VISUAL COGNITION
2023, VOL. 31, NO. 1, 1–17
https://doi.org/10.1080/13506285.2023.2188335

How scene encoding affects memory discrimination: Analysing eye movements data using data driven methods

F. Doidy[a], P. Desaunay[a,b], C. Rebillard[a], P. Clochon[a], A. Lambrechts[c], P. Wantzen[a], F. Guénolé[a,b], J. M. Baleyte[a,d], F. Eustache[a], D. M. Bowler[c], K. Lebreton[a] and B. Guillery-Girard[a]

[a]Normandie Université, UNICAEN, PSL Université Paris, EPHE, INSERM, U1077, CHU de Caen, GIP Cyceron, Neuropsychologie et Imagerie de la Mémoire Humaine, Caen, France; [b]Service de Psychiatrie de l'enfant et de l'adolescent, CHU de Caen, Caen, France; [c]Autism Research Group, Department of Psychology, City, University of London, London, UK; [d]Service de Psychiatrie de l'enfant et de l'adolescent, Centre Hospitalier Interuniversitaire de Créteil, Créteil, France

ABSTRACT
Encoding of visual scenes remains under-explored due to methodological limitations. In this study, we evaluated the relationship between memory accuracy for visual scenes and eye movements at encoding. First, we used data-driven methods, a fixation density map (using iMap4) and a saliency map (using GBVS), to analyse the visual attention for items. Second, and in a more novel way, we conducted scanpath analyses without a priori (using ScanMatch). Scene memory accuracy was assessed by asking participants to discriminate identical scenes (targets) among rearranged scenes sharing some items with targets (distractors) and new scenes. Shorter fixation duration in regions of interest (ROIs) at encoding was associated with a better rejection of distractors; there was no significant difference in the relative fixation time in ROIs at encoding, between subsequent hits and misses at test. Hence, density of eye fixations in data-driven ROIs seems to be a marker of subsequent memory discrimination and pattern separation. Interestingly, we also identified a negative correlation between average MultiDimensional Scaling (MDS) distance scanpaths and the correct rejection of distractors, indicating that scanpath consistency significantly affects the ability to discriminate distractors from targets. These data suggest that visual exploration at encoding participates in discrimination processes at test.

ARTICLE HISTORY
Received 19 November 2021
Accepted 2 March 2023

KEYWORDS
Eye tracking; fixation; scanpath; saliency; memory

The interest in using eye tracking to study memory processes is growing. The majority of studies analysed the link between encoding and the subsequent recognition without going further into the process of discrimination (Schurgin, 2018). Research on visual scenes adds a layer of difficulty, because of their complexity and diversity. Scenes are consensually defined as a human-scaled view of an environment, associating items arranged in specific locations with larger surfaces and elements that constitute the background, which together form a coherent semantic concept (see Williams & Castelhano, 2019, for a review). During visual exploration, eye movements bring the image of a stimulus to the fovea, and each eye movement to a position is preceded by a shift of attention to that position (Hoffman, 1998). Eye-tracking methodology provides the opportunity to capture these eye movements and therefore to explore overt attention for visual stimuli such as scenes, focusing on fixations, defined as the stationary states of the eyes during which gaze is held upon a specific location in the image, and saccades, that correspond to rapid eye movements between fixations, but provides only limited insights on covert attention (see Vidal et al., 2012, for a review).

