maryam Mokhtari, and Zahra Mohamad Hoseiny for their cooperation in coordinating patients, Shahrzad Mohammadpour for advice about model fitting, and Mohammadamin Motaharinia for help regarding programming issues and his comments on the first draft of this manuscript.

Funding

This study was supported by Isfahan University of Medical Sciences (grant number: 2400104).

Author contributions

Conceptualization: A.M, A.P, I.A, M.S. Paradigms development: A.M, A.P, M.S. Data acquisition: A.M, A.P, A.A. Statistical analysis: A.M, I.A, M.S. Data interpretation: A.M, A.P, I.A, M.S. Drafting/revising the manuscript: All authors. Supervision: I.A, M.S. Funding Acquisition: V.S, F.A, I.A. All the authors have read and approved the final version for publication.

Competing interest

The authors declare no competing interests.

Supplementary material

Supplementary material is available at *bioRxiv* online.

References

1. Dobson R, Giovannoni G. Multiple sclerosis – a review. *Eur J Neurol*. 2019;26(1):27-40. doi:10.1111/ene.13819
2. Benedict RHB, Amato MP, DeLuca J, Geurts JJG. Cognitive impairment in multiple sclerosis: clinical management, MRI, and therapeutic avenues. *Lancet Neurol*. 2020;19(10):860-871. doi:10.1016/S1474-4422(20)30277-5
3. Clemens L, Langdon D. How does cognition relate to employment in multiple sclerosis? A systematic review. *Mult Scler Relat Disord*. 2018;26:183-191. doi:10.1016/j.msard.2018.09.018
4. Vacchi L, Rocca MA, Meani A, et al. Working memory network dysfunction in relapse-onset multiple sclerosis phenotypes: A clinical-imaging evaluation. *Mult Scler J*. 2017;23(4):577-587. doi:10.1177/1352458516656809
5. Miller EK, Lundqvist M, Bastos AM. Working Memory 2.0. *Neuron*. 2018;100(2):463-475. doi:10.1016/j.neuron.2018.09.023
6. Nee DE, D'Esposito M. The Representational Basis of Working Memory. In: ; 2016:213-230. doi:10.1007/7854_2016_456
7. Hulst HE, Goldschmidt T, Nitsche MA, et al. rTMS affects working memory performance, brain activation and functional connectivity in patients with multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2017;88(5):386-394. doi:10.1136/jnnp-2016-314224
8. Covey TJ, Zivadinov R, Shucard JL, Shucard DW. Information processing speed, neural efficiency, and working memory performance in multiple sclerosis: Differential relationships with structural magnetic resonance imaging. *J Clin Exp Neuropsychol*. 2011;33(10):1129-1145. doi:10.1080/13803395.2011.614597
9. Rocca MA, Valsasina P, Hulst HE, et al. Functional correlates of cognitive dysfunction in multiple sclerosis: A multicenter fMRI Study. *Hum Brain Mapp*. 2014;35(12):5799-5814. doi:10.1002/hbm.22586
10. Costers L, Van Schependom J, Laton J, et al. The role of hippocampal theta oscillations in working memory impairment in multiple sclerosis. *Hum Brain Mapp*. 2021;42(5):1376-1390. doi:10.1002/hbm.25299
11. PARMENTER BA, SHUCARD JL, BENEDICT RHB, SHUCARD DW. Working memory deficits in multiple sclerosis: Comparison between the n -back task and the Paced Auditory Serial Addition Test. *J Int Neuropsychol Soc*. 2006;12(5):677-687. doi:10.1017/S1355617706060826
12. Pourmohammadi A, Motahharynia A, Shaygannejad V, Ashtari F, Adibi I, Sanayei M. Working memory dysfunction differs between secondary progressive and relapsing multiple sclerosis: effects of clinical phenotype, age, disease duration, and disability. *Mult Scler Relat Disord*. Published online November 2022:104411. doi:10.1016/j.msard.2022.104411
13. Stojanovic-Radic J, Wylie G, Voelbel G, Chiaravalloti N, DeLuca J. Neuroimaging and cognition using functional near infrared spectroscopy (fNIRS) in multiple sclerosis. *Brain Imaging Behav*. 2015;9(2):302-311. doi:10.1007/s11682-014-9307-y
14. Ma WJ, Husain M, Bays PM. Changing concepts of working memory. *Nat Neurosci*. 2014;17(3):347-356. doi:10.1038/nn.3655

