

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

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2

October 9, 2023 (this version)

Reviewed preprint version 1

July 27, 2023

Sent for peer review

April 6, 2023

Posted to bioRxiv

February 21, 2023

Ali Motahharynia, Ahmad Pourmohammadi, Armin Adibi, Vahid Shaygannejad, Fereshteh Ashtari, Iman Adibi , Mehdi Sanayei 

Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, Iran • Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, Iran • School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran • Department of Neurology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

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). However, recent observations based on the continuous reproduction paradigms have shown that assuming dynamic allocation of WM resources (resource model) instead of the binary hypothesis will give more accurate predictions in WM assessment. Moreover, continuous reproduction paradigms allow for assessing the distribution of error in recalling information, providing new insights into the organization of the WM system. In this study, we utilized two continuous reproduction paradigms, memory-guided localization (MGL) and analog recall task with sequential presentation, to investigate WM dysfunction in MS. Our results demonstrated an overall increase in recall error and decreased recall precision in MS. While sequential paradigms were better in distinguishing healthy control from relapsing-remitting MS, MGL were more accurate in discriminating MS subtypes (relapsing-remitting from secondary progressive), providing evidence about the underlying mechanisms of WM deficit in progressive states of the disease. Furthermore, computational modeling of the results from the sequential paradigm determined that imprecision in decoding information and swap error (mistakenly reporting the feature of other presented items) were responsible for WM dysfunction in MS. In summary, this study offered a sensitive measure for assessing WM deficit and provided new insight into the organization of the WM system in MS population.

eLife assessment

This paper provides **valuable** information regarding visuospatial working memory performance in patients with MS compared to healthy controls, using a relatively novel continuous measure of visual working memory. There are some weaknesses with the way the clinical groups were matched, but those limitations are acknowledged and the strength of evidence for the claims is otherwise **convincing**. The paper will be of interest to those working in the field of clinical neuroscience.

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 (Benedict et al., 2020; Dobson & Giovannoni, 2019). 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 (Benedict et al., 2020; Clemens & Langdon, 2018). One of the frequently affected domains of cognition in MS is working memory (WM) (Vacchi et al., 2017), which due to its essential role in several cognitive processes (Miller et al., 2018; Nee & D'Esposito, 2016), is one of the main areas of MS research (Costers et al., 2021; Covey et al., 2011; Hulst et al., 2017; Rocca et al., 2014; Vacchi et al., 2017).

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 (Costers et al., 2021; PARMENTER et al., 2006; Pourmohammadi et al., 2022; Rocca et al., 2014; Stojanovic-Radic et al., 2015; Vacchi et al., 2017). The basis of these change detection paradigms is the slot model of WM (Ma et al., 2014). In this quantized model, WM is considered a short-term storage for a limited number of items (Cowan, 2001; Ma et al., 2014; MILLER, 1956), storing the information 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 (Ma et al., 2014). Nonetheless, recent observations from analog recall paradigms assessing the precision of WM determined the dynamic allocation of WM resources (Bays et al., 2009; Bays & Husain, 2008; Gorgoraptis et al., 2011; Schneegans & Bays, 2016). Each stored item in this framework possesses a fraction of WM storage in which the allocated space changes dynamically between them (Bays & Husain, 2008; Ma et al., 2014; Schneegans & Bays, 2016). 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 (Peich et al., 2013; Zokaei et al., 2015). Also, assessing the distribution of error in these paradigms further helped in uncovering the underlying structure of the visual WM system (Liang et al., 2016; Lugtmeijer et al., 2021; McMaster et al., 2022; Peich et al., 2013; Zokaei et al., 2020). An analog recall task is a paradigm in which subjects need to simultaneously recall multiple features of items 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) (McMaster et al., 2022; Peich et al., 2013; Zokaei et al., 2020). 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 (Bays et al., 2009). They are identified as i: the Gaussian variability in response around the target value (target response proportion), ii: Gaussian variability in response around the non-target value, i.e., mistakenly reporting feature of other presented items (swap error), and iii: random responses (uniform response proportion) (Bays et al., 2009; Zokaei et al., 2020).

Studies based on analog recall paradigms unraveled new insights into the sources of recall error in neurodegenerative disorders. It was determined that random response and swap error contribute to the impairment of visual WM in Parkinson's and Alzheimer's diseases, respectively (Liang et al., 2016; Zokaei et al., 2020). However, regardless of the influential impact of these study designs on the 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 based on the slot model (Pourmohammadi et al., 2022 [DOI](#)). Here, we aimed to evaluate recall precision (quality) based on the resource model. In this regard, we developed two analog recall paradigms, a memory-guided localization (MGL) and two analog recall tasks with sequential bar presentation (3-bar and 1-bar). Primarily using the simplistic design of MGL, recall error (absolute error) and precision of visual WM were assessed. Similarly, recall error and precision were evaluated using the two designed analog recall paradigms with sequential presentation, i.e., the low memory load, 1-bar, and high memory load, 3-bar conditions, respectively. Furthermore, the classifying performance of these paradigms in distinguishing different groups was assessed. Finally, the distribution of recall error in analog paradigms with sequential presentation was further investigated using the probabilistic model of Bays, Catalao, and Husain (Bays et al., 2009 [DOI](#)). In both MGL and sequential paradigms (low and high memory load conditions), 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 target response proportion and swap error (non-target proportion) 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 materials and methods section. In the sequential paradigm with 3-bar, from a total of 76 patients (42 RRMS and 34 SPMS) and 49 healthy controls who participated, three healthy controls, three RRMS, and two SPMS were excluded. The demographic and clinical data of participants were summarized in **Tables 1** [DOI](#) and **2** [DOI](#).

Recall Error in Multiple Sclerosis

In the MGL paradigm, recall error was evaluated as a function of distance using a mixed-model analysis of variance (ANOVA). Recall error was significantly different between groups [$F(2,61) = 14.57, P < 10^{-5}$] and distances [$F(2,61) = 85.03, P < 10^{-23}$], **Fig. 2A** [DOI](#). A significant interaction was also observed between group and distance [$F(4,61) = 7.24, P < 10^{-4}$]. Tukey post hoc test determined that recall error was significantly higher in SPMS [$1.86^\circ \pm 0.92^\circ$ visual degree] compared to healthy control [$0.97 \pm 0.26, P < 10^{-4}$] and RRMS [$1.09 \pm 0.27, P < 10^{-3}$]. 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** [DOI](#). No significant interaction was observed between group and delay interval [$F(8,61) = 0.69, P = 0.70$]. After adjusting for cognitive abilities, assessed by the Montreal cognitive assessment (MoCA) screening tool (cognitive performance was significantly different between groups), the effect of group on recall error remained significant (Supplementary Table 1).

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

	HC (n = 23)	RRMS (n = 16)	SPMS (n = 25)	<i>P</i>
Gender (F:M)	13:10	14:2	17:8	0.12
Age (year)	35.9 ± 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
Cognitive ability (NL:MCI)	23:0	14:2	19:6	< 0.05 *
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 *
DMT (Platform: Non-platform)	N/A	2:14	0:25	0.07

All data, except for the categorical information, are presented as mean ± standard deviation.

HC = Healthy control, RRMS = Relapsing-remitting multiple sclerosis, SPMS = Secondary progressive multiple sclerosis, NL = Normal, MCI = Mild cognitive impairment, EDSS = Expanded disability status scale, DMT = Disease modifying therapy, N/A = Not applicable

**P* < 0,05

Table 1.

Demographic and clinical profiles of participants in the MGL paradigm.

	HC (n = 46)	RRMS (n = 39)	SPMS (n = 32)	<i>P</i>
Gender (F:M)	16:30	23:16	22:10	< 0.008 *
Age^a (year)	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 *
Cognitive ability[#] (NL:MCI)	42:4	28:11	18:13	< 0.003 *
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 *
DMT (Platform: Non-platform)	N/A	5:34	1:31	0.14

All the data, except for the categorical information, are presented as mean ± standard deviation.

HC = Healthy control, RRMS = Relapsing-remitting multiple sclerosis, SPMS = Secondary progressive multiple sclerosis, NL = Normal, MCI = Mild cognitive impairment, EDSS = Expanded disability status scale, DMT = Disease modifying therapy, N/A = Not applicable

[#]One value for SPMS is missing.

^aDunn's test: (Healthy vs. RRMS: $P = 0.27$, Healthy vs. SPMS: $P < 10^{-6}$ *, and RRMS vs SPMS: $P < 0.005$ *)

^bDunn's test: (Healthy vs. RRMS: $P < 10^{-5}$ *, Healthy vs. SPMS: $P < 10^{-5}$ *, and RRMS vs. SPMS: $P = 1$)

* $P < 0.05$

Table 2.

Demographic and clinical profiles of participants in the sequential paradigms.

