Semantic interference mechanisms on long-term visual memory and their eye-movement signatures in mild cognitive impairment

Semantic interference mechanisms on long-term visual memory and their eye-movement signatures in Mild Cognitive Impairment

Moreno I. Coco[1,2 *], Gabriella Merendino[3], Giuseppe Zappala'[3] and Sergio Della Sala[4]

[1]School of Psychology, The University of East London, London, UK

[2]Faculdade de Psicologia, Universidade de Lisboa, Lisbon, Portugal

[3]Neurological Ward, Dementia and related disorders Unit, Garibaldi Hospital, Catania, Italy

[4]Human Cognitive Neuroscience, Psychology, University of Edinburgh, Edinburgh, UK

* Correspondence should be sent to moreno.cocoi@gmail.com

Abstract

Objective: Long-term visual memory representations, measured by recognition performance, degrade as a function of semantic interference, and their strength is related to eye-movement responses. Even though clinical research has examined interference mechanisms in pathological cognitive ageing and explored the diagnostic potential of eye-movements in this context, little is known about their interaction in long-term visual memory.

Method: An eye-tracking study compared a Mild Cognitive Impaired group with healthy adults. Participants watched a stream of 129 naturalistic images from different semantic categories, presented at different frequencies (1, 6, 12, 24) to induce semantic interference (SI), then asked in a 2-Alternative Forced Choice paradigm to verbally recognize the scene they remembered (old/novel).

Results: Recognition accuracy of both groups was negatively impacted by SI, especially in the healthy adults. A wider distribution of overt attention across the scene predicted better recognition, especially by the MCI participants, although these fixation patterns were influenced by SI. MCI compensated the detrimental effect of SI by focusing overt attention during encoding and so accruing distinctive details of the scene. During recognition, MCI participants widened overt attention to boost retrieval. Independently of the group: (a) the re-instatement of fixations indicated a more successful recall and increased as a function of SI; and (b) attending visually salient regions negatively impacted on recognition accuracy, although the reliance on such regions grew as SI increased.

Conclusions: Effects of SI on long-term memory were reduced in MCI participants. They used different oculo-motor strategies compared to healthy adults to compensate its detrimental effects.

Key Points:

Question: Is visual memory of Mild Cognitive Impaired participants impacted by semantic interference similarly to healthy adults, and will their eye-movements reveal any difference?

Findings: Semantic interference effects are reduced in people with MCI and their eye-movements show subtle compensatory strategies associated to the formation and access of visual memories.

Importance: Visual semantic interference may be an interesting candidate mechanism to better study pathological memory processes.

Next Steps: Better characterise the nature of visual interference by developing new metrics to disentangle its perceptual and conceptual components.

Keywords: Long-term visual memory; Mild cognitive impairment; semantic interference; low-level visual saliency; eye-tracking

Semantic interference mechanisms on long-term visual memory and their eye-movement signatures in Mild Cognitive Impairment

Introduction

In a seminal study, Standing (1973) demonstrated that humans can remember thousands of images, highlighting our impressive capacity to retain and recollect rich and detailed visual information from memory. After this initial observation , which became accepted wisdom (Reed, 2012), the topic of long-term visual memory has been relatively neglected.

In the last two decades, however, probably fuelled by advances in computational methods, there has been a surge of interest on what makes images memorable (Bylinskii et al., 2015; Isola et al., 2014) and on the cognitive mechanisms involved in the encoding and successful access of visual information from memory (e.g., Bainbridge et al., 2019; Evans & Baddeley, 2018; see Rust & Mehrpour, 2020, for a recent review). An insightful observation by Brady and co-workers is that the successful later recognition of an object, or a scene, strongly depends on how many other objects (or scenes) from the same semantic category were also encoded in memory (Brady et al., 2008; Konkle et al., 2010a, 2010b). In these studies, participants were asked to watch a stream of visual stimuli (i.e., scenes) from different semantic categories (e.g, kitchens or deserts) at different frequencies (1, 4, 16, 64). Their memory was then tested in a 2-Alternative Forced Choice paradigm, whereby a seen image (old) was presented side-by-side with an unseen image (novel) from the same semantic category. The results showed that the fidelity of visual representations in memory degrades as a function of number of related images: the greater the number of images from the same semantic category were originally encoded, the worse their retrieval, suggesting a semantic interference effect on memory recognition.

The impact of interference on memory representations is not a new discovery, as it has been a topic of investigation since very early research (Müller & Pilzecker, 1900; for a review see Dewar et al., 2007). Several factors contributing to interference have been identified (e.g.,

retroactive vs. proactive, see Craig et al., 2015 for an overview of the distinction). Here, we focus on the interference on memory recognition that is generated by the similarity of stimuli at encoding (see McGeoch & McDonald, 1931 and Baddeley & Dale, 1966 for early work on semantic interference in verbal recall tasks).

Semantic memory and the mechanisms of interference have played a prominent role also in clinical research investigating the neuropsychological markers of pathological ageing. People affected by mild cognitive impairment (henceforth MCI), with ascertained memory impairments (Mitchell & Shiri-Feshki, 2009), often leading to more severe forms of Alzheimer's type dementia (henceforth AD), display a degraded performance in tasks tapping into semantic memory, such as word-to-picture matching (Adlam et al., 2006) or picture naming (e.g., Duong et al., 2006), and have greater difficulties recalling historical events (e.g., Leyhe et al., 2010) or faces of famous people (e.g., Barbeau et al., 2012). When looking more specifically at semantic interference effects, however, evidence is not as clear cut. Loewenstein et al. (2004) compared people with MCI, AD, and healthy adults on the Fuld Object Memory Evaluation task, which measures verbal memory recall of visually and tactilely presented real objects, and observed a clear effect of semantic interference in all groups. In this study, participants were first exposed to 10 common objects (and asked to touch, watch and name), then similarly presented with 10 novel objects from the same semantic category and asked to verbally recall both lists of objects. The results showed that the verbal recall of MCI, and especially AD participants, was significantly more impacted by the semantic interference than the healthy age-matched control group, even after accounting for differences in overall memory function. On the other hand, Mulatti et al., (2014) adopted a picture naming task, devised by Howard et al. (2006), and compared cumulative semantic interference effects in people with MCI and healthy adults. In line with Howard et al., the results showed that the naming latencies for healthy adults became progressively slower as the number of objects from the same semantic category increased; however, such an effect was not observed in the MCI participants. The differential semantic interference effects observed in these two studies may relate

to the evidence that people with AD show preserved information about the use of objects even when unable to name them (Bartolo et al., 2016). Hence, they may experience interference during the Fuld task, because it taps into specific object knowledge. Regardless of these mixed results, it has yet to be determined whether the detrimental effect exerted by visual semantic interference on the long-term memory of younger adults (Konkle et al., 2010b) would manifest in a healthy older population, and in people with MCI, and if so, what would the magnitude of this effect be. Indeed, there is only a handful of studies investigating long-term visual memory in people suffering from dementia, and despite some preliminary evidence that their capacity is well kept (e.g., Karlsson et al., 2003), it is a largely unexplored area of research. Our first objective is to compare long-term memory for naturalistic scene information for people with MCI and healthy age-matched adults, while examining the impact that semantic interference may exert.

Another important aspect about the memorability of images relates to the role played by extrinsic responses such as eye-movements (Bylinskii et al., 2015) and their link to memory processes (see Hannula, 2018, Hannula et al., 2010 and Ryan & Shen, 2020, for comprehensive reviews). For example, a more spread-out distribution of fixations on a scene upon first inspection predicts better later recognition (e.g., Damiano & Walther, 2019). In addition, memory retrieval is boosted by attending to similar locations (i.e., fixation re-instatement) during encoding and recognition (e.g., Foulsham & Kingstone, 2013), especially for older adults (e.g., Wynn et al., 2018). What it is still unclear is whether eye-movement responses could additionally inform about the impact of higher-level mechanisms, such as semantic interference, on memory processes, and further, help reveal any differences associated with pathological ageing.

This proposition is supported by recent clinical research aimed at isolating abnormal patterns of oculo-motor control in pathologically aged populations (see Molitor et al., 2015 for a review). In this literature, most differences on eye-movement patterns seem to be on saccadic responses and using simple tasks, such as fixating at dots placed at different eccentricities, which mostly tap into low-level mechanisms of oculo-motor control (Fletcher & Sharpe, 1986; Yang et al.,

2013; but see Boucart et al., 2014 for an example of a search task using naturalistic visual stimuli). So, even though oculo-motor responses reflect memory processes and may reveal underlying pathological conditions (Lagun et al., 2011), it is yet to be established whether high-level semantic interference mechanisms may manifest in eye-movement patterns and differ because of pathological cognitive ageing. Addressing this question constitutes the second objective of the current study.

Finally, visual information does not only carry high-level semantic information, but also, low-level perceptual information (e.g., colour, edges or luminosity), which can be computationally quantified as visual saliency (see Itti & Koch, 2000, for a well-known model). This type of information is known to guide the allocation of overt attention when the task has no specific target objects to look (e.g. free-viewing, Parkhurst et al., 2002), and has little impact on the memorability of an image (Isola et al., 2011). In the context of pathological ageing, only a few studies have examined whether the access and use of low-level information may be compromised. What is known originates primarily from visual search tasks, where people with dementia seem to show difficulty searching for visually salient targets (Tales et al., 2004) and need enhanced low-level perceptual features (e.g., contrast) to perform at a comparable level to healthy adults (Cronin-Golomb et al., 2007). However, to the best of our knowledge, it is yet to be established how visual saliency may influence the long-term visual memory of people with MCI. This constitutes the third objective of the current study.

The present study borrows the paradigm developed by Konkle et al., (2010b), in order to assess the magnitude of semantic interference effects on memory recognition in people with MCI and a healthy control group. It also examines a few key eye-movement measures to isolate: (a) how, and to what extent, the allocation of overt attention contributes to memory processes, (b) whether they reveal clear differences between healthy adults and people with MCI, (c) if semantic interference leaves a signature on such measures and (d) if low-level visual saliency can contribute to explaining pathological memory processes.

If semantic interference is a robust mechanism, spared by healthy and pathological cognitive ageing, we would expect to replicate the findings of Konkle et al. (2010b), as well as our own results with younger adults (Mikhailova et al., 2021). These studies showed that the greater the frequency of the semantic category a scene is taken from, the worse is the subsequent retrieval of such a scene. If pathological ageing instead significantly impacts on the mechanisms of semantic interference, in light of the mixed results discussed above, we may observe either a greater effect on memory recognition in the MCIs (Loewenstein et al., 2004) or a reduction of this effect (Mulatti et al., 2014).