15. MILLER GA. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev*. 1956;63(2):81-97. http://www.ncbi.nlm.nih.gov/pubmed/13310704
16. Cowan N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci*. 2001;24(1):87-114. doi:10.1017/S0140525X01003922
17. Bays PM, Husain M. Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science (80-)*. 2008;321(5890):851-854. doi:10.1126/science.1158023
18. Bays PM, Catalao RFG, Husain M. The precision of visual working memory is set by allocation of a shared resource. *J Vis*. 2009;9(10):7-7. doi:10.1167/9.10.7
19. Gorgoraptis N, Catalao RFG, Bays PM, Husain M. Dynamic Updating of Working Memory Resources for Visual Objects. *J Neurosci*. 2011;31(23):8502-8511. doi:10.1523/JNEUROSCI.0208-11.2011
20. Schneegans S, Bays PM. No fixed item limit in visuospatial working memory. *Cortex*. 2016;83:181-193. doi:10.1016/j.cortex.2016.07.021
21. Peich MC, Husain M, Bays PM. Age-related decline of precision and binding in visual working memory. *Psychol Aging*. 2013;28(3):729-743. doi:10.1037/a0033236
22. Zokaei N, Burnett Heyes S, Gorgoraptis N, Budhdeo S, Husain M. Working memory recall precision is a more sensitive index than span. *J Neuropsychol*. 2015;9(2):319-329. doi:10.1111/jnp.12052
23. Zokaei N, Sillence A, Kienast A, et al. Different patterns of short-term memory deficit in Alzheimer's disease, Parkinson's disease and subjective cognitive impairment. *Cortex*. 2020;132:41-50. doi:10.1016/j.cortex.2020.06.016
24. McMaster JMV, Tomić I, Schneegans S, Bays PM. Swap errors in visual working memory are fully explained by cue-feature variability. *Cogn Psychol*. 2022;137:101493. doi:10.1016/j.cogpsych.2022.101493
25. Lugtmeijer S, Schneegans S, Lammers NA, et al. Consequence of stroke for feature recall and binding in visual working memory. *Neurobiol Learn Mem*. 2021;179:107387. doi:10.1016/j.nlm.2021.107387
26. Liang Y, Pertzov Y, Nicholas JM, et al. Visual short-term memory binding deficit in familial Alzheimer's disease. *Cortex*. 2016;78:150-164. doi:10.1016/j.cortex.2016.01.015
27. Rolls ET. The storage and recall of memories in the hippocampo-cortical system. *Cell Tissue Res*. 2018;373(3):577-604. doi:10.1007/s00441-017-2744-3
28. Sicotte NL, Kern KC, Giesser BS, et al. Regional hippocampal atrophy in multiple sclerosis. *Brain*. 2008;131(4):1134-1141. doi:10.1093/brain/awn030
29. Cree BAC, Arnold DL, Chataway J, et al. Secondary Progressive Multiple Sclerosis. *Neurology*. 2021;97(8):378-388. doi:10.1212/WNL.0000000000012323
30. Parra MA, Fabi K, Luzzi S, Cubelli R, Hernandez Valdez M, Della Sala S. Relational and conjunctive binding functions dissociate in short-term memory. *Neurocase*. 2015;21(1):56-66. doi:10.1080/13554794.2013.860177
31. Bastin C, Besson G, Simon J, et al. An integrative memory model of recollection and familiarity to understand memory deficits. *Behav Brain Sci*. 2019;42:e281. doi:10.1017/S0140525X19000621
32. Bastin C, Besson G. Aging and binding in short-term memory: processes involved in conjunctive and

relational binding. *Memory*. 2021;29(2):193-209. doi:10.1080/09658211.2021.1873390

33. Figueroa-Vargas A, Cárcamo C, Henríquez-Ch R, et al. Frontoparietal connectivity correlates with working memory performance in multiple sclerosis. *Sci Rep*. 2020;10(1):9310. doi:10.1038/s41598-020-66279-0

34. Della Sala S, Parra MA, Fabi K, Luzzi S, Abrahams S. Short-term memory binding is impaired in AD but not in non-AD dementias. *Neuropsychologia*. 2012;50(5):833-840. doi:10.1016/j.neuropsychologia.2012.01.018