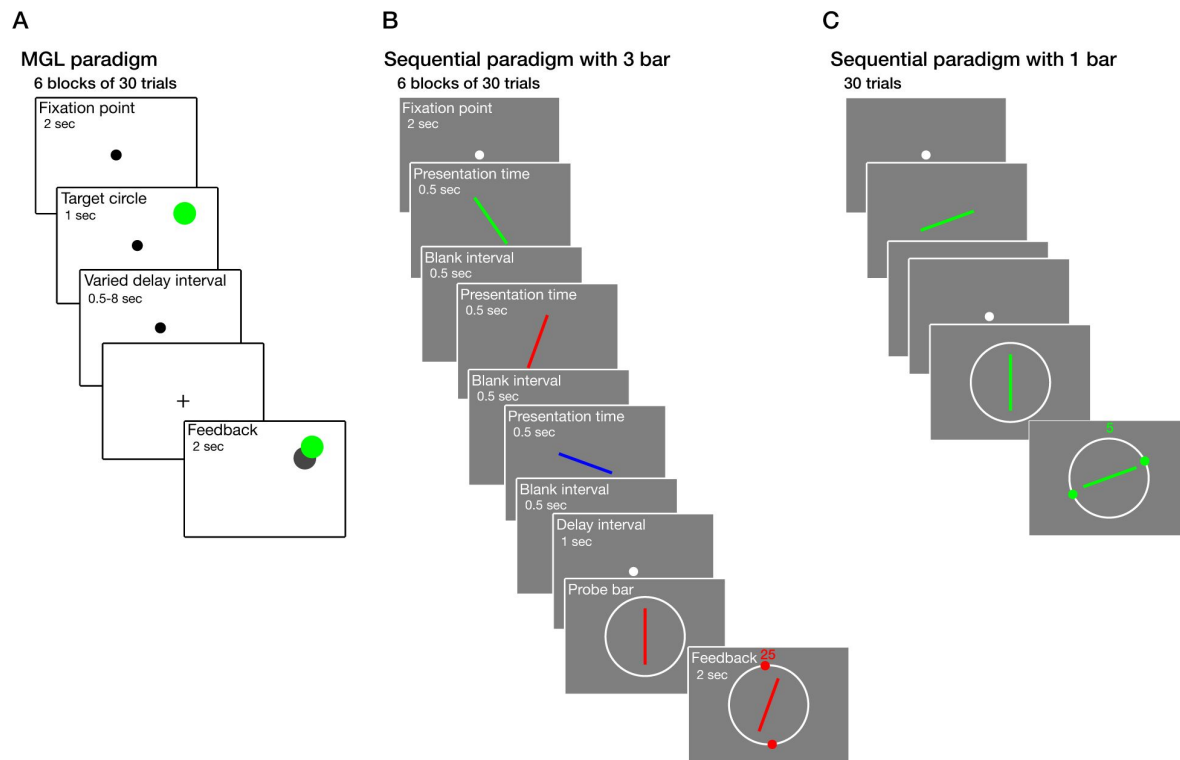


Figure 1.

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 sequential paradigm with 3 bar (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-bar paradigm (low memory load condition) has the same structure as the 3-bar paradigm except for presenting one bar instead of three.

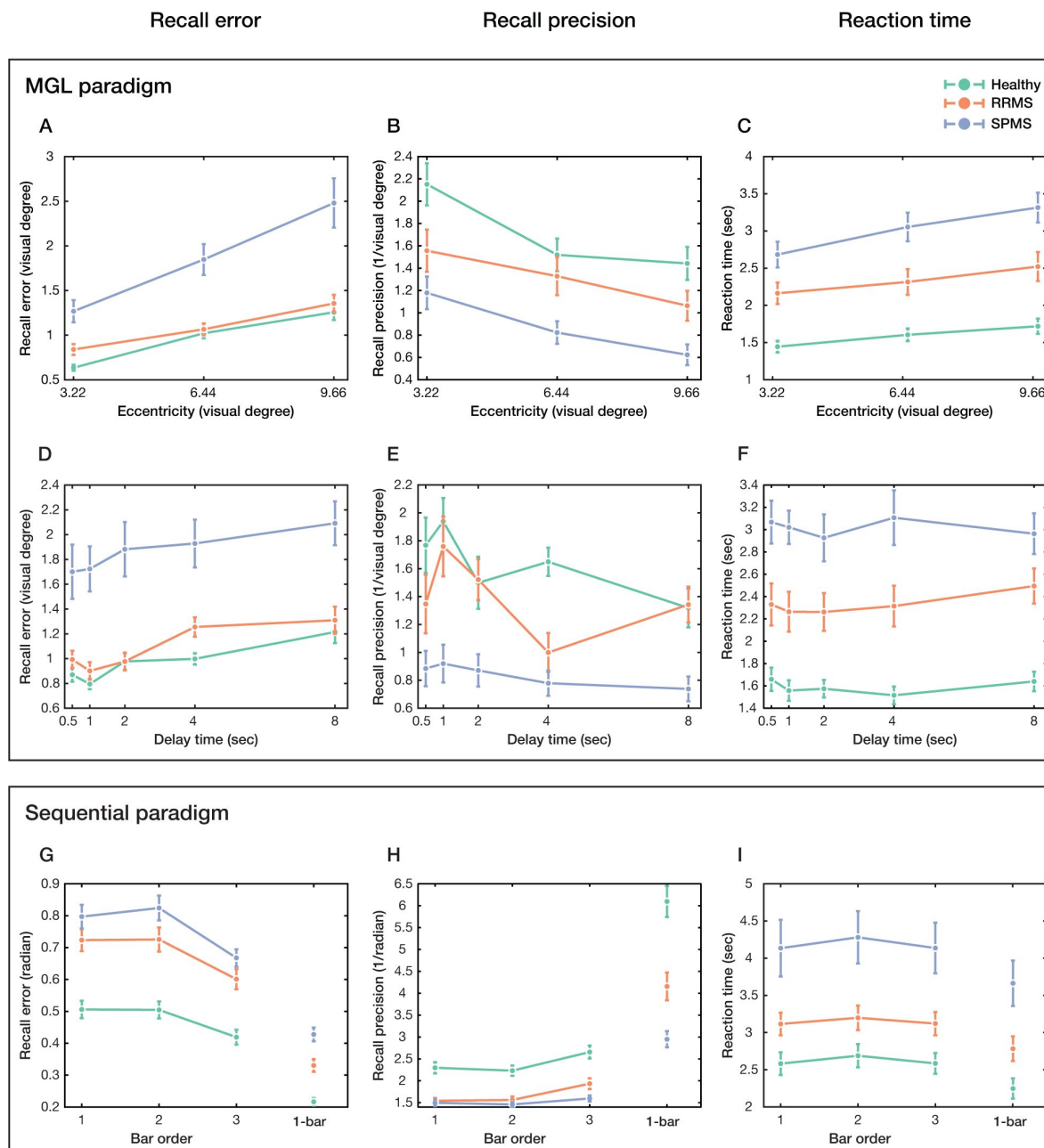


Figure 2.

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 the 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 the sequential paradigms with 3 bar (left) and 1 bar (right). Data are represented as mean $\pm$ SEM.

Recall error was evaluated for sequential tasks using the same method. In the ‘high memory load’ condition (i.e., sequential paradigm with 3-bar), 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 [$0.48 \pm 0.16, P < 10^{-5}, P < 10^{-8}$, respectively]. However, no significant difference was detected between RRMS and SPMS groups [$P = 0.14$]. After adjusting for gender, age, education, and cognitive abilities (they were significantly different between groups), the groups’ effect on recall error remained significant (Supplementary Table 3). 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 group and bar order [$F(4,114) = 0.16, P = 0.96$]. The statistical results of RT were summarized in Supplementary Table 4.

In the ‘low memory load’ condition (i.e., 1-bar 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 [$0.22 \pm 0.08, P < 10^{-4}, P < 10^{-8}$, respectively]. We also observed a significant difference between RRMS and SPMS [$p < 10^{-3}$]. After adjusting for gender, age, education, and cognitive abilities, the groups’ effect on recall error remained significant (Supplementary Table 5). 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 4.

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 group 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 /°] than in both RRMS [$1.32 \pm 0.51, P < 0.039$] and healthy control [$1.70 \pm 0.60, P < 10^{-5}$]. Recall precision was not significantly different between RRMS and healthy control [$P = 0.08$]. Moreover, this effect remained significant after adjusting for cognitive abilities (Supplementary Table 1). We also determined the effects of delay interval 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 ; 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: $P = 0.34$]. This effect remained significant after adjusting for cognitive abilities (Supplementary Table 1).

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 [$1.68 \pm 0.50, P < 10^{-6}$] and SPMS patients [$1.52 \pm 0.31, P < 10^{-7}$], no difference was observed between RRMS and SPMS participants [$P = 0.49$]. After adjusting for gender, age, education, and cognitive abilities, the effect of groups on recall precision remained significant (Supplementary Table 3).

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 [$4.16 \pm 1.98, P < 10^{-4}$] and SPMS [$2.95 \pm 1.05, P < 10^{-8}$]. Moreover, there was a significant difference between RRMS and SPMS [$P < 0.031$], which after adjusting for gender, age, education, and cognitive abilities, remained significant (Supplementary Table 5).

Distribution of Error in Recalling Information in Multiple Sclerosis

The distribution of errors in recalling information was assessed further to investigate the underlying mechanisms of WM impairment in MS. In this regard, the recorded data from the sequential paradigms was fitted to a probabilistic model developed by Bays et al. (Bays et al., 2009). In line with the results of recall error and precision, for the high memory load condition, circular standard deviation (SD) of von Mises distribution of recall error 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 [$F(4,114) = 1.19, P = 0.31$]. von Mises SD of recall error was lower in healthy control [0.51 ± 0.12] than both RRMS [$0.69 \pm 0.20, P < 10^{-5}$] and SPMS [$0.76 \pm 0.15, P < 10^{-8}$]. There was no difference between RRMS and SPMS in the von Mises SD parameter [$P = 0.16$]. In the low memory load condition, von Mises SD was affected by groups [$F(2,114) = 33.07, P < 10^{-11}$], **Fig. 3A**. Again, von Mises SD was lower in healthy control [0.25 ± 0.10] than both RRMS [$0.38 \pm 0.14, P < 10^{-5}$] and SPMS [$0.47 \pm 0.12, P < 10^{-8}$]. There was a significant difference between RRMS and SPMS in von Mises SD [$P < 0.02$]. After adjusting for gender, age, education, and cognitive abilities, the groups' effect on von Mises SD remained significant in both high and low memory load conditions (Supplementary Tables 3 and 5).