Encoding of natural scenes depends on both the visual properties of the scene, i.e., saliency, semantic integrity (i.e., semantically congruent vs. incongruent objects within a scene context), spatial associations, and the task set, i.e., search, memorization, or free viewing (see Castelhano & Krzyś, 2020; and Williams & Castelhano, 2019, for reviews). Eye-tracking studies on visual scene encoding have so far assessed the link between eye movements and subsequent recognition performance (using old/new paradigms) or awareness (using remember/know paradigms). Some studies identified that the number of fixation points made during scene encoding is associated

CONTACT B. Guillery-Girard berengere.guillery-girard@ephe.psl.eu, berengere.guillery@unicaen.fr Unité INSERM – EPHE – Université de Caen – U1077, Pôle des Formations et de Recherche en Santé (PFRS), "Neuropsychologie et Imagerie de la mémoire humaine", 2, rue des Rochambelles, F-14032 Caen Cedex CS 14032

with greater subsequent recognition performance (Choe et al., 2017), or memory awareness (Kafkas & Montaldi, 2012). In addition, the spatial distribution of clustered eye fixations during encoding has been found to predict subsequent remembering awareness of visual scenes relative to knowing, due to the recognition of distinct details (Sharot et al., 2008). With regards to the duration of fixations, shorter fixation durations during face encoding (considering the whole face as the area of interest) have been found to predict subsequent hits over misses (Parag & Vakil, 2018). Eye fixation can also be guided by visual saliency, which encompasses low-level scene features. The *bottom-up saliency map hypothesis* (Itti & Koch, 2001) points to the role of salient low-level features in capturing attention during free-viewing or explicit memory tasks (Underwood & Foulsham, 2006), but not in a search task (Foulsham & Underwood, 2011). Henderson and Hayes (2018) go beyond this model by showing that both meaning and image saliency predicted the distribution of attention during explicit encoding.

To date, no eye-tracking study on scenes encoding has assessed memory discrimination, which differs from memory recognition, on methodological, behavioural, and neural levels. Memory discrimination tasks require participants to recognize previously encoded stimuli among highly similar lures and new stimuli (i.e., old/lure discrimination), and rely on detailed memory representations of the old stimuli, while memory recognition tasks (i.e., old/new discrimination) do not need such detailed representations, and can be achieved at least with gist-based representations of the old stimuli (Loiotile & Courtney, 2015). Memory discrimination and memory recognition tasks explore respectively pattern separation at encoding, and pattern completion at retrieval, which are episodic memory processes involving different hippocampal subfields (Liu et al., 2016). Pattern separation is based on dentate gyrus activity, and allows to assign different neural codes to events with overlapping content, which results in distinct representations stored in memory, and supports subsequent old/lure discrimination (Hainmueller & Bartos, 2020; Rolls, 2016). Pattern completion is based on the activity of the cornus ammonis (CA3), and triggers, from partial cues, the retrieval of the whole representation of the encoded stimuli (Liu et al., 2016). Only one eye-tracking study has been conducted on memory discrimination, by Molitor et al. (2014), in which participants at test had to recognize old items (i.e., pictures of everyday objects) among similar distractors and new different items. Compared to hits, false alarms were associated with fewer fixations to the target at encoding, suggesting that errors were driven by diminished encoding (the *poor encoding hypothesis*).

The methodology of more recent research on eye fixations in relation to recognition rates and awareness has been based on fixation maps, which correspond to duration-weighted fixation density. Damiano and Walther (2019) used a paradigm in which observers were asked either to fixate or to explore scenes. During the study phase, exploration, relative to fixation, led to higher subsequent recognition rates, while at test, exploration, relative to fixation, led to the same hit rates but a higher rejection rate of new scenes. This demonstrates that fixations do not completely reflect visual attention, and what is encoded/recognized. In contrast to these results, Wynn et al. (2020) evaluated the level of similarity between fixation maps at encoding and retrieval (i.e., gaze reinstatement) during a pattern completion paradigm, and found similar gaze reinstatement for hit and false alarm responses at test. Other studies have replicated previous findings on spatial distribution of fixations, finding that a broader exploration during encoding leads to higher recognition performance (Damiano & Walther, 2019), and stronger familiarity awareness (Ramey et al., 2020a). Looking at the inter-observer congruency of fixation maps, Lyu et al. (2019) identified that scene memory is related to the consistency of fixation maps across viewers at encoding, which is itself related to the signal-to-noise fixation ratio between preferentially viewed regions of the scene and other regions.