35. Cecchini MA, Parra MA, Brazzelli M, Logie RH, Della Sala S. Short-term memory conjunctive binding in Alzheimer's disease: A systematic review and meta-analysis. *Neuropsychology*. Published online May 26, 2022. doi:10.1037/neu0000825

36. Zokaei N, Nour MM, Sillence A, et al. Binding deficits in visual short-term memory in patients with temporal lobe lobectomy. *Hippocampus*. 2019;29(2):63-67. doi:10.1002/hipo.22998

37. Pertzov Y, Miller TD, Gorgoraptis N, et al. Binding deficits in memory following medial temporal lobe damage in patients with voltage-gated potassium channel complex antibody-associated limbic encephalitis. *Brain*. 2013;136(8):2474-2485. doi:10.1093/brain/awt129

38. Parra MA, Della Sala S, Logie RH, Morcom AM. Neural correlates of shape–color binding in visual working memory. *Neuropsychologia*. 2014;52:27-36. doi:10.1016/j.neuropsychologia.2013.09.036

39. Rocca MA, Barkhof F, De Luca J, et al. The hippocampus in multiple sclerosis. *Lancet Neurol*. 2018;17(10):918-926. doi:10.1016/S1474-4422(18)30309-0

40. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018;17(2):162-173. doi:10.1016/S1474-4422(17)30470-2

41. World Medical Association Declaration of Helsinki. *JAMA*. 2013;310(20):2191. doi:10.1001/jama.2013.281053

42. Brainard DH. The Psychophysics Toolbox. *Spat Vis*. 1997;10(4):433-436. http://www.ncbi.nlm.nih.gov/pubmed/9176952

43. Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis*. 1997;10(4):437-442. http://www.ncbi.nlm.nih.gov/pubmed/9176953