Moreover, if oculo-motor responses are significantly affected by pathological ageing, as the literature on the topic seems to indicate (Molitor et al., 2015), we would expect differences between MCI participants and healthy adults to manifest also in eye-movement measures. In particular, MCI participants may not show a positive effect of fixation re-instatement on memory recognition, contrary to expectations on younger participants and healthy older adults (Foulsham & Kingstone, 2013; Wynn et al., 2018). Other differences between groups may emerge regarding how fixations distribute across the scene while encoding or recognition takes place. To address general patterns of viewing, we use attention maps (Henderson, 2003), which are two-dimensional aggregates of all fixations that occurred on a given scene, and use the information measure of entropy to extract a summary measure of their spread (refer to Method for greater details). Conceptually, a higher fixation entropy in the attention map would reflect a wider distribution of attention than an attention map with a lower fixation entropy. In the context of the current study, and in line with our study on younger adults (Mikhailova et al., 2021), during encoding, we expect fixation entropy to decrease as semantic interference increases. As the fidelity of memory representations blurs due to semantic interference, overt attention needs to be allocated onto fewer and more specific details of the scene that can act as distinctive cues to enhance its later memorability. This compensatory mechanism maybe more strongly required in MCI participants, whose attentional capacity to keep up with increasing semantic interference may be reduced. During recognition, on the other hand, we expect

fixation entropy to increase as a function of semantic interference and so maximize information recall i.e., larger sections of the scene need to be viewed to boost its recognition. In this phase, this compensatory mechanism may be more marked in the MCI group, capturing the evidence of more peripheral fixations (Rösler et al., 2005), but fail to significantly improve on their recognition performance.

With respect to visual saliency, previous work shows that it does not contribute to the memorability of images (Isola et al., 2011), but we may expect people with MCI instead to allocate their overt attention more prominently onto salient regions compared to healthy adults as an attempt to enhance low-level processing of visual information (Cronin-Golomb et al., 2007).

Methods

Participants

A total of 54 participants with a diagnosis of Mild Cognitive Impairment (MCI), a potential prodromal stage of Alzheimer's Disease, and 31 healthy age-matched controls, all native Italian speakers, were recruited from the Neurological Ward, Dementia and related disorders Unit at Ospedale Garibaldi, Catania (Italy). All participants were naive to the purpose of the current study and the stimuli used therein; they all voluntarily took part in the study; they received no honorarium and gave explicit informed consent. In Figure 1, we display a breakdown of the selection criteria applied to obtain the sample that was statistically analysed. We had to exclude 4 MCI participants and 4 healthy adults because they did not return for the follow-up session (more details below) and a further 23 MCI participants and 4 healthy adults because their recognition performance on the visual memory task was at chance. The MCI participants excluded for at-chance performance on the recognition memory had a similar age (Excluded = 70.68 ± 5.86, Included = 72.48 ± 9; $t(45) = -0.84$, $p = 0.4$) and years of education (Excluded = 10.27 ± 4.57, Included = 10.52 ± 5.19; $t(45) = -0.17$, $p = 0.9$) to the participants included. However, they had marginally lower performance on the MMSE (marginal) and significantly lower performance on the CORSI and on the REY for the

Delayed condition (see Table S1 in the Supplementary Material for a full comparison). This indicates that the long-term memory of the excluded participants was particularly compromised, and consequently were unable to perform the long-term visual memory task. The remaining 27 MCI participants (7 women, age = 72.48 ± 8.99) and 23 healthy adults (14 women, age = 68.08 ± 9.66) contributed to the analyses of the recognition performance. The eye-tracking data of these participants had to be further selected for quality purposes. We excluded 3 MCI participants and 3 healthy adults because we could not perform the first calibration, and 1 more MCI participant because of data loss in his/her eye-tracking record (> 30% of the total). Therefore, 23 MCI participants and 20 healthy adults contributed to the analyses of the eye-movement data. The diagnosis of MCI was structured following international guidelines (Arnáiz et al., 2004; Gauthier et al., 2006; Petersen, 2016) and based on the Mini Mental State Examination (see Grigoletto et al., 1999; Measso et al., 1993 for Italian norms) raw score ≥ 18, family and medical history interviews, as well as imaging data (e.g., MRI) and genetic data (e.g., ApoE), when available. Other inclusion criteria for the current study were: 1) between 50 and 90 years of age; 2) no less than 3 years of schooling (MCI = 10.51 ± 5.19; Control = 10.73 ± 4.92); 3) normal or corrected-to-normal vision with no history of eye surgery; 4) no history of neurological (other than memory disturbances) and/or psychiatric disorders; 5) no history of alcohol or substance abuse and/or use of medications likely to affect cognitive functioning; 6) able to understand the instructions and perform the task. The two groups were matched on age [$t(45.48) = 1.65$, $p = 0.1$)] and years of education [$t(47.40) = -0.15$, $p = 0.9$)]. In order to gather a better picture of the neuropsychological profiles of the two groups, we administered the following battery of tests spanning different cognitive functions: (a) Digit span and Corsi's Block Tapping Test to assess attention and short-term memory (verbal and non-verbal respectively), (b) Rey's Complex Figure, copy and delayed to assess visuo-spatial organization and non-verbal long-term memory, (c) Mesulam's Cancellation test to assess eye motor coordination and visuo-spatial scanning (letters and symbols) and (d) the Dubois' 5 Words test to assess free and cued recall (Dubois et al., 2002; see Girtler et al., 2012 for an application of

this test to an Italian sample). In Table 1, we report the descriptive statistics of these tests for the two groups, and assess statistical significance using a Welch two sample t-test (p-values are reported in the table). In Appendix A, we report a correlation analysis of the MCIs' individual performance on the neuropsychological tests, recognition accuracy, and semantic interference. The NHS South East Scotland Research Committee approved the protocol and material of this study prior to starting the data collection (Ref: 16/SS/0109).

INSERT FIGURE 1 AND TABLE 1 ABOUT HERE

Design and Stimuli

We selected 834 naturalistic images from the SUN database (Xiao et al., 2010) with a minimum of 550*550 pixels resolution, which are provided as thumbnails in the Supplementary Material. In the first session, we used 576 scenes from this pool drawn from 12 different semantic categories, 6 human-made environments (i.e., amusement park, bathroom, gas station, highways, kitchen, library) 6 six natural environments (i.e, beach, desert, field, forest, mountain, river). Similar to Konkle et al., (2010b) the frequency of scenes from each semantic category that were viewed by the participants was manipulated to induce semantic interference effects. The twelve types of semantic category scenes were randomly distributed with frequencies of 24, 12, 6, or 1 within the encoding and recognition phases. In the analyses, we treated SI as a continuous rather than categorical variable, with the aim of capturing its incremental impact on memory representations. For example, if the semantic category *highways* were set to have a total frequency of 6 for a certain participant, it meant that she/he viewed 6 different *highways* trials, randomly presented across the stream, during encoding. Then, during recognition she/he was tested on these same (old) 6 *highways* trials, plus a novel *highways* image displayed side-by-side, also randomly presented. Continuous SI, in this example, was obtained by incrementally counting the number of *highways* administered, i.e., from 1 to 6 for the encoding phase and from 7 to 12 for the recognition phase,

which correspond to the maximum number of *highways* trials a participant will be administered in this running example during her/his entire session.

Each participant saw 129 scenes during the encoding phase, and the same 129 (old) scenes side-by-side with 129 novel scenes from the same semantic categories during the recognition phase. In order to counterbalance for the amount of interference for each semantic category, we rotated the scenes in 4 different lists (e.g., if the kitchen category had an SI of 4 in list 1, the same category had an SI of 20 in list 2, and so on). Four further lists were then created by swapping old with novel scenes between the encoding and the recognition phase.

In a follow-up session, the same participants were called back to do the same task, after an average number of months of 3.88 ± 2.86 (SD), and administered the remaining 258 scenes from the original 834 image pool (129 for the encoding phase, and 129 novel scenes for the recognition phase). These scenes belonged to 129 different semantic categories according to the SUN classification, and hence were not expected to result in across-scenes semantic interference within a session. The purpose of this follow-up session was to estimate a by-participant baseline memory capacity, i.e. no-interference, and so compare it to the main session where the frequency of scenes for each semantic category was instead manipulated (refer to the Analysis section for more details). Two lists for this follow-up session were created by swapping old with novel scenes between the encoding and the recognition phase so to ensure that all scenes were seen in both conditions. Note also that the number of scenes administered was identical in the main session and the follow-up session (i.e., 129), but scenes were not repeated between sessions, i.e., we did not reuse scenes between the main and the follow-up session. SI was analysed following the same logic described above, but now, we have a frequency value of 1 for all encoding trials, as all scenes come from different semantic categories, and a value of 2 for all recognition trials as we are including in the count the trial viewed during encoding.

Apparatus

Visual stimuli were displayed on an AOC 19.5 inch flat-screen monitor (44.4 cm width and 23.7 height) with a 50Hz refresh rate at a resolution of 1600 × 900, and eye movements binocularly recorded using an EyeTribe eye-tracker (55 Hz sampling rate). The spatial resolution of the EyeTribe system reported by the manufacturer is 0.1 root mean square (RMS) and its accuracy and precision for scientific research validated by Ooms & Krassanakis, (2018). The experiment was implemented using OpenSesame (Version 3.1.9, Mathôt, Schreij, & Theeuwes, 2012) and the PyGaze Python plug-in (Dalmaijer et al., 2014) was used to acquire the eye-movement data. Each participant was calibrated on a 9-points grid at the beginning of each experimental session, and recalibrated if necessary. The visual angle deviation error (mean and standard deviation) accepted at the calibration for the two groups were*:* 0.79º ± 0.61º on the x-axis and 0.65º ± 0.38º on the y-axis for the healthy adults, and 1.24º ± 0.93º on the x-axis and 1.04º ± 0.57º on the y-axis for the MCI group.

INSERT FIGURE 2 ABOUT HERE

Procedure

During the encoding phase, each participant watched the stream of 129 images, randomly distributed (i.e., the semantic interference manipulation was not blocked). Each scene was presented for 3 seconds with an 800 ms fixation crosshair inter-trial interval, and semantic interference was manipulated as previously described. Participants had a 10 minutes break after the encoding phase. In the recognition phase, participants were shown 258 images side-by-side (i.e., 129 trials), half of which had been seen during encoding (old images) and the other half were novel (new images), again in randomized order and with equal frequency per semantic category (e.g., for 6 kitchen images in the encoding phase, there were 6 old and 6 novel kitchen images in the recognition phase). In this phase, on each trial, participants were asked to choose the scene they remembered by saying out loud either *uno* (Italian for number *one*) or *due* (number two) and so indicate either the left or

the right image respectively (see Figure 2 for an illustration of the task). The experimenter would log the response by pressing either 1 or 2 on the keyboard. The follow-up session followed the same experimental procedure but at its end, participants were now administered the battery of neuropsychological tests described above.

The size of the images was not the same for the encoding and the recognition phase. During encoding, the image was presented in the centre of the screen at a resolution of 700 × 700 pixels, which corresponds to 16.38° × 16.38° degrees of visual angle. During recognition, the two images were shown at a resolution of 560 × 560 off-centre, to the left and to the right (i.e., 12.97° × 12.97° of visual angle each picture) at an equal distance of 280 pixels from the centre of the display (i.e., 6.57° of visual angle). This distance guarantees that each image was presented in extra-foveal vision during a recognition trial, and so participants had to make a saccade towards it to explore an image of the pair. The long-term visual memory task lasted approximately 30 minutes, and the follow-up session about 30 minutes longer, due to the neuropsychological assessment.