Research focusing on scanpaths, i.e., the spatial distributions of eye gaze, were initially guided by the *scanpath theory*, which postulates that scanpath similarity between encoding and retrieval is predictive of memory performance (Noton & Stark, 1971). Subsequent studies have shown that scanpaths are idiosyncratic, showing a greater individual consistency between encoding and retrieval, than between different observers viewing the same image (e.g., Foulsham et al., 2012; French et al., 2017). Similarly to research using fixation maps, more recent research on scanpaths tends to identify specific patterns of eye movement associated with memory formation, and distinct roles for eye movements during encoding and retrieval. Analysing scanpath rehearsal during study phase, Meghanathan et al. (2019) described different types of refixations (the repetition in time of fixation sequences), which supported memory-encoding strategies in free viewing. Comparing scanpath for novel and repeated scenes during a search task, Wynn et al. (2016) evidenced a scanpath repetition of initial and final but not middle fixations during repeated scene; early scanpath similarity increased search efficiency by reducing search time at test. These results have therefore challenged the scanpath theory. Arizpe et al. (2019) used faces to propose an alternative model, suggesting that an increasing number of fixations during encoding enables the gradual integration of disparate information into a coherent representation, that can be activated during recognition within a small number of fixations. Scanpath analysis allows a population to be distinguished with typical development from participants with neuropsychiatric disorders (see Armstrong & Olatunji, 2012; and Toh et al., 2011, for reviews). Notably, during scene encoding and recognition, Shakespeare et al. (2015) have evidenced a greater *consistency* of scanpaths between healthy participants compared to participants with neurological conditions, suggesting that healthy participants looked more appropriately at task-relevant regions of the scene.

To summarize, both fixation map and scanpath eye-tracking studies using scenes suggest that eye-movements during encoding are predictive of recognition rates and awareness at test, but so far no study has explored memory discrimination via pattern separation processes. Hence, the present methodological study aimed to determine what patterns of eye movements during scene encoding are related to memory discrimination for visual scenes, focusing on fixations and scanpaths that predict correct recognition and false alarms. We designed an *old/rearranged/new* paradigm suitable to test memory accuracy and pattern separation processes, including an incidental encoding phase using drawings of scenes, followed by a surprise recognition phase in which participants had to recognize “target” scenes (same items and background), among “distractor” scenes (same items but new background) and new scenes (items and background were new). Incidental encoding of scenes fosters memory for visual features rather than for the whole scene, with images containing many features being richly encoded when compared to other kind of images (Evans & Baddeley, 2018). Drawings and photographs of scenes share large similarities; however, drawings allow a better handling of the content including many details and are more attractive compared to photographs (Park et al., 2019). First, we used data-driven methods, a fixation density map (using iMap4), to evaluate the visual attention for items. We proposed to extend the Molitor et al. (2014) results to fixation durations and predicted an association between increased false alarm rate and longer fixation durations on details shared between targets and distractors. Such an association would support the presence of an attentional bias at encoding, that would increase interference during the recognition phase. We also generated a saliency map (using GBVS) to ensure that our data-driven fixation maps were not related to low-level features, mainly saliency. Second, and in a more novel way, we conducted scanpath analyses without semantic a priori expectations in the definition of ROIs. Memory discrimination relies on detailed memory representations, and presumably on extensive exploration of visual scenes at encoding. Hence, without a priori analyses, which are based on a high number of similar ROIs, each containing a variable amount of semantic information, seem more suitable than a priori analyses which focus on specific items, for assessing participants’ exploration of the whole visual scene. We considered that inter-observer scanpath consistency would be a valuable marker of the quality of the memory trace and predicted that greater scanpath consistency at encoding would be associated with higher discrimination at test.