44. Fisher NI. *Statistical Analysis of Circular Data*. Cambridge University Press; 1993. doi:10.1017/CBO9780511564345

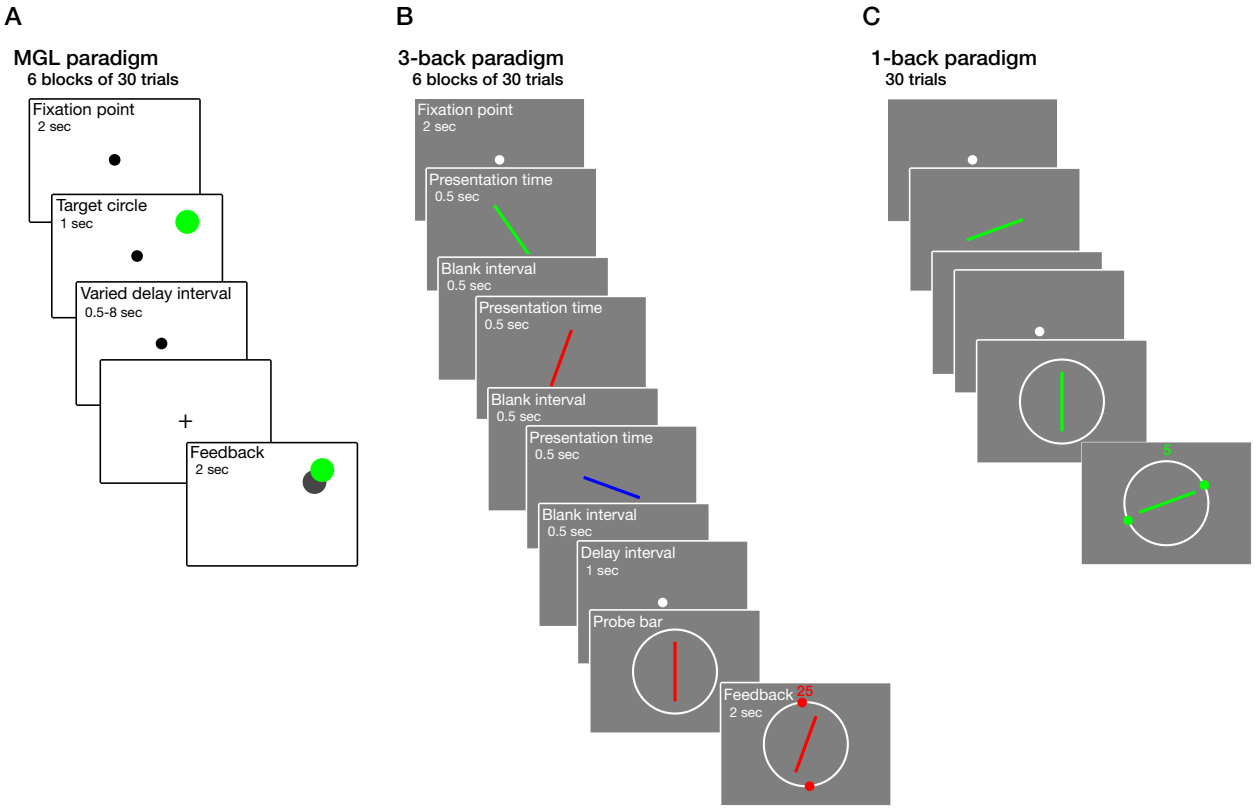

Figure 1. Schematic design of visual working memory (WM) paradigms. (A) In the memory-guided localization (MGL) paradigm, participants were asked to memorize and then localize the position of the target circle following a random delay interval of 0.5, 1, 2, 4, or 8 seconds. Following their response, visual feedback was presented. **(B)** In the 3-back paradigm (high memory load condition), a sequence of three colored bars was presented consecutively. Participants were asked to match the orientation of the probe bar to the previously presented bar with the same color. Visual feedback was displayed following their response. **(C)** The 1-back paradigm (low memory load condition) has the same structure as 3-back except for presenting one bar instead of three.