Data analysis

Data processing

Out of the 12,900 recognition trials considered in the analysis (i.e., 50 participants × 129 scenes × 2 sessions), we excluded 5 images (23 trials) that were recognized at (or below) chance level, and 150 further trials that had a response time either faster than 1% or slower than 99% of all trials, as computed independently for each participant. The number of recognition trials analysed was 5,854 trials for the healthy adults, with a by-participant average of 254.52 ± 0.79, and 6,873 trials for the MCI group, with a by-participant average of 254.55 ± 0.75.

All eye-movement data was collected after performing calibration at a low-resolution (55 Hz), which approximately corresponds to a sample every 33 ms. For this reason, we could not use any velocity/speed threshold algorithm to extract saccade onsets, and consequently detect fixations, because most saccades are likely to be as fast as, or even faster than a single datapoint. Moreover, the eye-movement measures that we are interested in for this study do not relate to any specific area-of-interest but rather tap into global patterns of viewing behaviour, and so we opted to compute our dependent measures directly

from the raw samples. Out of the total 10,794 eye-movement trials (43 participants × 129 trials × 2 sessions), we excluded a further 1,945 trials from the encoding phase and 2,177 trials from the recognition phase because of machine error or when more than 35 % of the eye-tracking samples of the trial were out of range. Thus, for the encoding phase, we considered a total of 4,147 trials for the healthy adults with a by-participant average of 207.35 ± 67.04 [48 – 255], and a total of 4,702 for the MCI group with a by-participant average of 204.43 ± 59.88 [116 – 255]; for the recognition phase, we considered a total of 4,104 for the healthy adults with a by-participant average of 205.2 ± 68.35 [12 – 255] and a total of 4,513 trials for the MCI group with a by-participant average of 196.21 ± 67.77 [39 – 255].

INSERT FIGURE 3 ABOUT HERE

Dependent variables

The impact of semantic interference on the memorability of images is statistically analysed as: *recognition accuracy* (a binary variable coded as 0 = Incorrect; 1 = Correct), which reflects the choice (1, 2) of the seen (old) image, and *interference slope*, which we calculate, independently for each participant, by fitting a linear model predicting recognition accuracy as a function of the frequency of scenes in their semantic category (i.e., SI) expressed as a continuous variable. A negative beta coefficient associated to the frequency manipulation implies that the memory performance worsens with additional scenes of the same category (see Konkle et al., 2010a, for a similar approach). Additionally, we investigate whether the impact of semantic interference may relate to the overall recognition accuracy of each participant. To do so, we created a quasi-experimental variable (Performance) by classifying each participant in either low-performer or high-performer based on the median of her/his reference group, i.e., independently for MCI and healthy adults. This further analysis is performed on a *d-prime* transformation of the recognition accuracy, and on the *recognition difference* of each participant under interference (6, 12, 24) against her/his accuracy in the follow up no-interference session (i.e., interference of 1).

The dependent measures computed from the eye-movement data and explained in greater details below are: (a) the *entropy* of the fixation distribution displayed by a participant while encoding or recognizing a seen (old) image, (b) the *scan-pattern similarity* between the eye-movements of the participant viewing the same scene at encoding and recognition and (c) the *Normalized Scanpath Saliency* (NSS, Peters et al., 2005) to measure the correspondence between fixation positions and low-level visual saliency (see Figure 3 for an illustration of the approaches used to compute fixation entropy and visual saliency). For (a) the fixation probability map was computed for each trial by fitting a Gaussian filter at the coordinate location of each fixation sample with a standard deviation of 1º of visual angle to approximate the size of the fovea. Then, the entropy of the resulting fixation probability map was calculated as $\sum_{x,y} p(S_{x,y}) log_2 p(S_{x,y})$, where $p(S_{x,y})$, is the normalised fixation probability at the coordinates of the fixation (x, y) in the scene *S* (see Castelhano et al., 2009 or Coco & Keller, 2014, for related examples). For (b), a grid of 5*5 equal sized quadrants fully covering the size of each scene was used to map the coordinates of each fixation sample to each quadrant uniquely identified with a categorical label (e.g., *aa*, *ab*, *ac*, etc., see Nuthmann & Einhäuser, 2015 for a similar approach). In this way, we obtained a categorical sequence of fixated locations (i.e., a scan-pattern). Then, we used the Longest Common Subsequence (LCS, Gusfield, 1997), to find the longest ordered common subsequence of the two scan-patterns (encoding and recognition) and derived a similarity score by taking the ratio between the length of the LCS and the geometric mean of the lengths of the two sequences. The similarity value obtained ranges from 1 for most similar to 0 for least similar (see Coco & Keller, 2012 for an application to eye-movement data). For (c), a low-level visual saliency map of each scene was computed using Fast and Efficient Saliency model (FES, Tavakoli et al., 2011), which was standardized to have zero mean and unit standard deviation before taking the average of

saliency values corresponding to all fixation positions observed in each trial, i.e., the NSS (Peters et al., 2005).

The data and R script to run the analyses and visualise the results of this manuscript are available on the Open Science Framework at https://osf.io/x6jbs/

Statistical analyses

Linear and generalized linear mixed-effects models (G/LMM), as implemented by the `lme4` package (Bates et al., 2015) in `R` (version 4.0.2), were used to analyse the data. In our modelling approach, we started with a full fixed-effect structure (i.e., all main effects and interactions) and a maximal random effect structure (i.e., random variables included both as intercepts and uncorrelated slopes; Barr et al., 2013). Then, fixed and random parameters were evaluated and backwards-reduced using the `lmerTest` package (Kuznetsova et al., 2017) to retain the model that was parsimonious in the number of parameters (Matuschek et al., 2017). The main predictors centred to reduce collinearity (see Jaeger, 2008, for an explanation of this strategy) were: (1) Group (Control = -.5, MCI = .5), (2) SI, standardised using z-scores and computed as a frequency of each particular scene category, ranged from 1 to 24 for the encoding data, and 2 to 48 for the recognition data, which are the maximum number of scenes belonging to the same category that could be sequentially seen up to the end of the two phases respectively[1]. For the analyses of d-prime and recognition difference we treat SI as a categorical, rather than as continuous variable (1 – reference level; 6 - for the difference score analysis) because these dependent measures are aggregated by participant, and we add the Performance (Low, High; Low - reference level) as a predictor in the model to examine whether the effect of semantic interference

[1] If the semantic interference condition is 1, a participant would view a scene at encoding, and then again side by side with a novel scene from the same category (2AFC) during recognition (i.e., 1 + 1 = 2). If the semantic interference condition is instead 24, a participant would view 24 scenes from the same category at encoding phase, and 24 novel scenes at recognition , i.e., a possible maximum of 24 + 24 = 48.

would be stronger (or weaker) according to the overall recognition performance of the participants. For the analysis of the dependent measures extracted from eye-movement data, we consider the Accuracy[2] (Incorrect = -.73, Correct = .26) which allows us to examine how such responses relate to the memorability of the images. The random effects considered in this study were Participants (50 for the model predicting response accuracy, and 43 for the models predicting eye-movement measures) and the semantic Category of the scene (141; 12 from the main session and 129 from the follow-up session). Moreover, we computed standardized beta estimates, which make it possible to compare the relative effect of each predictor in relation to the dependent variable, using the package `effectsize` (Mattan et al., 2020) and calculated effect sizes such as Cohen's *d*, using the package `esc` (Lüdecke, 2019). In the tables, we report the standardised (β) and unstandardized beta coefficients (*B*), standard errors (*SE*), and *t*-values only for those predictors that were significant after model selection. The *p*-values are based on Satterthwaite approximation of the effective degrees of freedom for LMMs models and on asymptotic Wald tests for GLMM models and their level of significance are reported in the Tables, next to the t-value, using asterisks (e.g., * = $p < 0.05$). The analysis of the *interference slope* was carried out using a general linear model because this measure was computed using by-participant linear models (see the Dependent Measure subsection above), and so it will be reported directly in the text. Finally, as gender was slightly unbalanced between groups (e.g., more males than females in the MCI sample) we included gender as a random effect in our LMM models, but this showed no significant improvement on the fit of the models for any of the dependent measures reported in this study[3].

[2] The values for incorrect and correct trials are not exactly -.5 and .5 because there are more correct than incorrect responses, and this difference arises more clearly when the variable is centred.

[3] In future studies, we aim at better matching gender within groups to explore more systematically the effects of gender in semantic interference.

Results

Recognition performance

Figure 4A shows the percentage of accurately recognized scenes as a function of semantic interference while Figure 4B displays the interference slope observed for the two groups. We found main effects of SI and Group, whereby the recognition of a scene decreased for increasing SI, and MCI participants displayed an overall poorer performance. Moreover, healthy adults were affected by SI more than MCI participants, as shown by the steeper decrease in memory recognition for increasing SI (refer to Table 2 for the model coefficients). The weaker effect of SI on recognition memory in the MCI group is confirmed on the interference slope, which is significantly more positive compared to healthy adults [$\beta = .014$, $t(49) = 2.76$, $p < .01$]. For the d-prime, we found it to be significantly lower in the MCI participants compared to the healthy adults, and stronger for the high-performing compared to the low-performing participants (see Figure 4C and refer to Table 3). D-prime was significantly lower only when interference was at its maximum level, i.e., 24, compared to no-interference (i.e., 1), but significantly less in the MCI group. Additionally, high-performing participants had significantly reduced d-primes at higher levels of interference compared to low-performing participants (i.e., 12 and 24). Crucially, we did not find any significant three-way interaction between group, performance, and SI, which indicates that the reduction of semantic interference effect in the MCI group may be ascribable to the low performers. On the recognition difference, we corroborated this overall pattern of results, as we found it significantly greater for increasing levels of SI (i.e., 12 and 24), but smaller in the MCI participants compared to the healthy adults, even though a greater difference was observed in the high performers of this group (see Figure 4D and refer to Table 3).

INSERT TABLES 2 AND 3; FIGURE 4 ABOUT HERE

Eye-movement behaviour

When looking at eye-movement measures and starting from the *entropy of fixation distribution* during encoding, we found a wider spread of attention (i.e., higher entropy) for correctly recognized scenes, which significantly narrows (i.e., lower entropy) as SI increased (two-way interaction Accuracy:SI). Fixation entropy was also higher on healthy adults compared to MCI participants for increasing SI (two-way interaction Group:SI). During recognition, we confirmed a higher fixation entropy for the correctly as compared to incorrectly recognised scenes. In contrast to what was observed during encoding, however, fixation entropy significantly increased as SI also increased, especially for correctly recognised scenes (two-way interaction Accuracy:SI), but significantly less for the MCI group when the scene was correctly recognized (three-way interaction Accuracy:Group:SI). This was the case even though their fixation entropy tended to be wider for correctly recognised scenes compared to healthy adults (marginal two-way interaction Accuracy:Group, refer to Figure 5 for a visualisation and Table 4 for the model output).