According to the study by Bays et al. (Bays et al., 2009), there are three sources of error for recalling information. The sources of these errors were defined as Gaussian variability in response value around i: target (target response proportion, **Fig. 3B**) and ii: non-target values (swap error, **Fig. 3C**) and iii: random responses (uniform response proportion, **Fig. 3D**). In the high memory load condition, target response proportion 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 group and bar order [$F(4,114) = 0.43, P = 0.78$]. Target proportion was higher in healthy control [0.88 ± 0.11] than both RRMS [$0.79 \pm 0.12, P < 0.003$] and SPMS [$0.76 \pm 0.14, P < 10^{-4}$]. There was no significant difference in target proportion between RRMS and SPMS groups [$P = 0.54$]. After adjusting for gender, age, years of education, and cognitive abilities, the effect of groups on target proportion remained significant (Supplementary Table 3). Moreover, the nearest neighbor analysis was performed to further evaluate the effect of target proportion in the absence of swap error. Removing the effect of swap error allowed us to assess the isolated effect of orientation recall. The findings from the nearest neighbor analysis showed a similar pattern of results (Supplementary Fig.1). The isolated effect of orientation was significantly different 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 order [$F(4,114) = 0.58, P = 0.67$]. After adjusting for gender, age, education, and cognitive abilities, the results of nearest neighbor analysis remained significant (Supplementary Table 3).

In the low memory load condition, groups significantly affected target response proportion [$F(2,114) = 3.11, P < 0.049$], **Fig. 3B**. While target proportion in the healthy group [0.98 ± 0.03] was significantly higher than SPMS patients [$0.94 \pm 0.09, P < 0.04$], after adjusting for gender, age, years of education, and cognitive abilities, this effect became insignificant (Supplementary Table 5). In addition, the target proportion did not significantly differ between healthy vs. RRMS [$0.96 \pm 0.08, P = 0.48$] and RRMS vs. SPMS population [$P = 0.37$].

In line with the above findings, swap error was higher in the MS population. In the high memory load condition, swap 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 group and bar order [$F(4,114) = 1.45, P = 0.22$]. Swap error was lower in healthy control [0.07 ± 0.06] than in both RRMS [$0.11 \pm 0.09, P < 0.05$] and SPMS patients [$0.14 \pm 0.09, P < 0.002$]. There was no significant difference in swap error between RRMS and SPMS [$P = 0.41$]. After adjusting for gender, age, years of education, and cognitive abilities, the groups' effect on swap error remained significant (Supplementary Table 3). Moreover, while in the high memory load condition, the uniform response proportion was different between groups [$F(2,114) = 5.50, P < 0.006$], **Fig. 3D**, no such differences were observed between bar orders [$F(2,114) = 0.81, P = 0.45$].

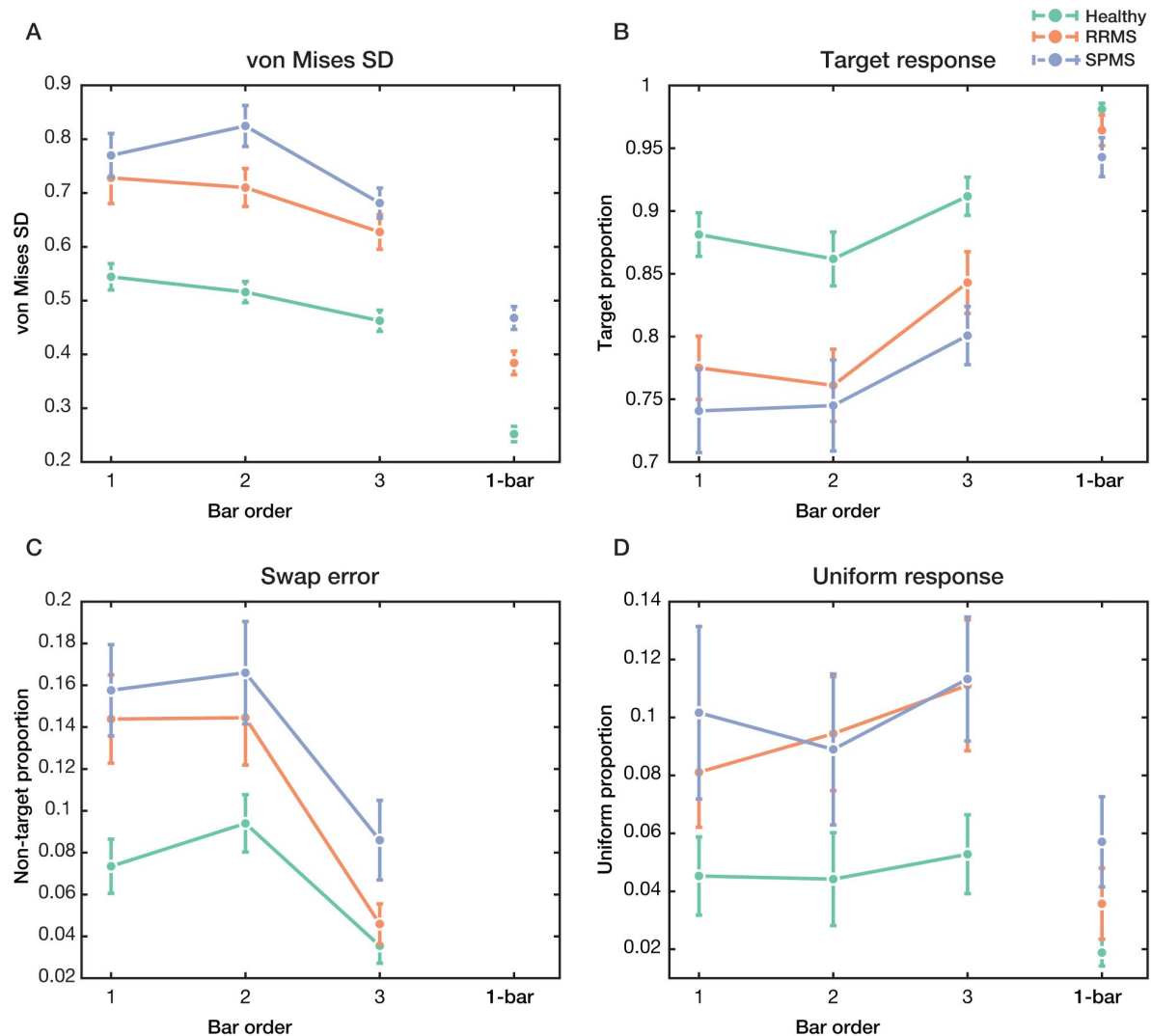


Figure 3.

The sources of recall error in high and low memory load conditions (3-bar and 1-bar, respectively).

(A) von Mises SD (circular standard deviation of von Mises distribution), (B) Target proportion (probability of response around the target value), (C) swap error (probability of response around the non-target values), and (D) uniform proportion (probability of random response) for healthy control and MS subtypes in the sequential paradigms with 3 bar (left of each subplot) and 1 bar (right of each subplot). Data are represented as mean ± SEM.

or the interaction between them [$F(4,114) = 0.18, P = 0.95$]. Post hoc analysis revealed that uniform proportion was lower in healthy control [0.05 ± 0.08] than both RRMS [$0.09 \pm 0.08, P < 0.03$] and SPMS [$0.10 \pm 0.08, P < 0.02$]. Moreover, there was no significant difference in uniform response proportion between RRMS and SPMS groups [$p = 0.95$]. Additionally, although after adjusting for gender, the effect of groups on uniform proportion remained significant, adding age, years of education, or cognitive abilities made this effect insignificant (Supplementary Table 3). The result of uniform response proportion in the low memory load condition is mathematically same as the target proportion (uniform proportion = $1 - \text{target proportion}$, since there was no swap error in the 1-bar condition).

Dissociable Function of MGL and Sequential Paradigms

The classifying ability of MGL and sequential paradigms in differentiating healthy control from MS patients was assessed based on recall error parameters using the receiver operating characteristic (ROC) analysis (Fig. 4A–C). The accuracy of MGL and sequential paradigms with 3-bar and 1-bar 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) and sequential paradigms with 3-bar (Fig. 4E) and 1-bar (Fig. 4F). While the MGL paradigm had good performance 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-bar 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-bar paradigm showed an in-between pattern of results. The 1-bar paradigm accurately discriminates healthy control from SPMS [94.4%], while it also performed well 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 analog recall paradigms with sequential presentation. Our results align with the previous reports regarding WM deficit in MS (Costers et al., 2021; PARMENTER et al., 2006; Pourmohammadi et al., 2022; Stojanovic-Radic et al., 2015; Vacchi et al., 2017). Complementary to these reports, which envisaged a binary model for storing information (slot-based model), we assessed recall precision in WM using analog reproduction paradigms (resource-based model). Further, by utilizing the unique design of our sequential paradigm, which allows us to assess the distribution of error in recalling information, we introduced a new mechanistic insight into the visual WM dysfunction in MS.