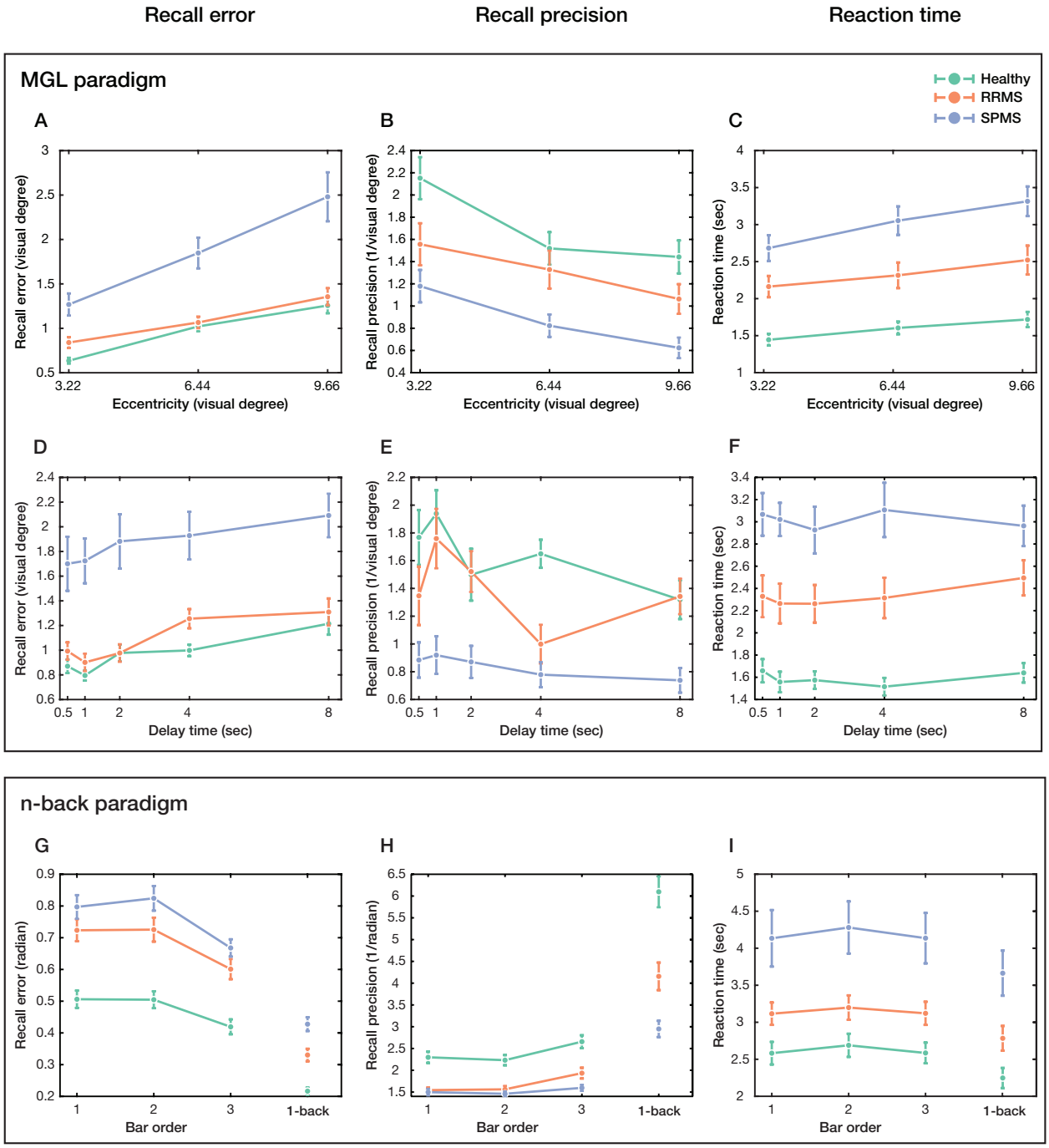

Figure 2. Recall error and precision of healthy control and multiple sclerosis (MS) subtypes (relapsing-remitting, RRMS and secondary progressive, SPMS) in visual WM paradigms. **(A)** Recall error, **(B)** recall precision, **(C)** and reaction time as a function of distance for MGL paradigm. **(D, E, and F)** The same as a function of delay interval. **(G)** Recall error and **(H)** precision and **(I)** reaction time as a function of bar order in 3-back paradigm (left) and 1-back paradigm (right). Data are represented as mean ± SEM.