INSERT TABLES 4 AND 5; FIGURES 5 AND 6 ABOUT HERE

The *scan-pattern similarity* between encoding and recognition was higher when the scene was correctly recognized, and it increased as a function of SI (refer to Figure 6A and Table 5 for the model output). Crucially, there was no significant difference between the two groups either with respect to the main effect or the interaction with the other predictors.

Finally, the correspondence between fixation positions and saliency maps (i.e., *NSS*) was significantly lower, in both phases of encoding and recognition, for correctly recognized scenes (see Figure 6B and refer to Table 5). Moreover, only during encoding, reliance on low-level visual saliency significantly increased as a function of SI. On this measure, we also did not find any significant effect of group, either as a main effect or in interaction with any other predictor.

Discussion

Long-term memory for visual information has an incredible capacity, and its fidelity is known to depend upon intrinsic properties of the scenes (or objects), as well as reflected by extrinsic responses associated to the processing of such information (e.g., Bainbridge et al., 2019; Brady et al., 2011; Bylinskii et al., 2015; Goetschalckx et al., 2018; Isola et al., 2014; Standing, 1973). On one hand, a general high-level intrinsic property of a scene is the semantic category it belongs to, which was shown to interfere with memory recognition: the more scenes from the same category that are encoded by participants into memory, the less accurate is their recognition (Konkle et al., 2010b). On the other hand, the formation and access of memory has been shown to be significantly and consistently reflected by the patterns of eye-movement responses as stimuli are encoded or recalled from memory (e.g., Hannula et al., 2010; Ryan & Shen, 2020). Even though research on neurodegenerative disorders of old-age, e.g., MCI, has long been interested in the diagnostic aspect of semantic interference (e.g., Loewenstein et al., 2004; Mulatti et al., 2014), and recently regarded eye-movement responses as possibly revelatory about it (e.g., Lagun et al., 2011; Molitor et al., 2015), there has been, to the best of our knowledge, no research simultaneously addressing these two key insights, especially in the context of long-term memory for visual information.

The aim of this study was to advance our understanding about semantic interference mechanisms on recognition memory of visual information in healthy and pathological cognitive ageing while examining possible links between overt attention, low-level perceptual properties of visual information and memory processes. We adapted the experimental design developed by Konkle et al., 2010b to test and compare people with MCI and healthy age-matched controls. We implemented the same experimental manipulation of semantic interference, i.e., different number of scenes were presented for each semantic category, but their frequency was reduced in total numbers to make the task shorter and more feasible to our older groups of participants (i.e., 1, 6, 12, 24, instead of 1, 4, 16, 64). We also added eye-tracking to the procedure and so monitored eye-movement responses while scenes were firstly encoded, and then recognized. Additionally, we examined the correspondence between eye-movement and low-level visual saliency of the scenes to explore the role that stimulus-driven properties may bear on the overt allocation of attention and on memory processes.

We found that the recognition memory of both groups was susceptible to semantic interference effects, whereby increasing numbers of scenes viewed from the same semantic category (semantic interference) resulted into poorer recognition performance. This result corroborates the original finding by Konkle et al., (2010b) and our own results on a sample of young healthy adults but tested under a different memory paradigm (Old/New vs. 2AFC, Mikhailova et al., 2021). We critically contribute to this area of research by showing that the mechanism of semantic interference on memory recognition manifests also in older participants even though it appears to be significantly reduced for people with MCI. One possible explanation for this result is that MCI participants have an inefficient access to semantic knowledge and so do not experience cumulative interference effects, i.e., the fidelity of memory representations for visual scenes does not tend to blur, which resonates with the findings reported by Mulatti et al. (2014) using lexical material in an object naming

task. An alternative explanation of reduced semantic interference in the MCI group may relate to each individual performance on the recognition memory and on the associated neuropsychological profile. We explored this hypothesis by classifying participants into low- and high- performers based on the median accuracy of their respective group (healthy adults, MCI), and compared their d-primes, as well as, the recognition difference of each participant's individual performance with, and without, semantic interference (from the follow-up session). We observed a nearly identical effect of semantic interference on both high-performing groups but showed a marked difference in the low-performing groups, with the MCI group showing no-interference effect compared to the healthy adults. Even though this result suggests that the reduced effect of semantic interference in the MCI participants was driven by those with an overall poorer recognition accuracy, it still remains a possibility that the mechanism of semantic interference, per-se, was to some extent compromised as a result of their neuropathological condition. In fact, when looking at correlations between interference slopes and individual neuropsychological results of the MCI participants (see Appendix A), we found that those with a better performance on the 5 Words test (X5WT), which is a verbal recall task with a semantic cueing component, also were more strongly impacted by semantic interference [$r(25) = -.52$, $p = 0.008$]. This result, however, also cannot be conclusive as cognitive capacity (MMSE) and recognition accuracy in the visual memory task both positively correlated with the 5 Words test. Admittedly, these correlation analyses highlight two limitations of our study: (1) some of our participants may have presented with early AD, and (2) a more extensive neuropsychological testing was needed to better frame whether the cognitive impairment of our MCI participants was purely amnestic, or multifarious in nature. In general, it should be noted that several participants were excluded from both groups and especially in the MCI group (i.e., 54 % of the sample), which indicates that long-term memory capacity of visual information in older populations does not seem to be as striking as previous research

on younger participants showed (e.g., Brady et al., 2008; Isola et al., 2011; Standing, 1973). This observation also casts doubts on the reliability of previous research with AD patients demonstrating instead a rather high recognition performance by this population in a similar task (Karlsson et al., 2003). Future research should try to further reduce the total number of scenes participants are asked to memorise, so aiming at increasing their capacity, as well as manipulating the consolidation phase of the task to minimize retroactive interference, e.g., by utilising a rest condition (Dewar et al., 2012).

When looking at the eye-movement behaviour, we found some commonalities between the two groups but also some intriguing differences. Confirming previous literature looking at the similarity of scan-patterns between the encoding and recognition phase of the task, we found that a higher re-instatement similarity score was predictive of better memory recognition (Foulsham & Kingstone, 2013; Wynn et al., 2018). Moreover, to corroborate this finding, we can report that fixation re-instatement increased as a function of semantic interference, which indicates that as the fidelity of memory representations degrades, there is a greater need to rely on the sequential oculo-motor coding of visual information to help distinguish among them. Contrary to what was originally expected, however, we did not find any significant difference between the two groups. This result may indicate that the consistency of sequential eye-movements remains a key mechanism to maintain memory and support memory recognition, even when neuropathological conditions are affecting it. Future studies may examine in greater detail the featural information that is stored when sequences of eye-movements are generated, perhaps by testing memory of participants directly on the locations of the scene that have been actively viewed, rather than just their overall recognition.

When looking at the entropy of fixation distribution, we found that a larger exploration of the scene, either at encoding or at recognition, was associated with better memory performance, which is largely in line with recent research on the topic (Damiano

& Walther, 2019). In line with our study on young adults (Mikhailova et al., 2021), the semantic interference manipulation exerted a general effect of focusing overt attention during encoding while widening it during recognition, especially for scenes that would be later correctly recognized. As the fidelity of visual representations blurs due to the increasing interference during encoding, participants need to focus on more specific details of the scenes (i.e. lower entropy) to make them more distinctively memorable, and this was especially the case for MCI participants, who may have a reduced ability to accrue visual information from the context (Mosimann et al., 2004). When participants were later asked to recognize a scene encoded in a highly interfering semantic category, they needed to acquire as much information as possible to trigger its correct retrieval (i.e., higher entropy), but this strategy did not prove to be successful in MCI participants. In fact, even though MCI participants needed to explore more widely the scene during recognition than healthy adults to correctly retrieve it from memory, possibly due to their hypothesised reduction in focal attention (Rösler et al., 2005), they did not compensate in the face of increasing semantic interference. An interesting follow-up germinating from this work could systematically compare the amount of information that is gained from the periphery by using gaze-contingency methods to blur extra-foveal regions of the visual field at different eccentricities. This paradigm may reveal a cut-off point between MCI and healthy-control participants at which they reach equal recognition performance.

Finally, we explored how much and to what extent MCIs and healthy adults attended to low-level visual features of scenes, and whether this would relate to image memorability and semantic interference. Contrary to Isola et al., (2011), who found no effect of visual saliency, but in line with our sibling study on younger adults (Mikhailova et al., 2021), we found that attending to scene locations which were more visually salient, in either phases, was linked to poorer recognition. Interestingly, we did not find any significant difference between the two groups, which departs from previous literature showing instead a weaker

reliance of people with dementia to low-level visual features of stimuli (Cronin-Golomb et al., 2007). We acknowledge that the results presented in this study are based on the natural visual saliency of scenes, and perhaps, a more controlled study where visual saliency is systematically manipulated is needed to provide more conclusive results.

Our results lend support to theories advocating a unitary approach to the study of memory and overt attention, whereby the explicit responses of the latter may inform about the ongoing processes of the former (Ryan et al., 2020; Ryan & Shen, 2020). We found a few significant differences in the eye-movement of the MCI participants and healthy adults, confirming that it may prove important to add this component to clinical research (Molitor et al., 2015), though such differences were more subtle than expected.

In sum, this study shows that the long-term capacity to retain naturalistic information in older populations is less efficient than previously assumed to be, and recognition is strongly susceptible to semantic interference, especially in individuals not affected by neuropathological conditions. People with MCI seem to utilise the oculo-motor system similarly to healthy adults; but also display subtle differences, which may partly help explain the reduced impact that semantic interference had on their long-term memory.

Acknowledgments

We would like to thank Prof. Alan Baddeley and Dr. Timothy Brady, who also hosted a short visit of MIC to his lab, for the insightful comments and feedback to previous versions of this work. We are also grateful to Dr. Melanie Vitkovitch for thoroughly reading and commenting on the revised version of this manuscript and Prof. Frank Keller for giving the EyeTribe to MIC as a gift while he was a Research Associate in his lab. This research was supported by Fundação para a Ciência e Tecnologia under Grant [PTDC/PSI-ESP/30958/2017] to MIC.

Appendix A

Correlational analyses of neuropsychological tests and recognition memory measures in the MCI group.

In this section, we briefly discuss the most important correlations observed between the performance at the neuropsychological tests, the recognition accuracy and interference slopes of our MCI participants. In Figure 1A, we report the Pearson's correlations of all measures, standardised and centred, leaving in only those values that were significant at $p < 0.05$, and blanking anything else that was below the stated level of significance. The findings that we highlight are: (a) MMSE positively correlates with recognition accuracy and on tests tapping into memory processes, especially when involving long-term mechanisms (e.g., REY, Delayed); and (b) semantic interference is negatively correlated with (1) the MMSE, indicating that cognitive capacity matters for this mechanism to manifest, (2) the REY (Delayed), showing that indeed it semantic interference is a mechanism relying on long-term processes, and (3) the 5 Words test, which is a task tapping maintenance and retrieval of semantic information.