Although the results from both MGL and sequential paradigms showed an overall decrease in recall precision and increased recall error in MS population, 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 sequential paradigm with 3-bar (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-bar paradigm (low memory load condition), in which all three groups performed with different levels of precision. ROC analysis further confirmed these results, which determined a dissociation between the classifying

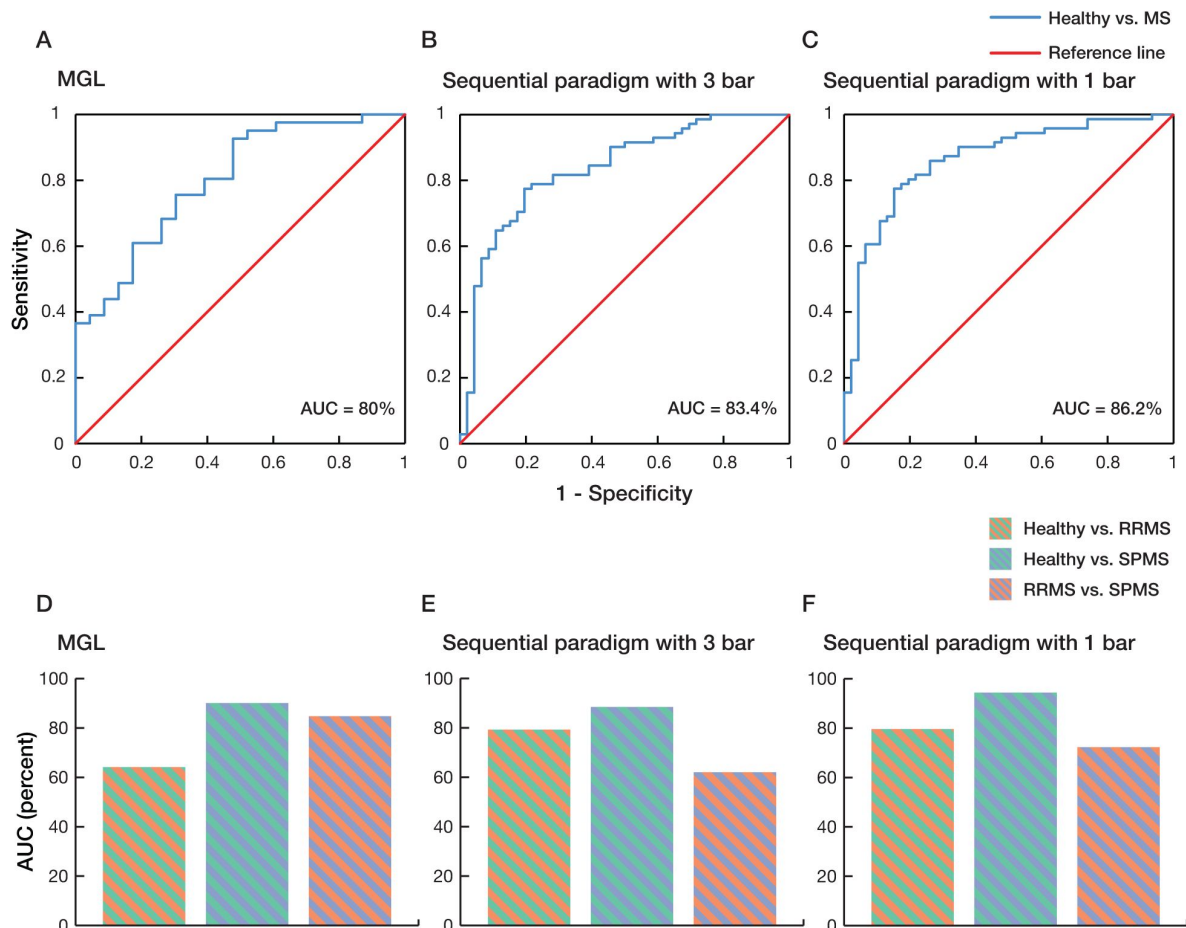


Figure 4.

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 and sequential paradigms with (B) 3 bar and (C) 1 bar 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 the curve (AUC) for (D) MGL and sequential paradigms with (E) 3 bar and (F) 1 bar.

performance of MGL and sequential paradigms. Additionally, although the low memory load condition was better than the high memory load condition in distinguishing MS subtypes, its classifying performance was not as good 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 sequential paradigms), in which spatial WM was assessed in the MGL paradigm. As the spatial WM process was associated with the function of the hippocampus (Rolls, 2018 [\[1\]](#)), the observed difference could indicate more hippocampal disruption in SPMS patients. This finding is in line with previous studies that showed more hippocampal regional loss (Sicotte et al., 2008 [\[2\]](#)) and increased hippocampus neuroinflammatory activity in SPMS compared to RRMS (Cree et al., 2021 [\[3\]](#)), suggesting that assessment of the spatial WM could be a specific marker for disorganization of the WM system in SPMS.

Another explanation is the long delay interval 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 finding is in line with our previous study, which showed that change detection paradigms with long delay intervals were promising in differentiating multiple sclerosis subtypes (Pourmohammadi et al., 2022 [\[4\]](#)). 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 the sequential paradigm). However, since the stimulus presentation time was adequate in sequential paradigms and the stimuli were presented consecutively, it did not seem that the inadequate time for encoding information was responsible for this difference (Peich et al., 2013 [\[5\]](#); Zokaei et al., 2015 [\[6\]](#)).

Additionally, one may argue that the observed dissociation could be due to the extra binding process needed in the sequential paradigm with 3-bar stimuli. However, the results from the low memory load condition and our findings from the sequential paradigm 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. Nevertheless, 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) (Parra et al., 2015 [\[7\]](#)). This issue becomes more interesting when we realize that the relational binding function is mainly centered on the hippocampus (Liang et al., 2016 [\[8\]](#); Parra et al., 2015 [\[7\]](#)), 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 (Bastin et al., 2019 [\[9\]](#); Dobson & Giovannoni, 2019 [\[10\]](#); Figueroa-Vargas et al., 2020 [\[11\]](#); Lugtmeijer et al., 2021 [\[12\]](#); Vacchi et al., 2017 [\[13\]](#)).

We applied a probabilistic model to disentangle the error distribution in recalling information in the sequential paradigms to further investigate the underlying mechanism behind the visual WM dysfunction in MS. Our finding revealed that in addition to imprecision in decoding information, swap error (mistakenly reporting a non-target feature) contributed to WM dysfunction in individuals with MS. Swap error has been associated with various mechanisms, including the variability in cue feature dimension, cue-independent sources, and strategic guessing (McMaster et al., 2022 [\[14\]](#)). The recent study by McMaster et al. comprehensively investigated these hypotheses and determined that the variability in cue feature dimension could solely account for the swap error mechanism (McMaster et al., 2022 [\[14\]](#)). The neural correlates of swap error could be explained by the imprecision in decoding information in the cue feature dimension (i.e., color in

this study) due to the noise in neural activity. Hence, a non-cued item may mistakenly be considered as the actual cue, leading to the reporting of the corresponding non-target feature as a response (Schneegans & Bays, 2017 [DOI](#)). Therefore, the swap error observed in this study can be attributed to the variability in the color dimension. The neural underpinning of this observation in MS population could be related to their impaired neural activity. Recent studies have demonstrated disturbed neural activity during WM tasks in MS (Costers et al., 2021 [DOI](#); Figueroa-Vargas et al., 2020 [DOI](#)). For instance, Costers et al. showed a disturbed theta band oscillation, which plays a crucial role in WM encoding information, in the hippocampal area of MS population (Costers et al., 2021 [DOI](#)). While this evidence provides a primary insight into the possible underlying mechanisms of WM dysfunction in MS, future electrophysiological studies using the cued recall paradigms could further elucidate this matter.

Swap errors (misbinding errors) have been observed in various neurological disorders, including different types of Alzheimer's disease (Cecchini et al., 2022 [DOI](#); Della Sala et al., 2012 [DOI](#); Liang et al., 2016 [DOI](#); Zokaei et al., 2020 [DOI](#)), epileptic patients with temporal lobe lobectomy (Zokaei et al., 2019 [DOI](#)), and voltage-gated potassium channel complex antibody (VGKC-Ab) limbic encephalitis (Pertzov et al., 2013 [DOI](#)). From a neuroanatomical perspective, previous studies proposed that impairment in hippocampus and medial temporal lobe regions contribute to problems in relational binding (Della Sala et al., 2012 [DOI](#); Liang et al., 2016 [DOI](#); Pertzov et al., 2013 [DOI](#); Zokaei et al., 2019 [DOI](#)). Furthermore, whereas regions related to conjunctive binding have yet to be better understood, occipital and parietal regions seem to be mainly related to conjunctive processing (Cecchini et al., 2022 [DOI](#)). Additionally, a recent study by Valdés Hernández and colleagues also demonstrated the possible role of globus pallidus in conjunctive binding (Valdés Hernández et al., 2020 [DOI](#)). On the other hand, a study by Vacchi et al. demonstrated less activation of the superior and inferior parietal lobule during an N-back fMRI task in individuals with MS (Vacchi et al., 2017 [DOI](#)). Moreover, globus pallidus was shown to be involved in MS (Fujiwara et al., 2017 [DOI](#)). Hence, considering the convergence of brain regions involved in conjunctive processing (assessed through the current sequential paradigm) with those shown to be affected in MS, the parietal regions and globus pallidus could be the candidate regions responsible for the observed impairment. Additionally, based on the evidence showing the involvement of hippocampal regions in MS (Rocca et al., 2018 [DOI](#); Sicotte et al., 2008 [DOI](#)), we also expect to see relational processing impairment; however, the current study design did not allow us to evaluate this condition. Further neuroimaging studies are needed to validate these findings. 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 these findings, our study had some limitations that should be addressed in future research. 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 deficit in MS. Concurrent 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 assessing the source of WM deficit regarding relational binding could elucidate this statement.

Additionally, although clinical and cognitive assessments were performed to mitigate the possible confounding effects of cognitive, visual, and motor impairments on the outcomes of the study, it cannot be concluded that no confounding effects occurred. In future studies, including a control condition matched to the experimental paradigm where only the memory components are removed could better clarify this issue. Finally, considering the aim of this study, which was to develop a practical apparatus for WM assessment in clinical settings, we did not use an eye

tracker. Although we attempted to minimize the effect of eye movement in both paradigms by instructing the participants to fixate on a central fixation point, using an eye tracker is necessary to validate these findings further.