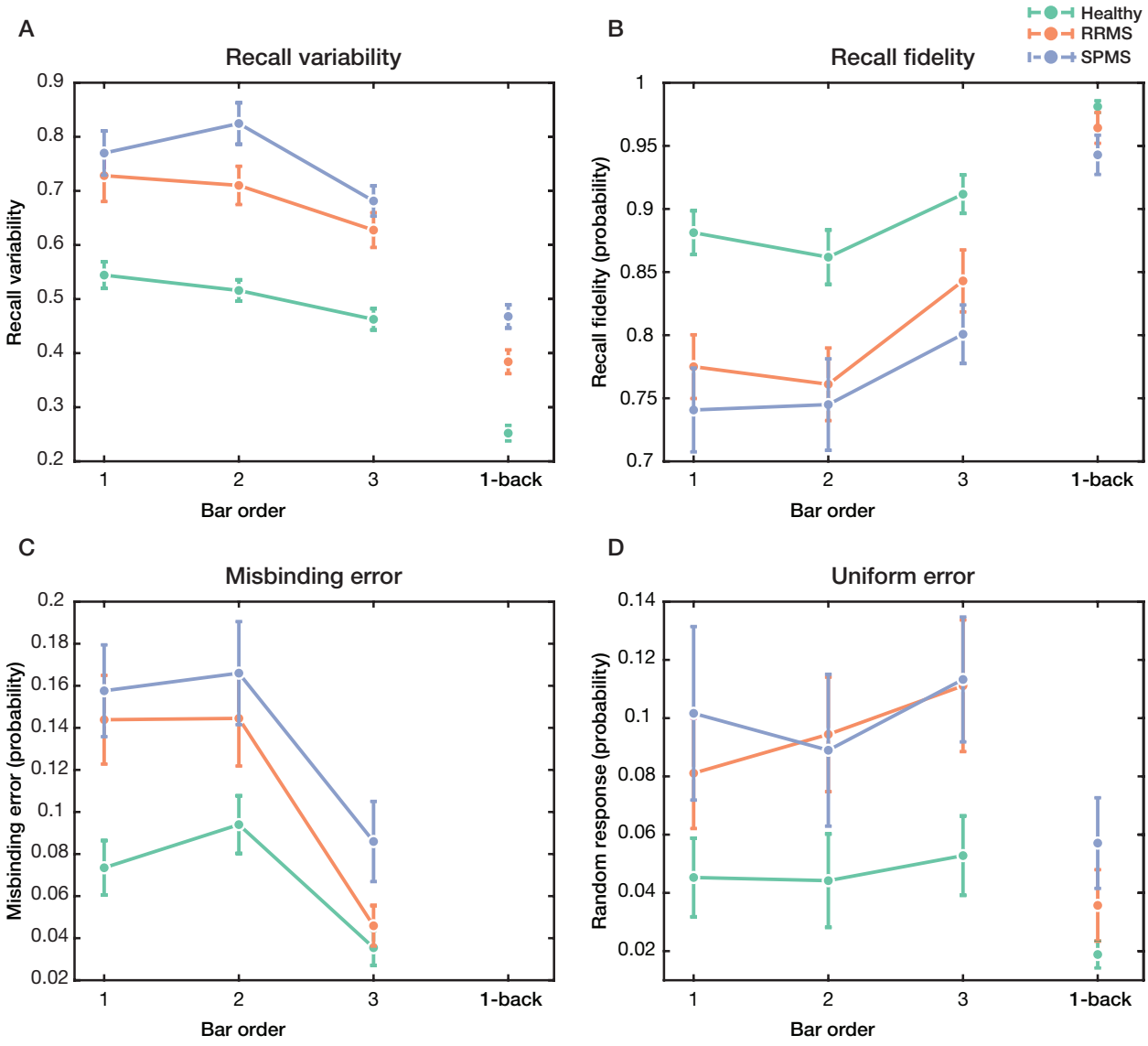

Figure 3. The sources of recall error in high and low memory load conditions (3-back and 1-back, respectively). (A) Recall variability (circular standard deviation of von Mises distribution), **(B)** recall fidelity (probability of response around the target value), **(C)** misbinding error (probability of response around the non-target values), and **(D)** uniform error (probability of random response) for healthy control and MS subtypes in 3-back paradigm (left of each subplot) and 1-back paradigm (right of each subplot). Data are represented as mean ± SEM.