INSERT FIGURE 1A ABOUT HERE

References

Adlam, A. L. R., Bozeat, S., Arnold, R., Watson, P., & Hodges, J. R. (2006). Semantic knowledge in mild cognitive impairment and mild Alzheimer's disease. *Cortex*, *42*(5), 675–684. https://doi.org/10.1016/S0010-9452(08)70404-0

Arnáiz, E., Almkvist, O., Ivnik, R. J., Tangalos, E. G., Wahlund, L. O., Winblad, B., & Petersen, R. C. (2004). Mild cognitive impairment: A cross-national comparison. *Journal of Neurology, Neurosurgery and Psychiatry*, *75*(9), 1275–1280. https://doi.org/10.1136/jnnp.2003.015032

Baddeley, A. D., & Dale, H. C. A. (1966). The effect of semantic similarity on retroactive interference in long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *5*(5), 417–420. https://doi.org/10.1016/S0022-5371(66)80054-3

Bainbridge, W. A., Hall, E. H., & Baker, C. I. (2019). Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-018-07830-6

Barbeau, E. J., Didic, M., Joubert, S., Guedj, E., Koric, L., Felician, O., Ranjeva, J. P., Cozzone, P., & Ceccaldi, M. (2012). Extent and neural basis of semantic memory impairment in mild cognitive impairment. *Journal of Alzheimer's Disease*, *28*(4), 823–837. https://doi.org/10.3233/JAD-2011-110989

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bartolo, A., Della Sala, S., & Cubelli, R. (2016). The sign of "undue contact" in the Object Use Test. *Cortex*, *75*, 235–236. https://doi.org/10.1016/j.cortex.2015.06.016

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boucart, M., Bubbico, G., Szaffarczyk, S., & Pasquier, F. (2014). Animal spotting in Alzheimer's

disease: An eye tracking study of object categorization. *Journal of Alzheimer's Disease*, *39*(1), 181–189. https://doi.org/10.3233/JAD-131331

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 4. https://doi.org/10.1167/11.5.4

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Bylinskii, Z., Isola, P., Bainbridge, C., Torralba, A., & Oliva, A. (2015). Intrinsic and extrinsic effects on image memorability. *Vision Research*, *116*, 165–178. https://doi.org/10.1016/j.visres.2015.03.005

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(6), 1–15. https://doi.org/10.1167/9.3.6.Introduction

Coco, M. I., & Keller, F. (2012). Scan Patterns Predict Sentence Production in the Cross-Modal Processing of Visual Scenes. *Cognitive Science*, *36*(7), 1204–1223. https://doi.org/10.1111/j.1551-6709.2012.01246.x

Coco, M. I., & Keller, F. (2014). Classification of visual and linguistic tasks using eye-movement features. *Journal of Vision*, *14*(3), 11. https://doi.org/10.1167/14.3.11

Craig, M., Dewar, M., & Sala, S. Della. (2015). Retroactive Interference. In *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* (Second Edi, Vol. 20). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.51042-3

Cronin-Golomb, A., Gilmore, G. C., Neargarder, S., Morrison, S. R., & Laudate, T. M. (2007). Enhanced stimulus strength improves visual cognition in aging and Alzheimer's disease. *Cortex*, *43*(7), 952–966. https://doi.org/10.1016/S0010-9452(08)70693-2

Dalmaijer, E. S., Mathôt, S., & Van der Stigchel, S. (2014). PyGaze: an open-source, cross-platform

toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, *46*(4), 913–921. https://doi.org/10.3758/s13428-013-0422-2

Damiano, C., & Walther, D. B. (2019). Distinct roles of eye movements during memory encoding and retrieval. *Cognition*, *184*(December 2018), 119–129. https://doi.org/10.1016/j.cognition.2018.12.014

Dewar, M., Alber, J., Butler, C., Cowan, N., & Della Sala, S. (2012). Brief Wakeful Resting Boosts New Memories Over the Long Term. *Psychological Science*, *23*(9), 955–960. https://doi.org/10.1177/0956797612441220

Dewar, M. T., Cowan, N., & Sala, S. Della. (2007). Forgetting due to retroactive interference: A fusion of Müller and Pilzecker's (1990) early insights into everyday forgetting and recent research on anterograde amnesia. *Cortex*, *43*(5), 616–634. https://doi.org/10.1016/S0010-9452(08)70492-1

Dubois, B., Touchon, J., Portet, F., Ousset, P. J., Vellas, B., & Michel, B. (2002). " The 5 words": a simple and sensitive test for the diagnosis of Alzheimer's disease. *Presse Medicale (Paris, France: 1983)*, *31*(36), 1696–1699.

Duong, A., Whitehead, V., Hanratty, K., & Chertkow, H. (2006). The nature of lexico-semantic processing deficits in mild cognitive impairment. *Neuropsychologia*, *44*(10), 1928–1935. https://doi.org/10.1016/j.neuropsychologia.2006.01.034

Evans, K. K., & Baddeley, A. (2018). Intention, attention and long-term memory for visual scenes: It all depends on the scenes. *Cognition*, *180*, 24–37. https://doi.org/10.1016/j.cognition.2018.06.022

Fletcher, W. A., & Sharpe, J. A. (1986). Saccadic eye movement dysfunction in Alzheimer's disease. *Annals of Neurology*, *20*(4), 464–471. https://doi.org/10.1002/ana.410200405

Foulsham, T., & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *Journal of Experimental Psychology: General*, *142*(1), 41–56. https://doi.org/10.1037/a0028227

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., Belleville, S., Brodaty, H., Bennet, D., Chertkow, H., Cummings, J. L., Leon, M. De, Feldman, H., Ganguli, M., Hampel, H., Scheltens, P., Tierney, M. C., Whitehouse, P., & Winblad, B. (2006). Mild cognitive impairment. *Lancet*, *367*, 1262–1270.

Girtler, N., Brugnolo, A., Campus, C., De Carli, F., Fama, F., Ghio, L., Robert, P., Barbieri, M. P., Delmonte, M., Venuti, S., Gabrielli, F., Rodriguez, G., & Nobili, F. (2012). The Short Cognitive Evaluation Battery in Cognitive Disorders of the Elderly - Italian Version. *Dementia and Geriatric Cognitive Disorders*, *29*(33), 255–265. https://doi.org/10.1159/000339230

Goetschalckx, L., Moors, P., & Wagemans, J. (2018). Image memorability across longer time intervals. *Memory*, *26*(5), 581–588. https://doi.org/10.1080/09658211.2017.1383435

Grigoletto, F., Zappalà, G., Anderson, D. W., & Lebowitz, B. D. (1999). Norms for the mini-mental state examination in a healthy population. *Neurology*, *53*(2), 315–320. https://doi.org/10.1212/wnl.53.2.315

Gusfield, D. (1997). Algorithms on strings, trees, and sequences: Computer science and computational biology. *ACM Sigact News*, *28*(4), 41–60.

Hannula, D. E. (2018). Attention and long-term memory: Bidirectional interactions and their effects on behavior. In *Psychology of Learning and Motivation - Advances in Research and Theory* (1st ed., Vol. 69). Elsevier Inc. https://doi.org/10.1016/bs.plm.2018.09.004

Hannula, D. E., Althoff, R. R., Warren, D. E., Riggs, L., Cohen, N. J., & Ryan, J. D. (2010). Worth a glance: using eye movements to investigate the cognitive neuroscience of memory. *Frontiers in Human Neuroscience*, *4*, 166. https://doi.org/10.3389/fnhum.2010.00166

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504. https://doi.org/10.1016/j.tics.2003.09.006

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition*, *100*(3), 464–482. https://doi.org/10.1016/j.cognition.2005.02.006

Isola, P., Parikh, D., Torralba, A., & Oliva, A. (2011). Understanding the Intrinsic Memorability of Images. *Advances in Neural Information Processing Systems 24*. https://doi.org/10.1167/12.9.1082

Isola, P., Xiao, J., Parikh, D., Torralba, A., & Oliva, A. (2014). What makes a photograph memorable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(7), 1469–1482. https://doi.org/10.1109/TPAMI.2013.200

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Karlsson, T., Johansson, I., Adolfsson, R., Nilsson, L.-G., & Dubuc, S. (2003). Recognition Memory in Alzheimer's Disease A Demonstration of a Remarkable Memory Capacity in Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, *15*, 6–9. https://doi.org/10.1159/000066671

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010a). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology. General*, *139*(3), 558–578. https://doi.org/10.1037/a0019165

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010b). Scene memory is more detailed than you think: the role of categories in visual long-term memory. *Psychological Science*, *21*(11), 1551–1556. https://doi.org/10.1177/0956797610385359

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lagun, D., Manzanares, C., Zola, S. M., Buffalo, E. A., & Agichtein, E. (2011). Detecting cognitive

impairment by eye movement analysis using automatic classification algorithms. *Journal of Neuroscience Methods*, *201*(1), 196–203. https://doi.org/10.1016/j.jneumeth.2011.06.027

Leyhe, T., Müller, S., Eschweiler, G. W., & Saur, R. (2010). Deterioration of the memory for historic events in patients with mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia*, *48*(14), 4093–4101. https://doi.org/10.1016/j.neuropsychologia.2010.10.011

Loewenstein, D. A., Acevedo, A., Luis, C., Crum, T., Barker, W. W., & Duara, R. (2004). Semantic interference deficits and the detection of mild Alzheimer's disease and mild cognitive impairment without dementia. *Journal of the International Neuropsychological Society*, *10*(1), 91–100. https://doi.org/10.1017/S1355617704101112

Lüdecke, D. (2019). *esc: Effect Size Computation for Meta Analysis (Version 0.5.1)*. https://doi.org/10.5281/zenodo.1249218

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Mattan S, B.-S., Makowski, D., & Lüdecke, D. (2020). Compute and interpret indices of effect size. In *CRAN*. https://github.com/easystats/effectsize

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McGeoch, J. A., & McDonald, W. T. (1931). Meaningful Relation and Retroactive Inhibition. *The American Journal of Psychology*, *43*(4), 579–588.

Measso, G., Cavarzeran, F., Zappalà, G., Lebowitz, B. D., Crook, T. H., Pirozzolo, F. J., Amaducci, L. A., Massari, D., & Grigoletto, F. (1993). The Mini-Mental State Examination: Normative Study of An Italian Random Sample. *Developmental Neuropsychology*, *9*(2), 77–85. https://doi.org/10.1080/87565649109540545

Mikhailova, A., Raposo, A., Della Sala, S., & Coco, M. I. (2021). Eye-movements reveal semantic interference effects during the encoding of naturalistic scenes in long-term memory. *(Under Review)*.