In summary, using the resource-based model paradigms, we demonstrated that recall error and precision were impaired in MS. We provided new insight regarding the progressive state of the disease by assessing the plausible mechanisms related to the dissociable behavior of MGL and sequential paradigms in classifying MS subtypes. Furthermore, applying a computational model capable of disentangling error distribution in recalling information allowed us to uncover new insight regarding WM deficit in the MS population. Our results determined that decreased signal-to-noise ratio and swap error were responsible for WM deficit in MS. Overall, 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) and 73 healthy control volunteers participated in this study. Patients were recruited in a full-census manner from the Comprehensive Multiple Sclerosis Clinic at Kashani Hospital in Isfahan between February 2021 and January 2022. The following criteria were used for participant selection: diagnosis based on the 2017 McDonald criteria (Thompson et al., 2018 [DOI](#)), age between 18 to 55 years, diagnosis within 1-18 years prior to entering the study, an expanded disability status scale (EDSS) score of 0-6.5, no history of clinical relapse or corticosteroid therapy within two months before entering the study, normal or corrected-to-normal visual acuity and color vision (based on the recorded bedside neurologic examination in their profile), and no impairment in factors that could interfere with the study including visual acuity, visual field, extraocular movement, scotoma, nystagmus, or tremor in the upper extremity, assessed as part of the EDSS examination. Additionally, normal performance in the Nine-Hole Peg Test (9-HPT < 45 seconds), the absence of a history of brain surgeries, major neurological disorders (stroke, epilepsy, brain tumor, or CNS infection), psychiatric disorders (major depressive disorder, bipolar disorder, or schizophrenia), drug or alcohol abuse, and chronic systemic disorders (diabetes, renal failure, liver failure, chronic obstructive pulmonary disorder (COPD), hypothyroidism, or hyperthyroidism) were required for participation in this study. Moreover, the control group should not have a family history of MS in their first-degree relatives.

2. Procedure

The study began with a clinical interview and neurologic examination to collect clinical and demographic information from the participants. Participants also completed 9-HPT and the Persian version of the MoCA test (Badrkhahan et al., 2020 [DOI](#)). MoCA is a standard cognitive screening tool with a scoring system ranging from 0 to 30, where distinct ranges correspond to different levels of cognitive function. Scores of ≥ 26 indicate normal cognitive performance, 18-25 denotes mild cognitive impairment, 10-17 determine moderate cognitive impairment, and ≤ 10 is considered severe impairment.

Participants received verbal and written instructions before beginning the assessment. A 10-trial training block was then performed to ensure their understanding of the tasks. After this, the actual assessment started.

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 ("World Medical Association Declaration of Helsinki," 2013 [DOI](#)) and was approved by the Iranian National Committee of Ethics in Biomedical

3. Visual working memory paradigms

Visual WM was assessed using two analog recall tasks, memory-guided localization (MGL) and analog recall paradigms with sequential bar presentation. 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 (Brainard, 1997; Pelli, 1997).

3.1. 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 (absolute error), Euclidian distance between the target's location and subject response in visual degree, and reaction time (RT) were recorded for further assessment.

3.2. Sequential paradigm with bar stimuli

Two designs of analog report tasks with sequential bar presentation, i.e., the low memory load condition (1-bar) and high memory load condition (3-bar), 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' cued 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 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). The 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 (Fig. 1C). After the high memory load condition, subjects participated in 30 trials of the low memory load condition. Due to the 1-bar design of low memory load condition, the swap error was absent, so fewer trials were needed (Peich et al., 2013). 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 (absolute error), angular difference between the target value and subject response, and RT were recorded.

4. 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 $\pm$ standard deviation (SD). Data with extreme outliers (values greater than 3rd quartile + 3 $\times$ interquartile range or less than 1st quartile – 3 $\times$ interquartile range) in MGL and sequential paradigm with 3-bar stimuli were excluded from further analysis. The level of significance was set at p -value < 0.05 .

Clinical and demographic profiles of the participants, except for the gender, cognitive ability, and treatment history, were compared using one-way ANOVA or Kruskal-Wallis H test (three groups comparison) and independent sample t-test or Mann-Whitney U test (two groups comparison). The cognitive ability of participants was classified according to the mentioned ranges of MoCA score. Patients' disease-modifying therapy (DMT) was classified into platform and non-platform treatments. In our study, platform treatments include interferon beta-1a and glatiramer acetate, and non-platform treatments contain rituximab, ocrelizumab, fingolimod, dimethyl fumarate, and natalizumab. These categorical variables, along with gender, were compared using the Chi-squared test. The post hoc Tukey's and Dunn's multiple comparison tests were performed following the significant results of ANOVA and Kruskal-Wallis H test. Also, Bonferroni correction for multiple tests was performed following Dunn's post hoc analysis, and the adjusted p -value was reported.

For the MGL paradigm, recall error, recall precision (defined as the reciprocal of the standard deviation of recall error), and RT were compared between groups (Healthy, RRMS, and SPMS) among different conditions (distance or delay, mixed model ANOVA, between- and within-subjects comparisons). For the sequential paradigm, since the data was in a circular space, based on the method proposed by Fisher (Fisher, 1993 [\[1\]](#)), we used the analog report MATLAB toolbox of Bays Lab (<https://www.paulbays.com/toolbox/> [\[2\]](#)) to calculate the circular mean of recall error for each subject. Similarly, recall precision, defined as the reciprocal of the circular standard deviation of recall error, was calculated for each subject. For both high and low memory load conditions (3-bar and 1-bar), recall error, recall precision, and RT were compared between groups with respect to the order of presented bars (mixed model and one-way ANOVA). To further investigate the sources of error and uncover the involved mechanisms in visual WM impairment, the Mixture Model (<https://www.paulbays.com/toolbox/> [\[2\]](#)), a probabilistic model developed before (Bays et al., 2009 [\[3\]](#); Schneegans & Bays, 2016 [\[4\]](#)), was utilized. The Mixture Model considers three possible sources for information recall. They are defined as the Gaussian variability in reporting the target and non-target values and random responses (Bays et al., 2009 [\[3\]](#); Schneegans & Bays, 2016 [\[4\]](#)). In our study, they referred to reporting the orientation of the target bar (target proportion), misreporting the orientation of the other two non-target bars instead of the target bar (swap error), and random response (uniform proportion). Using the Mixture Model, the probabilities of target, non-target, and uniform responses and the von Mises distribution concentration score were calculated for each subject. The sources of error were evaluated by comparing the probability of target, non-target, and uniform responses between groups among different bar orders (mixed model ANOVA). Also, von Mises SD, defined as the circular standard deviation of the concentration score of von Mises distribution, was assessed using the same method. Moreover, nearest neighbor analysis was performed to further evaluate the isolated effect of orientation (Pertzov et al., 2013 [\[5\]](#)). In this set of analyses, we removed the effect of swap error 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-bar design of the low memory load condition, swap error was not present; hence, only the target proportion, uniform proportion, and von Mises SD were compared.

For each comparison, hierarchical regression analyses were performed to evaluate the possible confounding effect of demographic and clinical variables (significantly different between groups) on results. Finally, the dissociable function of MGL and sequential paradigms in distinguishing

healthy control from MS patients, healthy control from MS subtypes, and MS subtypes from each other were assessed by performing ROC analysis. The area under the curve (AUC) for recall error is calculated, which serves as a measure of classifying accuracy for the paradigms.

Acknowledgements

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.

Additional information

Funding

This study was supported by Isfahan University of Medical Sciences (grant number: 2400104). The funder had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Ali Motahharynia, Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing—original draft, Writing—review & editing, Visualization; Ahmad Pourmohammadi, Conceptualization, Methodology, Software, Investigation, Data curation, Writing—review & editing; Armin Adibi, Investigation, Writing—review & editing; Vahid Shaygannejad, Resources, Writing—review & editing, Funding acquisition; Fereshteh Ashtari, Resources, Writing—review & editing, Funding acquisition; Iman Adibi, Conceptualization, Methodology, Formal analysis, Resources, Writing—review & editing, Supervision, Funding acquisition; Mehdi Sanayei, Conceptualization, Methodology, Software, Formal analysis, Writing—review & editing, Supervision

Competing interests

The authors declare that no competing interests exist.

Ethics

This study followed the latest update of the Declaration of Helsinki and was approved by the Iranian National Committee of Ethics in Biomedical Research (Approval ID: IR.MUI.MED.REC.1400.441).

Additional files

Supplementary files

Supplementary file is available online.

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/> .