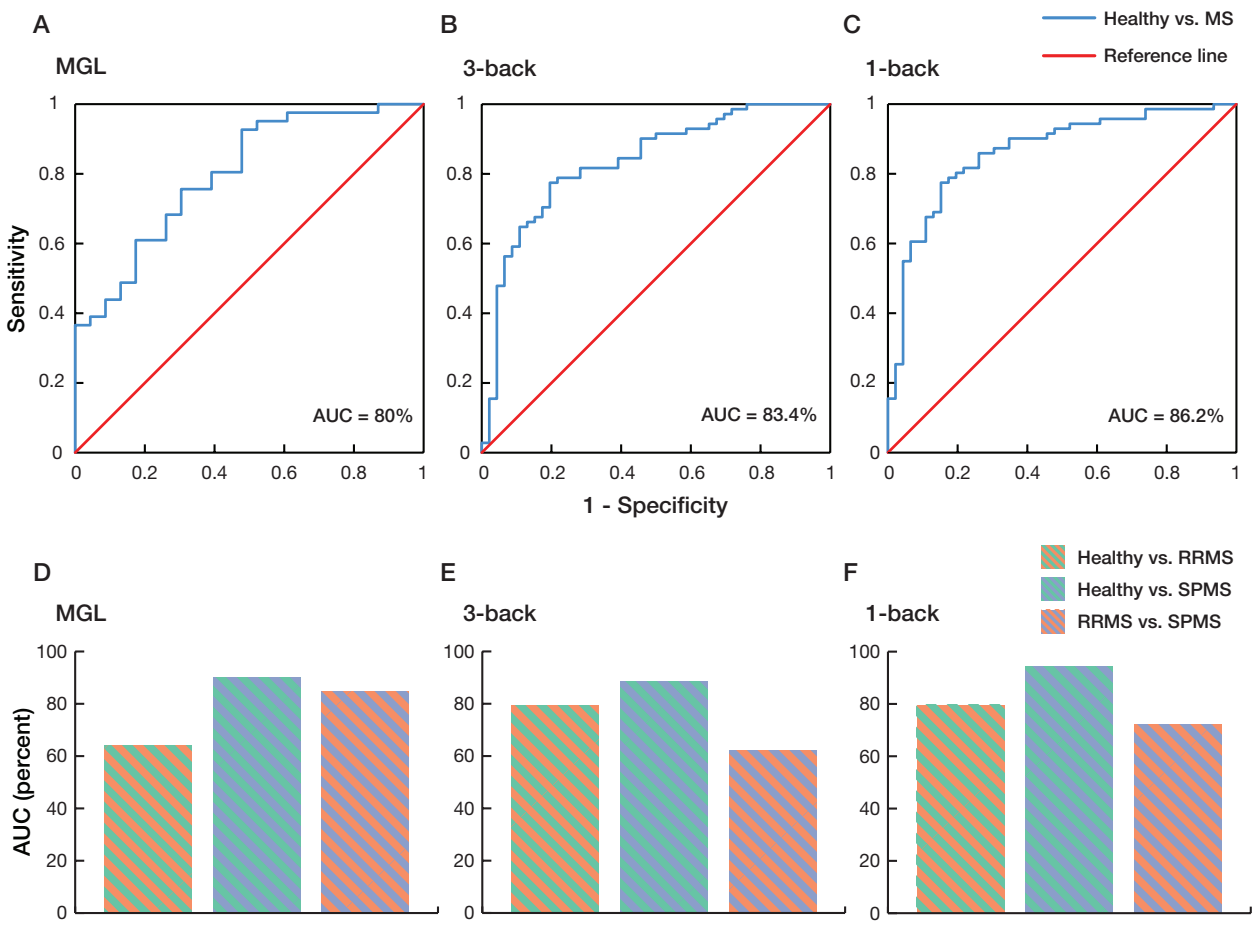

Figure 4. Classifying performance of visual WM paradigms in differentiating healthy control from MS and MS subtypes, and MS subtypes from each other. Receiver operating characteristic (ROC) curve demonstrated the accuracy of **(A)** MGL, **(B)** 3-back, and **(C)** 1-back paradigms in distinguishing healthy control from MS patients. The precision of these paradigms in dissociating healthy control from MS subtypes (RRMS and SPMS) and MS subtypes from each other is represented as the area under curve (AUC) for **(D)** MGL, **(E)** 3-back, and **(F)** 1-back paradigms.

723 **Table 1A. Demographic and clinical profiles of participants in the MGL paradigm**

	HC (n = 23)	RRMS (n = 16)	SPMS (n = 25)	*P*
Gender (F:M)	13:10	14:2	17:8	0.12
Age	35.91 (8.34)	37.25 (6.63)	39.28 (5.56)	0.25
Education (years)	13.30 (2.74)	13.69 (3.34)	13.56 (3.22)	0.86
Disease duration (years)	N/A	8.562 (3.20)	11.56 (3.28)	< 0.02 *
EDSS	N/A	1.28 (0.79)	2.740 (1.23)	< 0.0002 *

724

725 HC = Healthy control, RRMS = Relapsing-remitting multiple sclerosis, SPMS = Secondary progressive multiple sclerosis, EDSS = Expanded
726 disability status scale, N/A = Not applicable
727 ***P* < 0.05**

728

729

730 **Table 1B. Demographic and clinical profiles of participants in the n-back paradigms**

	HC (n = 46)	RRMS (n = 39)	SPMS (n = 32)	*P*
Gender (F:M)	16:30	23:16	22:10	< 0.008 *
Age[a]	30.5 (10.37)	32.03 (6.72)	39.00 (6.43)	$< 10^{-6}$ *
Education[b] **(years)**	16.95 (2.23)	13.87 (3.41)	13.67 (2.73)	$< 10^{-7}$ *
Disease duration (years)	N/A	6.60 (3.84)	9.37 (4.43)	< 0.007 *
EDSS	N/A	1.49 (1.01)	3.86 (1.74)	$< 10^{-7}$ *

731

732 [a]**Dunn's test:** (Healthy vs. RRMS: $P = 0.27$, Healthy vs. SPMS: $P < 10^{-6}$*, and RRMS vs SPMS: $P < 0.005$*)
733 [b]**Dunn's test:** (Healthy vs. RRMS: $P < 10^{-5}$*, Healthy vs. SPMS: $P < 10^{-5}$*, and RRMS vs. SPMS: $P = 1$)
734 ***P* < 0.05**