Mitchell, A. J., & Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia - Meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica*, *119*(4), 252–265. https://doi.org/10.1111/j.1600-0447.2008.01326.x

Molitor, R. J., Ko, P. C., & Ally, B. A. (2015). Eye movements in Alzheimer's disease. In *Journal of Alzheimer's Disease* (Vol. 44, Issue 1, pp. 1–12). https://doi.org/10.3233/JAD-141173

Mosimann, U. P., Felblinger, J., Ballinari, P., Hess, C. W., & Müri, R. M. (2004). Visual exploration behaviour during clock reading in Alzheimer's disease. *Brain*, *127*(2), 431–438. https://doi.org/10.1093/brain/awh051

Mulatti, C., Calia, C., De Caro, M. F., & Della Sala, S. (2014). The cumulative semantic interference effect in normal and pathological ageing. *Neuropsychologia*, *65*, 125–130. https://doi.org/10.1016/j.neuropsychologia.2014.10.007

Müller, G. E., & Pilzecker, A. (1900). *Experimentelle beiträge zur lehre vom gedächtniss* (J. Barth (ed.)).

Nuthmann, A., & Einhäuser, W. (2015). A new approach to modeling the influence of image features on fixation selection in scenes. *Annals of the New York Academy of Sciences*, *1339*(1), 82–96. https://doi.org/10.1111/nyas.12705

Ooms, K., & Krassanakis, V. (2018). Measuring the spatial noise of a low-cost eye tracker to enhance fixation detection. *Journal of Imaging*, *4*(8). https://doi.org/10.3390/jimaging4080096

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*(1), 107–123. https://doi.org/10.1016/S0042-6989(01)00250-4

Peters, R. J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Research*, *45*(18), 2397–2416.

https://doi.org/10.1016/j.visres.2005.03.019

Petersen, R. C. (2016). Mild Cognitive Impairment. *CONTINUUM: Lifelong Learning in Neurology*, *22*(2 Dementia), 404–418. https://doi.org/10.1007/s12445-012-0228-y

Reed, S. K. (2012). *Cognition: Theories and applications.* CENGAGE learning.

Rösler, A., Mapstone, M., Hays-Wicklund, A., Gitelman, D. R., & Weintraub, S. (2005). The "zoom lens" of focal attention in visual search: Changes in aging and Alzheimer's disease. *Cortex*, *41*(4), 512–519. https://doi.org/10.1016/S0010-9452(08)70191-6

Rust, N. C., & Mehrpour, V. (2020). Understanding Image Memorability. *Trends in Cognitive Sciences*, *24*(7), 557–568. https://doi.org/10.1016/j.tics.2020.04.001

Ryan, J. D., & Shen, K. (2020). The eyes are a window into memory. *Current Opinion in Behavioral Sciences*, *32*, 1–6. https://doi.org/10.1016/j.cobeha.2019.12.014

Ryan, J. D., Shen, K., & Liu, Z. (2020). The intersection between the oculomotor and hippocampal memory systems : empirical developments and clinical implications. *Annals of the New York Academy of Sciences*, *1464*(1), 115–141. https://doi.org/10.1111/nyas.14256

Standing, L. (1973). Learning 10,000 PICTURES. *Quarterly Journal of Experimental Psychology*, *25*(25), 207–222. https://doi.org/10.1080/14640747308400340

Tales, A., Muir, J., Jones, R., Bayer, A., & Snowden, R. J. (2004). The effects of saliency and task difficulty on visual search performance in ageing and Alzheimer's disease. *Neuropsychologia*, *42*(3), 335–345. https://doi.org/10.1016/j.neuropsychologia.2003.08.002

Tavakoli, H. R., Rahtu, E., & Heikkil, J. (2011). Fast and efficient saliency detection using sparse sampling and kernel density estimation. In *Scandinavian conference on image analysis* (pp. 666–675). Springer Berlin Heidelberg.

Thompson, S. A., Graham, K. S., Patterson, K., Sahakian, B. J., & Hodges, J. R. (2002). Is knowledge of famous people disproportionately impaired in patients with early and questionable Alzheimer's disease? *Neuropsychology*, *16*(3), 344–358. https://doi.org/10.1037/0894-4105.16.3.344

Wynn, J. S., Olsen, R. K., Binns, M. A., Buchsbaum, B. R., & Ryan, J. D. (2018). Fixation reinstatement supports visuospatial memory in older adults. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(7), 1119. https://doi.org/10.1037/xhp0000522

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). SUN database: Large-scale scene recognition from abbey to zoo. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3485–3492. https://doi.org/10.1109/CVPR.2010.5539970

Yang, Q., Wang, T., Su, N., Xiao, S., & Kapoula, Z. (2013). Specific saccade deficits in patients with Alzheimer's disease at mild to moderate stage and in patients with amnestic mild cognitive impairment. *Age*, *35*(4), 1287–1298. https://doi.org/10.1007/s11357-012-9420-z

Table 1: Neuropsychological battery of tests administered at the end of the follow-up session: Mini Mental State Examination (MMSE); Rey's Complex Figure (REY); Digit span (DIGIT); Corsi's Block Tapping Test (CORSI); Mesulam's Cancellation test; Dubois' 5 Words test. Means and standard deviation (in parentheses) for the MCI participants and healthy aged-matched controls. The p-value is obtained by comparing the two groups using a Welch two sample t-test, in bold when the difference was significant.

Neuropsychological Test		MCI (N = 27)	Control (N = 23)	p-value
MMSE		**24.58 (3.45)**	**28.74 (1.66)**	**< .001**
REY	**Copy**	**26.57 (7.26)**	**32.3 (3.2)**	**< .001**
	Delayed	**6.28 (6.27)**	**11.07 (4.41)**	**0.01**
DIGIT	**Forwards**	**4.12 (0.83)**	**4.78 (1.04)**	**0.02**
	Backwards	**2.68 (0.69)**	**3.39 (0.78)**	**< .001**
CORSI	**Forwards**	**3.88 (0.93)**	**4.57 (0.79)**	**0.01**
	Backwards	**3 (0.87)**	**3.61 (0.94)**	**0.02**
Cancellation Test	Letter	57.04 (4.81)	58.61 (1.97)	0.14
	Symbol	55.4 (4.04)	56.22 (4.77)	0.53
5 Words test		**8.12 (1.92)**	**10 (0)**	**< .001**

Table 2: Generalized and linear mixed-effects model outputs for the *recognition accuracy*

Dependent Variable	Predictor	B	SE	β	d	z
Response Accuracy	Intercept	1.13	0.08	1.13	4.24	14.02***
	Group	-0.66	0.15	-0.32	0.69	-4.38***
	SI	-0.27	0.03	-0.27	0.56	-9.44***
	Group:SI	0.21	0.04	0.1	0.21	5.13***

(*) $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note. Predictors (centered) were: Group (Control = -.5, MCI = .5), and Semantic Interference (SI, a z-scored continuous variable ranging from 2 to 48). The random effects introduced as intercept and slopes were Participants (50) and the semantic Category of the scene (141).

Table 3: Linear mixed-effects model outputs for the *d-prime* and the *recognition difference* between the recognition accuracy of participants without semantic interference (follow-up session) and with it (main session, 6, 12, 24).

Dependent variable	Predictor	β	SE	β	d	t
D-prime	Intercept	1	0.15	0.12	0.24	6.82***
	Group	-0.68	0.16	-0.88	3.83	-4.3***
	SI (6)	0	0.18	0.01	0.02	0.05
	SI (12)	-0.3	0.18	-0.39	0.84	-1.66(*)
	SI (24)	-0.42	0.18	-0.54	1.3	-2.32
	Performance	1.08	0.16	1.4	2.85	-2.32*
	Group:SI (6)	-0.08	0.19	-0.11	0.22	-0.43
	Group:SI (12)	0.3	0.19	0.4	0.84	1.57
	Group:SI (24)	0.39	0.19	0.5	1.18	2.01*
	SI (6):Performance	-0.18	0.19	-0.23	0.48	0.35
	SI (12):Performance	-0.59	0.19	-0.76	2.4	-3.03**
	SI (24):Performance	-0.78	0.19	-1.02	10.27	-4.02***
Recognition Difference	Intercept	-5.4	2.7	0.21	0.44	-2*
	Group	11.36	3.33	0.9	4.26	3.4**
	SI (12)	-3.38	1.69	-0.27	0.55	-1.99*
	SI (24)	-7.25	1.69	-0.57	1.41	-4.27***
	Performance	-2.43	3.33	-0.19	0.39	-0.73
	Group:Performance	-14.34	4.52	-1.14	4.1	-3.17**

Note. Predictors (centered) in the linear-mixed effect models were: Group (Control, MCI; Control as reference level), Performance (Low, High; Low as reference level), and Semantic Interference (1, 6, 12, 24; 1 as reference level). The only random effect introduced as intercept was Participants (50).

Table 4: Linear mixed-effects model outputs for the *fixation entropy* observed during encoding and recognition of the seen (old) image.

Fixation Entropy	Predictor	B	SE	β	d	t
Encoding	Intercept	10.93	0.05	0	0	221.93***
	Accuracy	0.05	0.01	0.05	0.1	4.14***
	Group	-0.05	0.1	-0.05	0.1	-0.5
	SI	-0.01	0.01	-0.02	0.04	-1.64(*)
	Accuracy:SI	-0.02	0.01	-0.02	0.04	-2.31*
	Group:SI	-0.03	0.01	-0.04	0.08	-4.51***
Recognition (old image)	Intercept	10.7	0.03	-0.01	0.03	312.33***
	Accuracy	0.15	0.02	0.15	0.31	6.9***
	Group	0.08	0.06	0.1	0.19	1.25
	SI	0.02	0	0.04	0.08	2.65*
	Accuracy:Group	0.07	0.04	0.04	0.04	1.68(*)
	Accuracy:SI	0.04	0.01	0.04	0.08	3.86**
	Group:SI	0.01	0.01	0.01	0.03	1.28
	Accuracy:Group:SI	-0.05	0.02	-0.26	0.05	-2.88**

Note. Predictors (centered) in the linear-mixed effect models were: Accuracy (Incorrect = -0.73; Correct = 0.26), Group (Control = -.5, MCI = .5), and Semantic Interference (SI, a z-scored continuous variable originally ranging from 2 to 48). The random effects introduced as intercept and slopes were Participants (43) and the semantic Category of the scene (141).

Table 5: Linear mixed-effects model outputs for the eye-movement measures of *scan-pattern similarity* between the sequence of fixations on the seen (old) image during encoding and the sequence of fixations on the same seen image during the recognition phase; and the *normalized scan-path* correspondence between fixation position and low-level visual saliency value observed during encoding and recognition of the seen (old) image.