References

- Badrkhahan S. Z., Sikaroodi H., Sharifi F., Kouti L., Noroozian M (2020) **Validity and reliability of the Persian version of the Montreal Cognitive Assessment (MoCA-P) scale among subjects with Parkinson's disease** *Applied Neuropsychology: Adult* **27**:431–439 <https://doi.org/10.1080/23279095.2019.1565762>
- Bastin C., Besson G., Simon J., Delhay E., Geurten M., Willems S., Salmon E (2019) **An integrative memory model of recollection and familiarity to understand memory deficits** *Behavioral and Brain Sciences* **42** <https://doi.org/10.1017/S0140525X19000621>
- Bays P. M., Catalao R. F. G., Husain M (2009) **The precision of visual working memory is set by allocation of a shared resource** *Journal of Vision* **9**:7–7 <https://doi.org/10.1167/9.10.7>
- Bays P. M., Husain M (2008) **Dynamic Shifts of Limited Working Memory Resources in Human Vision** *Science* **321**:851–854 <https://doi.org/10.1126/science.1158023>
- Benedict R. H. B., Amato M. P., DeLuca J., Geurts J. J. G (2020) **Cognitive impairment in multiple sclerosis: clinical management, MRI, and therapeutic avenues** *The Lancet Neurology* **19**:860–871 [https://doi.org/10.1016/S1474-4422\(20\)30277-5](https://doi.org/10.1016/S1474-4422(20)30277-5)
- Brainard D. H (1997) **The Psychophysics Toolbox** *Spatial Vision* **10**:433–436
- Cecchini M. A., Parra M. A., Brazzelli M., Logie R. H., Della Sala S. (2022) **Short-term memory conjunctive binding in Alzheimer's disease: A systematic review and meta-analysis** *Neuropsychology* <https://doi.org/10.1037/neu0000825>
- Clemens L., Langdon D (2018) **How does cognition relate to employment in multiple sclerosis? A systematic review** *Multiple Sclerosis and Related Disorders* **26**:183–191 <https://doi.org/10.1016/j.msard.2018.09.018>
- Costers L., Van Schependom J., Laton J., Baijot J., Sjøgård M., Wens V., De Tiège X., Goldman S., D'Haeseleer M., D'hooghe M. B., Woolrich M., Nagels G. (2021) **The role of hippocampal theta oscillations in working memory impairment in multiple sclerosis** *Human Brain Mapping* **42**:1376–1390 <https://doi.org/10.1002/hbm.25299>
- Covey T. J., Zivadinov R., Shucard J. L., Shucard D. W (2011) **Information processing speed, neural efficiency, and working memory performance in multiple sclerosis: Differential relationships with structural magnetic resonance imaging** *Journal of Clinical and Experimental Neuropsychology* **33**:1129–1145 <https://doi.org/10.1080/13803395.2011.614597>
- Cowan N (2001) **The magical number 4 in short-term memory: A reconsideration of mental storage capacity** *Behavioral and Brain Sciences* **24**:87–114 <https://doi.org/10.1017/S0140525X01003922>
- Cree B. A. C., Arnold D. L., Chataway J., Chitnis T., Fox R. J., Pozo Ramajo A., Murphy N., Lassmann H (2021) **Secondary Progressive Multiple Sclerosis** *Neurology* **97**:378–388 <https://doi.org/10.1212/WNL.00000000000012323>

- Della Sala S., Parra M. A., Fabi K., Luzzi S., Abrahams S. (2012) **Short-term memory binding is impaired in AD but not in non-AD dementias** *Neuropsychologia* **50**:833–840 <https://doi.org/10.1016/j.neuropsychologia.2012.01.018>
- Dobson R., Giovannoni G (2019) **Multiple sclerosis – a review** *European Journal of Neurology* **26**:27–40 <https://doi.org/10.1111/ene.13819>
- Figueroa-Vargas A., Cárcamo C., Henríquez-Ch R., Zamorano F., Ciampi E., Uribe-San-Martin R., Vásquez M., Aboitiz F., Billeke P (2020) **Frontoparietal connectivity correlates with working memory performance in multiple sclerosis** *Scientific Reports* **10** <https://doi.org/10.1038/s41598-020-66279-0>
- Fisher N. I. (1993) **Statistical Analysis of Circular Data** <https://doi.org/10.1017/CBO9780511564345>
- Fujiwara E., Kmech J. A., Cobzas D., Sun H., Seres P., Blevins G., Wilman A. H (2017) **Cognitive Implications of Deep Gray Matter Iron in Multiple Sclerosis** *American Journal of Neuroradiology* **38**:942–948 <https://doi.org/10.3174/ajnr.A5109>
- Gorgoraptis N., Catalao R. F. G., Bays P. M., Husain M (2011) **Dynamic Updating of Working Memory Resources for Visual Objects** *Journal of Neuroscience* **31**:8502–8511 <https://doi.org/10.1523/JNEUROSCI.0208-11.2011>
- Hulst H. E., Goldschmidt T., Nitsche M. A., de Wit S. J., van den Heuvel O. A., Barkhof F., Paulus W., van der Werf Y. D., Geurts J. J. G. (2017) **rTMS affects working memory performance, brain activation and functional connectivity in patients with multiple sclerosis.** *Journal of Neurology Neurosurgery & Psychiatry* **88**:386–394 <https://doi.org/10.1136/jnnp-2016-314224>
- Liang Y., Pertzov Y., Nicholas J. M., Henley S. M. D., Crutch S., Woodward F., Leung K., Fox N. C., Husain M (2016) **Visual short-term memory binding deficit in familial Alzheimer’s disease** *Cortex* **78**:150–164 <https://doi.org/10.1016/j.cortex.2016.01.015>
- Lugtmeijer S., Schneegans S., Lammers N. A., Geerligs L., de Leeuw F. E., de Haan E. H. F., Bays P. M., Kessels R. P. C. (2021) **Consequence of stroke for feature recall and binding in visual working memory** *Neurobiology of Learning and Memory* **179** <https://doi.org/10.1016/j.nlm.2021.107387>
- Ma W. J., Husain M., Bays P. M (2014) **Changing concepts of working memory** *Nature Neuroscience* **17**:347–356 <https://doi.org/10.1038/nn.3655>
- McMaster J. M. V., Tomić I., Schneegans S., Bays P. M (2022) **Swap errors in visual working memory are fully explained by cue-feature variability** *Cognitive Psychology* **137** <https://doi.org/10.1016/j.cogpsych.2022.101493>
- Miller E. K., Lundqvist M., Bastos A. M (2018) **Working Memory 2.0** *Neuron* **100**:463–475 <https://doi.org/10.1016/j.neuron.2018.09.023>
- Miller G. A (1956) **The magical number seven plus or minus two: some limits on our capacity for processing information** *Psychological Review* **63**:81–97
- Nee D. E., D’Esposito M (2016) **Nee, D. E., & D’Esposito, M. (2016). The Representational Basis of Working Memory (pp. 213–230). 10.1007/7854_2016_456 The Representational Basis of Working Memory :213–230** https://doi.org/10.1007/7854_2016_456

- Parmenter B. A., Shucard J. L., Benedict R. H. B., Shucard D. W (2006) **Working memory deficits in multiple sclerosis: Comparison between the n-back task and the Paced Auditory Serial Addition Test** *Journal of the International Neuropsychological Society* **12**:677–687 <https://doi.org/10.1017/S1355617706060826>
- Parra M. A., Fabi K., Luzzi S., Cubelli R., Hernandez Valdez M., Della Sala S. (2015) **Relational and conjunctive binding functions dissociate in short-term memory** *Neurocase* **21**:56–66 <https://doi.org/10.1080/13554794.2013.860177>
- Peich M.-C., Husain M., Bays P. M (2013) **Age-related decline of precision and binding in visual working memory** *Psychology and Aging* **28**:729–743 <https://doi.org/10.1037/a0033236>
- Pelli D. G (1997) **The VideoToolbox software for visual psychophysics: transforming numbers into movies** *Spatial Vision* **10**:437–442
- Pertsov Y., Miller T. D., Gorgoraptis N., Caine D., Schott J. M., Butler C., Husain M (2013) **Binding deficits in memory following medial temporal lobe damage in patients with voltage-gated potassium channel complex antibody-associated limbic encephalitis** *Brain* **136**:2474–2485 <https://doi.org/10.1093/brain/awt129>
- Pourmohammadi A., Motahharynia A., Shaygannejad V., Ashtari F., Adibi I., Sanayei M (2022) **Working memory dysfunction differs between secondary progressive and relapsing multiple sclerosis: effects of clinical phenotype, age, disease duration, and disability** *Multiple Sclerosis and Related Disorders* **104411** <https://doi.org/10.1016/j.msard.2022.104411>
- Rocca M. A., Barkhof F., De Luca J., Frisén J., Geurts J. J. G., Hulst H. E., Sastre-Garriga J., Filippi M., Barkhof F., Ciccarelli O., De Stefano N., Enzinger C., Filippi M., Frederiksen J. L., Gasperini C., Kappos L., Palace J., Rocca M. A., Rovira A., Yousry T. A. (2018) **The hippocampus in multiple sclerosis** *The Lancet Neurology* **17**:918–926 [https://doi.org/10.1016/S1474-4422\(18\)30309-0](https://doi.org/10.1016/S1474-4422(18)30309-0)
- Rocca M. A., Valsasina P., Hulst H. E., Abdel-Aziz K., Enzinger C., Gallo A., Pareto D., Riccitelli G., Muhlert N., Ciccarelli O., Barkhof F., Fazekas F., Tedeschi G., Arévalo M. J., Filippi M (2014) **Functional correlates of cognitive dysfunction in multiple sclerosis: A multicenter fMRI Study** *Human Brain Mapping* **35**:5799–5814 <https://doi.org/10.1002/hbm.22586>
- Rolls E. T (2018) **The storage and recall of memories in the hippocampo-cortical system** *Cell and Tissue Research* **373**:577–604 <https://doi.org/10.1007/s00441-017-2744-3>
- Schneegans S., Bays P. M (2016) **No fixed item limit in visuospatial working memory** *Cortex* **83**:181–193 <https://doi.org/10.1016/j.cortex.2016.07.021>
- Schneegans S., Bays P. M (2017) **Neural Architecture for Feature Binding in Visual Working Memory** *The Journal of Neuroscience* **37**:3913–3925 <https://doi.org/10.1523/JNEUROSCI.3493-16.2017>
- Sicotte N. L., Kern K. C., Giesser B. S., Arshanapalli A., Schultz A., Montag M., Wang H., Bookheimer S. Y (2008) **Regional hippocampal atrophy in multiple sclerosis** *Brain* **131**:1134–1141 <https://doi.org/10.1093/brain/awn030>
- Stojanovic-Radic J., Wylie G., Voelbel G., Chiaravalloti N., DeLuca J (2015) **Neuroimaging and cognition using functional near infrared spectroscopy (fNIRS) in multiple sclerosis** *Brain Imaging and Behavior* **9**:302–311 <https://doi.org/10.1007/s11682-014-9307-y>