Method

Participants

Forty-nine healthy young adults were recruited for this study (26 males, mean age 21.3 ± 2.41 years, mean level education 13.81 ± 1.86 years). Of these

49 participants, 42 were right-handed (assessed by the checklist of De Agostini & Dellatolas, 1988), and 41 had right ocular dominance. All participants had normal or corrected-to-normal vision. None reported recent use of alcohol or illicit drugs, current or past mental disorder, neurological disorder (including history of head trauma with loss of consciousness, or seizures), or current medical condition. An intellectual impairment was ruled out, using the Wechsler Adult Intelligence Scale-IV (Verbal Comprehension Index = 111.60 ± 17, Perceptual Reasoning Index = 101.62 ± 12.60). Forty-two participants were retained for analysis (1 participant was excluded because he did not look sufficiently at some pictures, i.e., total visit duration <3% of total scene duration for one picture, and <43% for two pictures) and six participants were excluded due to eye tracker dysfunction including calibration. When using Molitor's results as criteria, the required sample size was about 35 participants for Alpha = 0.05 (power goal = 0.90). We decided to increase the number of inclusions to have a minimum of 35 participants for analyses: 42 participants were retained for analyses, we aimed for this large number of participants to compensate for the small number of stimuli.

This study was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki, 2008), and was approved by the local ethics committee before it started (CPP Nord Ouest III, N° ID RCB: 2013-A01800-45); all participants signed for consent.

Material

A total of 43 coloured pictures (1024 × 717 pixels) created by a professional cartoonist were used (Figure 1): seven were used for the training phase before study phase, six were used to avoid primacy and recency effects at study phase, and the remaining 30 were used in the data analyses. These 30 item-background pictures belonged to one of 10 different semantic categories (rural, farm, forest, garden, mountain, snow, car park, swimming pool, beach, and city). Three pictures were created for each semantic category: a "target" picture identical at study and at test phase, a "distractor" picture, and a "new" picture. The "distractor" picture contained the same items in the foreground as the "target" picture combined with a new background from the same semantic category. The foreground items, common to both the target and distractor, were two or three per picture (3 for "forest" and "beach" categories, and 2 for other categories). The "new" picture combined new foreground items and a new background from the same semantic category as the "target."

Procedure

E-Prime software (PST, Pittsburgh, PA) was used to control stimulus presentation and to record participants' responses. Participants were sitting comfortably 79 cm from the screen in a dimly lit room during the whole experiment. Eye positions and gaze durations were measured at 60 Hz with a remote eye tracker (Tobii X120 A, Tobii Technology AB, Stockholm, Sweden). A box with infrared sources and a camera were set below a 22-inch TFT flat-screen monitor (HP, x22LED). Stimuli were presented in full screen at a resolution of 1024 × 768 pixels. The size of the projection screen was 47.75 × 26.92 cm^2, subtended 33.6° × 19.4° in visual angle. Before each recording, a 9-point eye-tracker calibration and validation procedure was performed.

The incidental memory task included a study phase followed by a surprise recognition phase (Figure 2). At study phase, a trial started with a dynamic fixation central white cross presented on a black background for 1280 ms. A picture then appeared on the screen for 5000 ms, followed by a black screen for 500 ms. The 10 "target" pictures were presented in pseudo-random order. Participants were asked to freely and carefully explore each picture for the full duration of the presentation. The test phase was provided after a 20-minute break during which participants performed unrelated tests (i.e., Rey's figure test, and a verbal fluency test, not affecting participants' ability to complete the test phase; no fatigue was observed). The test phase started with four familiarization trials comprising of four pictures (not presented in the study phase), followed by the 30 pictures of interest presented in pseudo-random order (10 "target," 10 "distractor," and 10 "new" pictures). Participants were instructed to identify "target" pictures and reject other kinds of pictures, i.e., "distractor" and "new" pictures. In both cases, participants responded by pressing one of two keys on a computer mouse. There was no time limit to respond. A trial started