Dependent Variable	Predictor	B	SE	β	d	t
Scan-Pattern Similarity	Intercept	0.27	0.008	-0.3	0.06	34.55^{***}
	Accuracy	0.02	0.003	0.07	0.13	6.41^{***}
	SI	0.006	0.002	0.05	0.1	3.4^{***}
NSS (Encoding)	Intercept	0.57	0.01	-0.04	0.09	37.97^{***}
	Accuracy	-0.02	0.006	-0.03	0.05	-2.55^{*}
	SI	0.01	0.004	0.04	0.08	2.29^{*}
NSS (Recognition)	Intercept	0.5	0.01	-0.06	0.12	42.16^{***}
	Accuracy	-0.02	0.007	-0.03	0.06	-3.18^{**}

Note. Predictors (centered) entered in the linear-mixed effect models were: Accuracy (Incorrect = -0.73; Correct = 0.26), Group (Control = -.5, MCI = .5), and Semantic Interference (a z-scored continuous variable originally ranging from 2 to 48). The random effects introduced as intercept and slopes were Participants (43) and the semantic Category of the scene (141).

Figure 1:

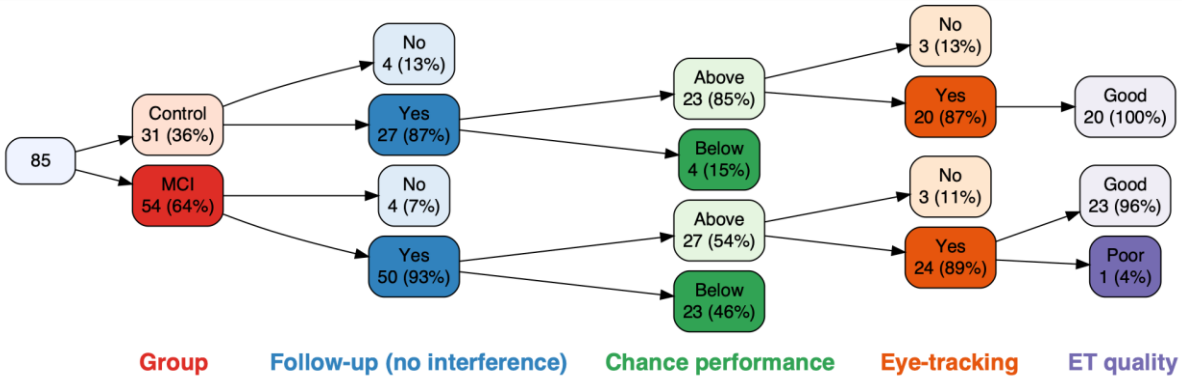

Figure 1: Selection tree illustrating the different steps (and reasons) of participants' exclusion, organized as columns.

Figure 2:

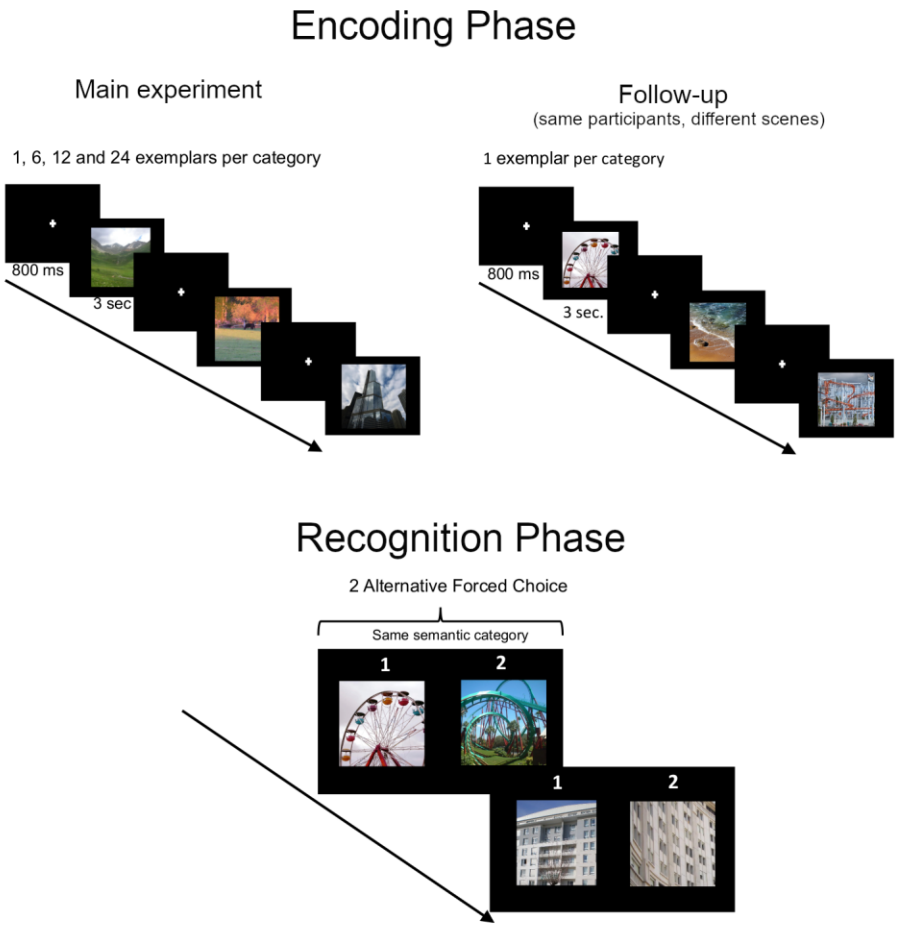

Figure 2: Illustration of the long-term visual memory task. During the encoding phase (upper row) participants were asked to watch a stream of naturalistic scenes, each from a different semantic category, and belonging to any of the four semantic interference blocks (1, 6, 12, 24). The only difference between the main experiment and the follow-up session is that in the latter there was no semantic interference (i.e., a semantic interference level of 1). In the recognition phase, participants verbally indicated the scene they remembered by speaking out loud either “one” (for the left scene) or “two” (for the second scene). We used a different set of scenes for the main experiment and for the follow-up session.

Figure 3:

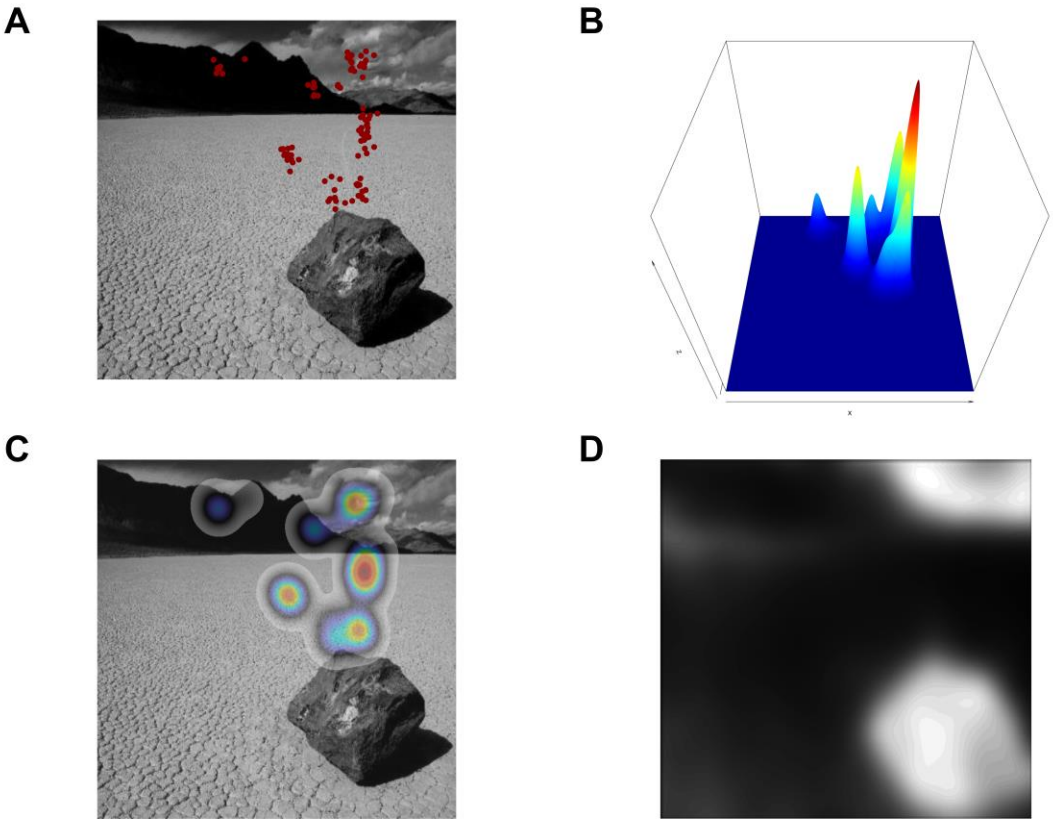

Figure 3: Illustration of the dependent measures considered in this study exemplified from a single trial. (A) A naturalistic scene used in our study with overlaid fixations as red dots. (B) Fixation distribution obtained by fitting Gaussians at fixation locations and normalized to be a probability map. Entropy can be computed over this map and used to represent the overall spread of overt attention across the scene (C) Another way of portraying the fixation distribution using a heatmap. (D) The low-level visual saliency map of the scene as computed using Fast and Efficient Saliency model (FES, Tavakoli, Rahtu and Heikkil, 2011).

Figure 4:

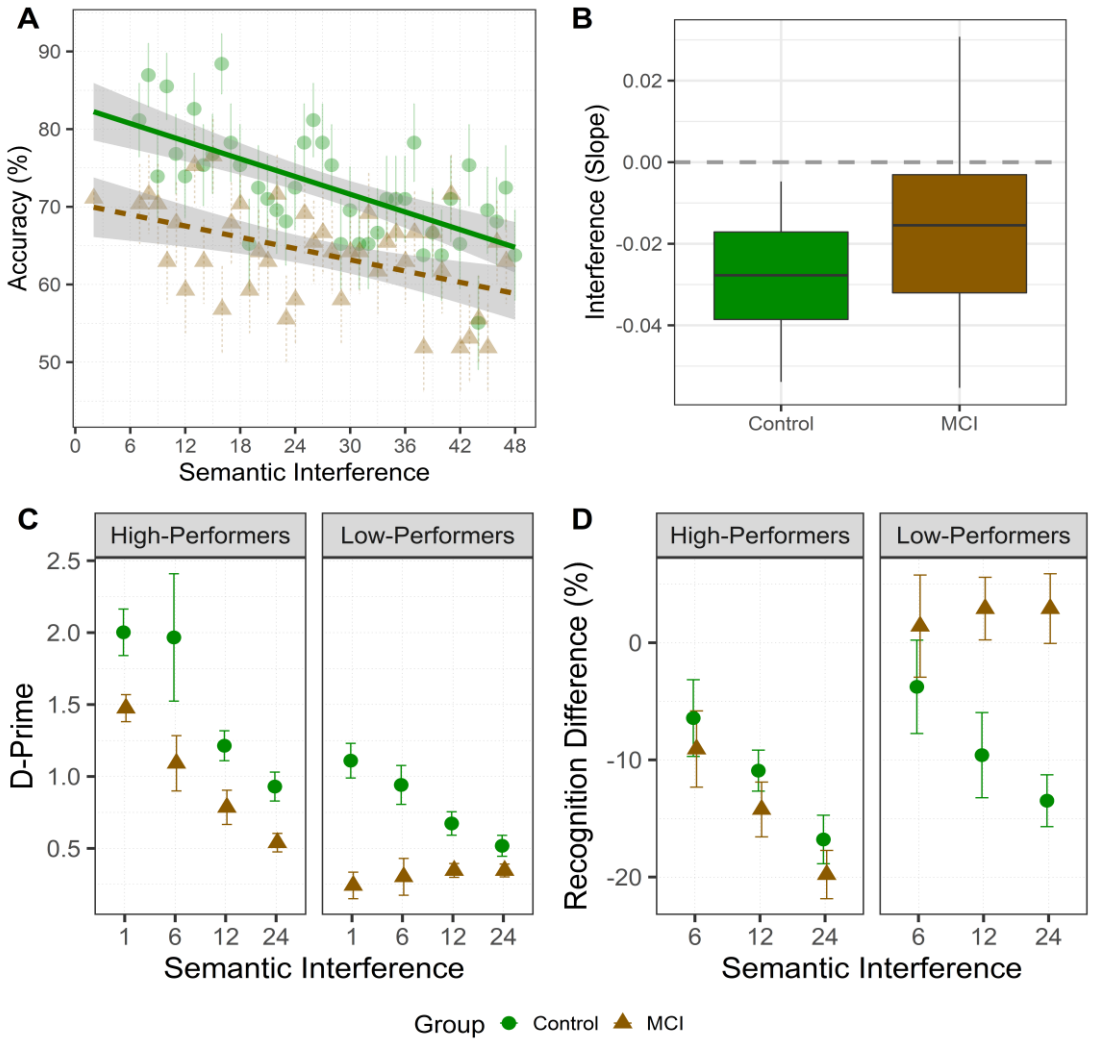

Figure 4: (A) Recognition accuracy (y-axis) as a function of semantic interference (x-axis). Each point represents the average recognition accuracy for each level of interference, represented as a continuous variable ranging from 2 to 48. Lines indicate the estimates from a linear model fit to the data and the shaded bands represent the 95% confidence intervals. (B) Beta coefficient of the semantic interference slope, by-participant, as linear predictor of recognition accuracy, and averaged for the two groups (MCI, Control). The whiskers represent the 25th and 75th percentile of the measure (lower and upper quartiles). (C) D-prime (y-axis) as a function of different levels of semantic interference (1, 6, 12, 24; x-axis). (D) Percentage of difference in recognition accuracy (y-axis) for different levels of semantic interference (6, 12, 24) compared to the no-interference condition (i.e., 1 from the follow-up session). In (A), (B) and (D) the Groups are marked using point type and color (Control = green circles; MCI = brown triangles). In (C) and (D) the recognition Performance of the participants within their group (Low, High) is organized as panels.

Figure 5:

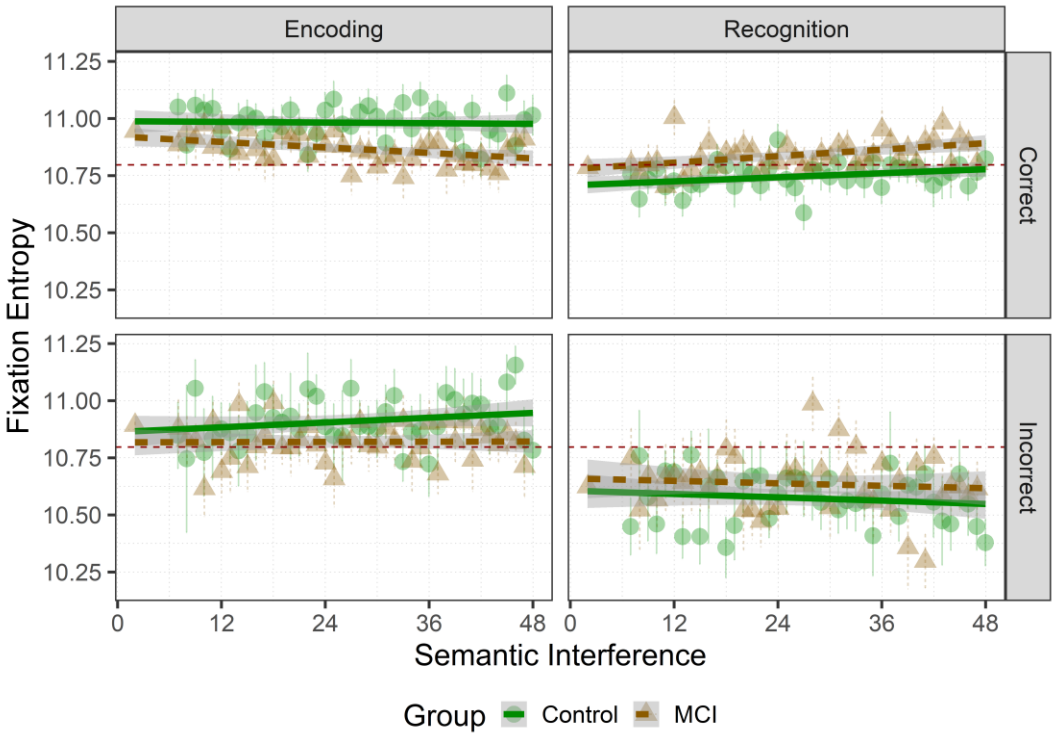

Figure 5: Entropy of the fixation distribution (y-axis) as a function of semantic interference (x-axis: a continuous variable ranging from 2 to 48) during encoding and recognition of the old (seen) image organized as columns for Correct (top-panel) and Incorrect (bottom panel) recognition responses. The Group of participants is marked using point type and color (Control = green circles; MCI = brown triangles), and each point represents the average of fixation entropy for each level of interference. Lines indicate the estimates from a linear model fit to the data and the shaded bands represent the 95% confidence intervals. The dark red dashed mid-line present in all the panels represent the average fixation entropy and it serves to more easily compare relative differences across the four panels.

Figure 6:

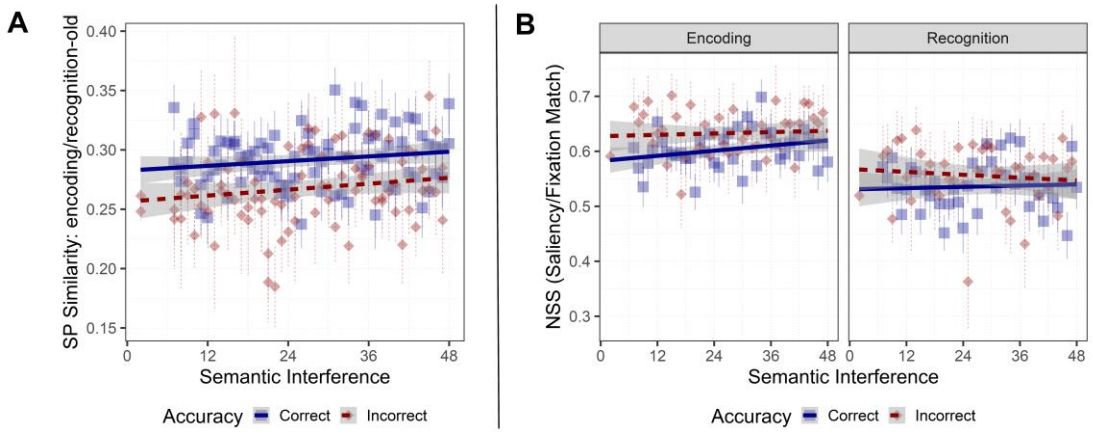

Figure 6: (A) Similarity between eye-movement scan-patterns during the encoding and recognition of an old (seen) image as a function of semantic interference. The similarity is computed using the Longest Common Subsequence (range from 0 to 1). (B) The Normalized Scan-Path Saliency between fixation position and low-level visual saliency of the scene for the two phases of the memory task. In both (A) and (B), recognition accuracy is visualized using point type and color (Correct = blue squares; Incorrect = red diamonds). Each point represents the average scan-pattern similarity or NSS for each level of interference, represented as a continuous variable ranging from 2 to 48 for correct vs incorrect recognition memory responses. Lines indicate the estimates from a linear model fit to the data and the shaded bands represent the 95% confidence intervals.

Figure 1A:

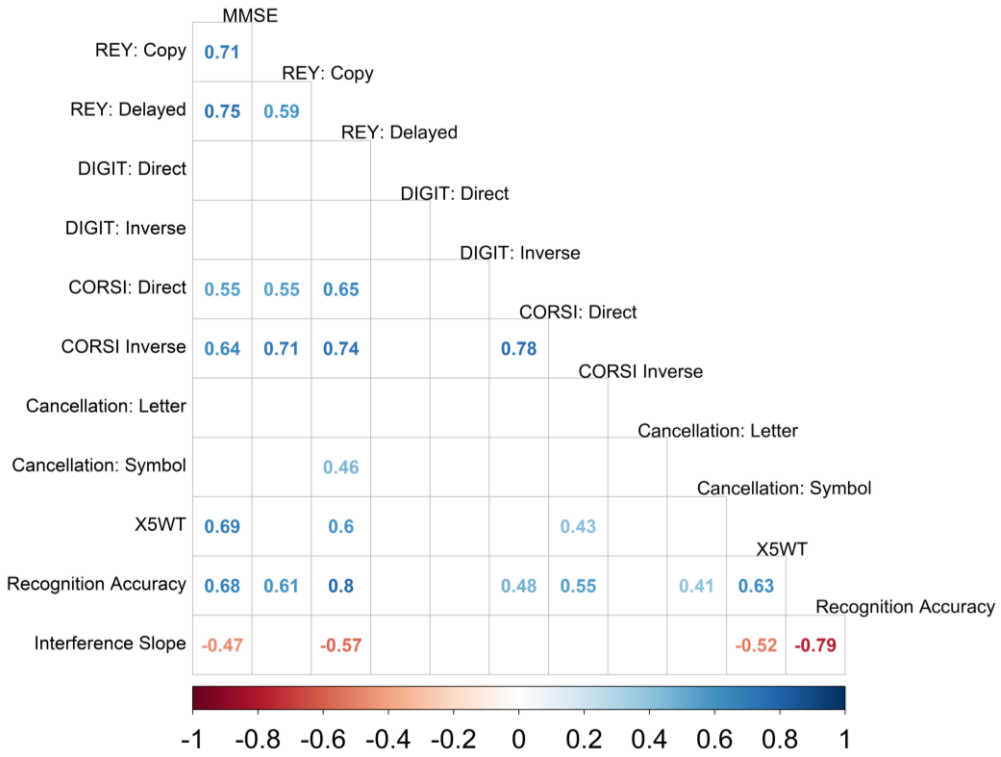

Figure 1A: Correlation of neuropsychological tests, the recognition accuracy in the long-term memory task, and the individual interference slope of the MCI participants. All measures have been standardised and centred before running the correlation. The correlation values presented are only those ones that are significant at $p < 0.05$, while the empty cells indicate all those correlations that were non-significant. The color palette indicates the strength of the correlation from red (−1) to blue (1).