Thompson A. J., Banwell B. L., Barkhof F., Carroll W. M., Coetzee T., Comi G., Correale J., Fazekas F., Filippi M., Freedman M. S., Fujihara K., Galetta S. L., Hartung H. P., Kappos L., Lublin F. D., Marrie R. A., Miller A. E., Miller D. H., Montalban X., Cohen J. A (2018) **Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria** *The Lancet Neurology* **17**:162–173 [https://doi.org/10.1016/S1474-4422\(17\)30470-2](https://doi.org/10.1016/S1474-4422(17)30470-2)

Vacchi L., Rocca M. A., Meani A., Rodegher M., Martinelli V., Comi G., Falini A., Filippi M (2017) **Working memory network dysfunction in relapse-onset multiple sclerosis phenotypes: A clinical-imaging evaluation** *Multiple Sclerosis Journal* **23**:577–587 <https://doi.org/10.1177/1352458516656809>

Valdés Hernández M. C., Clark R., Wang S.-H., Guazzo F., Calia C., Pattan V., Starr J., Della Sala S., Parra M. A. (2020) **The striatum, the hippocampus, and short-term memory binding: Volumetric analysis of the subcortical grey matter's role in mild cognitive impairment** *NeuroImage: Clinical* **25** <https://doi.org/10.1016/j.nicl.2019.102158>

World Medical Association Declaration of Helsinki (2013) **World Medical Association Declaration of Helsinki. (2013).** *JAMA*, **310**(), 2191. [10.1001/jama.2013.281053](https://doi.org/10.1001/jama.2013.281053) *JAMA* **310** <https://doi.org/10.1001/jama.2013.281053>

Zokaei N., Burnett Heyes S., Gorgoraptis N., Budhdeo S., Husain M (2015) **Working memory recall precision is a more sensitive index than span** *Journal of Neuropsychology* **9**:319–329 <https://doi.org/10.1111/jnp.12052>

Zokaei N., Nour M. M., Sillence A., Drew D., Adcock J., Stacey R., Voets N., Sen A., Husain M (2019) **Binding deficits in visual short-term memory in patients with temporal lobe lobectomy** *Hippocampus* **29**:63–67 <https://doi.org/10.1002/hipo.22998>

Zokaei N., Sillence A., Kienast A., Drew D., Plant O., Slavkova E., Manohar S. G., Husain M (2020) **Different patterns of short-term memory deficit in Alzheimer's disease, Parkinson's disease and subjective cognitive impairment** *Cortex* **132**:41–50 <https://doi.org/10.1016/j.cortex.2020.06.016>

Article and author information

Ali Motahharynia

Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, Iran, Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, Iran
ORCID iD: [0000-0002-1140-3257](https://orcid.org/0000-0002-1140-3257)

Ahmad Pourmohammadi

Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, Iran, Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, Iran, School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
ORCID iD: [0000-0002-4996-0569](https://orcid.org/0000-0002-4996-0569)

Armin Adibi

Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, Iran, Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, Iran
ORCID iD: [0000-0001-5428-9143](https://orcid.org/0000-0001-5428-9143)

Vahid Shaygannejad

Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, Iran,
Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, Iran,
Department of Neurology, School of Medicine, Isfahan University of Medical Sciences, Isfahan,
Iran

Fereshteh Ashtari

Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, Iran,
Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, Iran,
Department of Neurology, School of Medicine, Isfahan University of Medical Sciences, Isfahan,
Iran

Iman Adibi

Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, Iran,
Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, Iran,
Department of Neurology, School of Medicine, Isfahan University of Medical Sciences, Isfahan,
Iran

For correspondence: i.adibi@gmail.com

ORCID iD: [0000-0003-4409-9404](https://orcid.org/0000-0003-4409-9404)

Mehdi Sanayei

Center for Translational Neuroscience, Isfahan University of Medical Sciences, Isfahan, Iran,
Isfahan Neuroscience Research Center, Isfahan University of Medical Sciences, Isfahan, Iran,
School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran,
Iran

For correspondence: mehdi.sanayei@gmail.com

Copyright

© 2023, Motahharynia et al.

This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use and redistribution provided that the original author and source are credited.

Editors

Reviewing Editor

Jonathan Roiser

University College London, United Kingdom

Senior Editor

Jonathan Roiser

University College London, United Kingdom

Reviewer #1 (Public Review):

This study compares visuospatial working (VWM) memory performance between patients with MS and healthy controls, assessed using analog report tasks that provide continuous measures of recall error. The aim is to advance on previous studies of VWM in MS that have used binary (correct/incorrect) measures of recall, such as from change detection tasks, that

are not sensitive to the resolution with which features can be recalled, and to use mixture modelling to disentangle different contributions to overall performance. The results identify a specific decrease in the precision of VWM recall in MS, although the possibility that visual and/or motor impairments contribute to performance decrements on the memory task cannot be ruled out.

Reviewer #2 (Public Review):

The authors applied two visual working memory tasks, a memory-guided localization (MGL), examining short-term memory of the location of an item over a brief interval, and an N-back task, examining orientation of a centrally presented item, in order to test working memory performance in patients with multiple sclerosis (including a subgroup with relapsing-remitting and one with secondary progressive MS), compared with healthy control subjects. The authors used an approach in testing and statistically modelling visual working memory paradigm previously developed by Paul Bays, Masud Husain and colleagues. Such continuous measure approaches make it possible to quantify the precision, or resolution, of working memory, as opposed to measuring working memory using discretised, all-or-none measures. This represents an advance beyond prior work in this area.

The authors of the present study found that both MS subgroups performed worse than controls on the N-back task and that only the secondary progressive MS subgroup was significantly impaired on the MGL task. The underlying sources of error including incorrect association of an object's identity with its location or serial order, were also examined.

The application of more precise psychophysiological methods to test visual working memory in multiple sclerosis should be applauded. It has the potential to lead to more sensitive and specific tests which could potentially be used as useful outcome measures in clinical trials of disease-modifying drugs, for example.

The present study does not compare the continuous-report testing with a discrete measure task so it is unclear whether the former is more sensitive, or more feasible in this patient group, although this may not have been the purpose of the study.

Comments on the revised submission: My previous comments have been answered to the extent that is possible with the data available.

Author Response

The following is the authors' response to the current reviews.

eLife assessment

This paper provides valuable information regarding visuospatial working memory performance in patients with MS compared to healthy controls, using a relatively novel continuous measure of visual working memory. There are some weaknesses with the way the clinical groups were matched, but those limitations are acknowledged and the strength of evidence for the claims is otherwise convincing. The paper will be of interest to those working in the field of clinical neuroscience.

We are grateful to the editors and reviewers for their careful review of our manuscript and their dedicated time and effort. Their valuable feedback has been instrumental in improving the quality of our work.

Reviewer #1 (Public Review):

This study compares visuospatial working (VWM) memory performance between patients with MS and healthy controls, assessed using analog report tasks that provide continuous measures of recall error. The aim is to advance on previous studies of VWM in MS that have used binary (correct/incorrect) measures of recall, such as from change detection tasks, that are not sensitive to the resolution with which features can be recalled, and to use mixture modelling to disentangle different contributions to overall performance. The results identify a specific decrease in the precision of VWM recall in MS, although the possibility that visual and/or motor impairments contribute to performance decrements on the memory task cannot be ruled out.

Although we try to address this matter by clinical screening, as the reviewer mentioned, the possibility that visual and/or motor impairments contribute to performance decrements on the memory task cannot be ruled out. Therefore, in future studies, including a control condition matched to the experimental paradigm where only the memory components are removed is needed to elucidate this issue.

Reviewer #2 (Public Review):

The authors applied two visual working memory tasks, a memory-guided localization (MGL), examining short-term memory of the location of an item over a brief interval, and an N-back task, examining orientation of a centrally presented item, in order to test working memory performance in patients with multiple sclerosis (including a subgroup with relapsing-remitting and one with secondary progressive MS), compared with healthy control subjects. The authors used an approach in testing and statistically modelling visual working memory paradigm previously developed by Paul Bays, Masud Husain and colleagues. Such continuous measure approaches make it possible to quantify the precision, or resolution, of working memory, as opposed to measuring working memory using discretised, all-or-none measures. This represents an advance beyond prior work in this area.

The authors of the present study found that both MS subgroups performed worse than controls on the N-back task and that only the secondary progressive MS subgroup was significantly impaired on the MGL task. The underlying sources of error including incorrect association of an object's identity with its location or serial order, were also examined. The application of more precise psychophysiological methods to test visual working memory in multiple sclerosis should be applauded. It has the potential to lead to more sensitive and specific tests which could potentially be used as useful outcome measures in clinical trials of disease-modifying drugs, for example. The present study does not compare the continuous-report testing with a discrete measure task so it is unclear whether the former is more sensitive, or more feasible in this patient group, although this may not have been the purpose of the study.

The reviewer brought up an important point, but as they stated, it was not the focus of our current study. Nevertheless, it is a valuable suggestion for future research to compare continuous with discrete measure paradigms to assess their sensitivity and feasibility in the MS population.

The following is the authors' response to the original reviews.

We thank the editors and reviewers for their thorough reading of this manuscript and valuable suggestions. We appreciate the time and effort they have put into this manuscript to provide feedback for improving our work. Based on their comments, we carefully considered

their suggestions and revised the manuscript to address their concerns. Our one-by-one response to reviewer comments is as follows.

Reviewer #1 (Public Review):

This study compares visuospatial working memory performance between patients with MS and healthy controls, assessed using analog report tasks that provide continuous measures of recall error. The aim is to advance on previous studies of VWM in MS that have used binary (correct/incorrect) measures of recall, such as from change detection tasks, that are not sensitive to the resolution with which features can be recalled, and to use mixture modelling to potentially disentangle different contributions to overall performance. This aim is met in part, but there are some problems with the authors' interpretation of their findings:

- 1. How can the authors be confident the performance deficits in the patient groups are impairments of working memory and not visual or motor in nature? I appreciate there was some kind of clinical screening, but it seems like there should have been a control condition matched to the experimental tasks with only the memory components removed.*

We appreciate the reviewer's concern regarding the potential confounding effects of visual or motor impairment on the outcomes of our study.

While we acknowledge that a control condition with only the memory components removed could have further strengthened our results, we did not include one, which is a limitation of the current study.

To address this limitation, we conducted clinical screening to ensure that the observed deficit was due to working memory impairment and not visual or motor in nature. As part of the expanded disability status scale (EDSS) evaluation, we did not include individuals with issues such as visual acuity, visual field, and extraocular movement impairment, scotoma, nystagmus, and tremors in the upper extremity, which could interfere with the study. Moreover, participants were screened using the 9-Hole Peg Test (9-HPT) before entering the study. These evaluations helped us to ensure that participants with impaired visual or motor performance, which could potentially confound the study, were not included. Our effort to remove the confounding factors with clinical screening provided additional insight into the interpretability of the results. We have updated our inclusion/exclusion criteria accordingly and included this limitation in our discussion.

- 1. The participant groups are large, which is definitely a strength, but not particularly well-matched in terms of demographics, with notable differences in age (mean and spread), years of education and gender. These could potentially contribute to differences in performance between groups and tasks.*

We appreciate the reviewer's comment and agree that a matched control group would be ideal. However, we addressed this issue using hierarchical regression analysis.

Our study assessed visual working memory (VWM) resolution using two analog recall paradigms: the sequential paradigm with bar stimuli and memory-guided localization (MGL). While the demographic data of gender, age, and education in the MGL paradigm were matched between patients and control group, there was a significant difference in these factors between groups in the sequential paradigm.

To address this issue, we performed hierarchical regression analysis to compare recall parameters in the sequential paradigm with 3-bar and 1-bar stimuli, respectively. We assessed for the confounding effect of gender, age, and education, and the results were presented in supplementary tables 3 and 5.

In the sequential paradigm with 3-bar stimuli (high memory load condition), we found that all recall parameters were significantly different between groups. However, after adjusting for age and education, the result became insignificant for uniform response proportion. In the 1-bar paradigm (low memory load condition), while the results were significantly different between groups, after adjusting for gender, age, and education, target and uniform response proportions became insignificant (uniform proportion = $1 - \text{target proportion}$, since there was no swap error in the 1-bar condition).

1. The authors interpret the mixture model parameter described as "misbinding error" as reflecting failures of feature binding, and propose a link to hippocampus on that basis, however there is now quite strong evidence that these errors (often called swaps) are explained mostly or entirely by imprecision in memory for the cue feature (bar color in this case), e.g. McMaster et al. (2022), already cited in the ms.

We thank the reviewer for this valuable comment regarding interpreting the mixture model parameter, described as a "misbinding error" in our study.

Swap error has been attributed to different mechanisms, including the variability in cue feature dimension, cue-independent sources, and strategic guessing. As the reviewer mentioned, in a recent study by McMaster et al., a comprehensive evaluation of these hypotheses was performed and determined that the variability in cue feature dimension could solely explain the occurrence of swap error.

In response to this comment, we have added a discussion of this matter, the neural correlates of swap error, and the possible explanation for this phenomenon in multiple sclerosis (MS) population to the seventh paragraph of the discussion. Additionally, since our study did not include neuroimaging assessment, we have discussed the results from neuroanatomical points of view to further explain the possible structures involved in the occurrence of swap errors in MS. The seventh and eighth paragraphs of the discussion have been revised for further clarification.

1. The methodology of the ROC analyses should be described in more detail: it is not clear what measures are being used to classify participants or how.

This matter is clarified in the results and the last paragraph of materials and methods. In both paradigms, recall error was used for classification purposes.

1. There are a number of unusual choices of terminology that could potentially confuse or mislead the reader: The tasks are not "n-Back" tasks by the usual meaning; they are analog report tasks with sequential presentation. The terms recall "error", "variability", "precision" and "fidelity" are used idiosyncratically. Variability and precision usually refer to the same thing; they describe the dispersion or spread of errors. The measure described as recall error in the sequential tasks is presumably absolute (or unsigned) error. For the mixture model parameters I suggest describing them more explicitly in terms of the mixture attributes, e.g. "Von Mises SD", "Target proportion", "Non-target proportion" "Uniform proportion".

We thank the reviewer for this suggestion. We have made revisions to clarify the terminology used in our study.

The term "n-back" is changed to an analog recall paradigm with sequential presentation. Additionally, as mentioned in the materials and methods, the recall error in the MGL paradigm is the Euclidian distance between the target's location and subject response in visual degree. In the sequential paradigms, this value is the angular difference between the response and target value, in which both are absolute errors. To avoid confusion, we have added the term "absolute error" alongside the term "recall error" to provide a clear understanding of this measurement. Moreover, as the reviewer suggested, we changed "recall variability" to "von Mises SD" for a more precise description. We also changed the remaining terms to "target proportion", "swap error (non-target proportion)", and "uniform proportion".

Reviewer #2 (Public Review):

The authors applied two visual working memory tasks, a memory-guided localization (MGL), examining short-term memory of the location of an item over a brief interval, and an N-back task, examining orientation of a centrally presented item, in order to test working memory performance in patients with multiple sclerosis (including a subgroup with relapsing-remitting and one with secondary progressive MS), compared with healthy control subjects. The authors used an approach in testing and statistically modelling visual working memory paradigm previously developed by Paul Bays, Masud Husain and colleagues. Such continuous measure approaches make it possible to quantify the precision, or resolution, of working memory, as opposed to measuring working memory using discretised, all-or-none measures.

The authors of the present study found that both MS subgroups performed worse than controls on the N-back task and that only the secondary progressive MS subgroup was significantly impaired on the MGL task. The underlying sources of error including incorrect association of an object's identity with its location or serial order, were also examined.

The application of more precise psychophysiological methods to test visual working memory in multiple sclerosis should be applauded. It has the potential to lead to more sensitive and specific tests which could potentially be used as useful outcome measures in clinical trials of disease modifying drugs, for example.

However, there are some significant limitations which severely affect the scientific validity and interpretability of the study:

1. There is a striking lack of key clinical information:

1.1) There is a striking lack of key clinical information. The inclusion and exclusion criteria are unclear and a recruitment flowchart has not been provided. Therefore it is unclear what proportion of MS patients were ineligible due to, for example, visual impairment.

We thank the reviewer for raising this matter. To address this issue, we revised the first section of materials and methods to include detailed inclusion/exclusion criteria information. However, it is important to note that we recruited the patients in a full-census manner, where only the patients who fulfilled the inclusion criteria participated. Unfortunately, we did not record the number of patients who did not meet the inclusion criteria.

1.2) Basic clinical data such as EDSS scores, disease duration, treatment history, and performance on standard cognitive testing were not provided. Basic clinical and demographic data for each subgroup were not provided in a clear format. This severely limits the interpretability of the study and its significance for this clinical population. For example, might it be that the SPMS patients performed worse on the MGL task because they were more cognitively impaired than RRMS patients? That question might be easily answered, but the answer is unclear based on the data provided.

We appreciate the reviewer for bringing up this important concern. To further clarify the basic clinical and demographic data, we have revised tables 1 and 2 to include detailed information regarding gender, age, education, cognitive ability, disease duration, EDSS score, and disease-modifying therapy (DMT) for each group, respectively. The information is reported as mean $\pm$ standard deviation except for the categorical data.

Regarding the participants' cognitive ability, we added the Montreal cognitive assessment test results for both paradigms. MoCA is a standard cognitive screening tool that has a score of 0 to 30. The different ranges of MoCA scores related to the different levels of cognitive function, in which a score ≥ 26 is considered normal cognitive ability, 18-25 denotes mild cognitive impairment, 10-17 determines moderate cognitive impairment, and a score ≤ 10 is considered severe impairment.

First, we classify the participants based on their MoCA value and compare groups with each other. While the primary results showed that patient groups were more impaired compared to healthy controls, our results remained significant after adjusting for MoCA using hierarchical regression analysis. This suggests that the observed difference was not solely due to more cognitive impairment in the patients' population.

Moreover, the information regarding the treatment history of patients is added in the following format. DMT is classified into two groups, i.e., platform and non-platform treatments. In our study, the platform treatments include interferon beta-1a and glatiramer acetate, and non-platform treatments include rituximab, ocrelizumab, fingolimod, dimethyl fumarate, and natalizumab. In both paradigms, the patients did not significantly differ based on the received therapy. The MoCA assessment and treatment history information is added to tables 1 and 2 and supplementary tables 1, 3, and 5. Moreover, the second paragraph of materials and methods, second paragraph of statistical analysis in materials and methods, and the appropriate sections of the results are revised.

1. The study is completely agnostic to the underlying pathophysiology. There is no neuroimaging available, therefore it is unclear how the specific working memory impairments observed might relate to lesioned underlying brain networks which are crucial for specific aspects of working memory. This severely limits the scientific impact of the results. This limitation is acknowledged by the authors, but the authors did not put forward any hypotheses on how their results may be underpinned by the underlying disease processes.

We appreciate the reviewer for this valuable suggestion. To further strengthen the connection between our findings and the possible underlying mechanisms of WM dysfunction in MS, we have added a discussion from the neuroanatomical perspective in the eighth paragraph of the discussion section.

1. The present study does not compare the continuous-report testing with a discrete measure task so it is unclear if the former is more sensitive, or more feasible in this patient group, although this may not have been the purpose of the study.

The reviewer pointed out an interesting matter. However, this was not the focus of the current study. Nonetheless, it is a valuable suggestion for future work to compare continuous vs. discrete measure paradigms to determine their sensitivity and feasibility in the MS population.