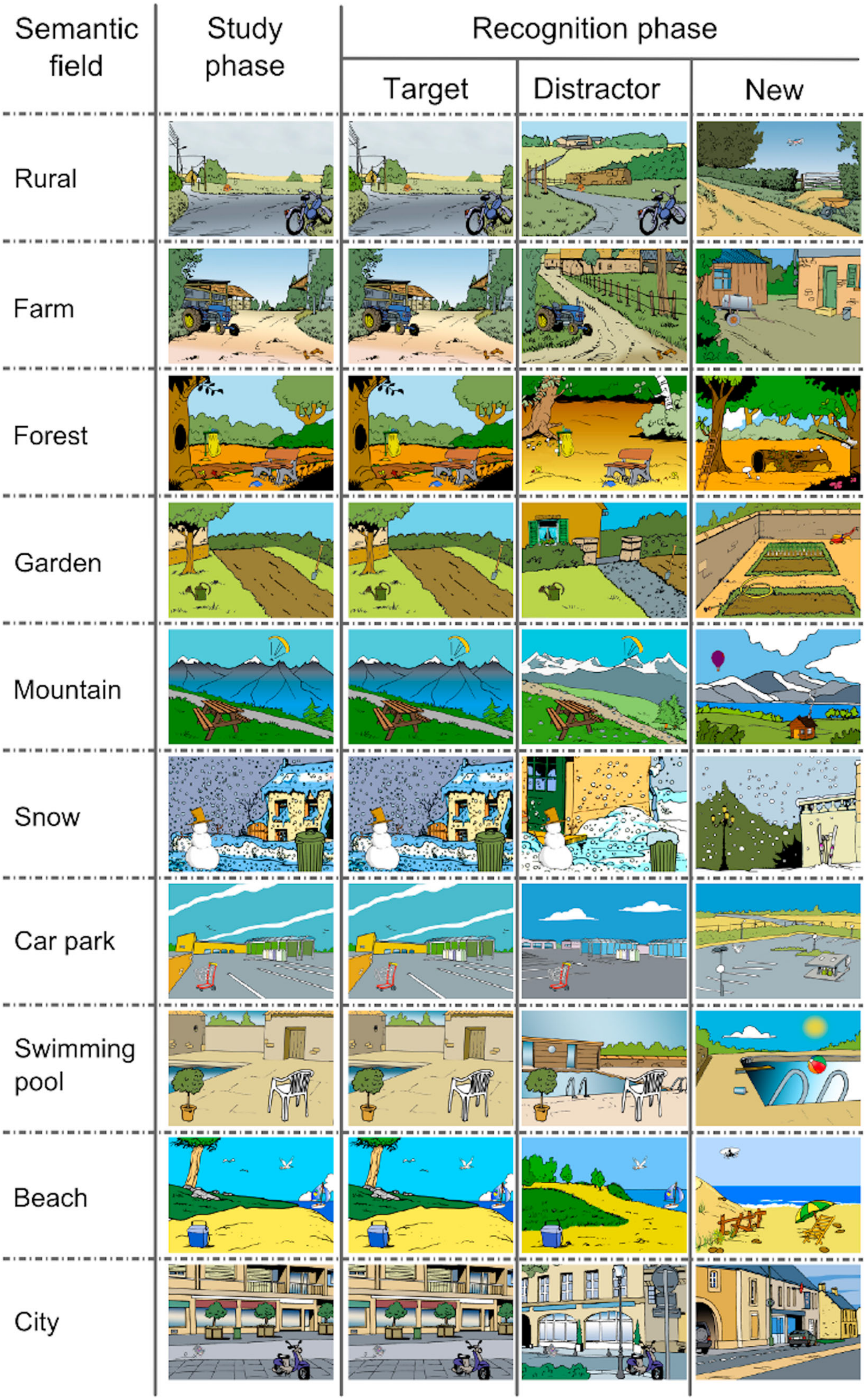

Figure 1. Scenes presented to participants during study and recognition phases (only scenes used for the eye-tracking data analysis are presented).

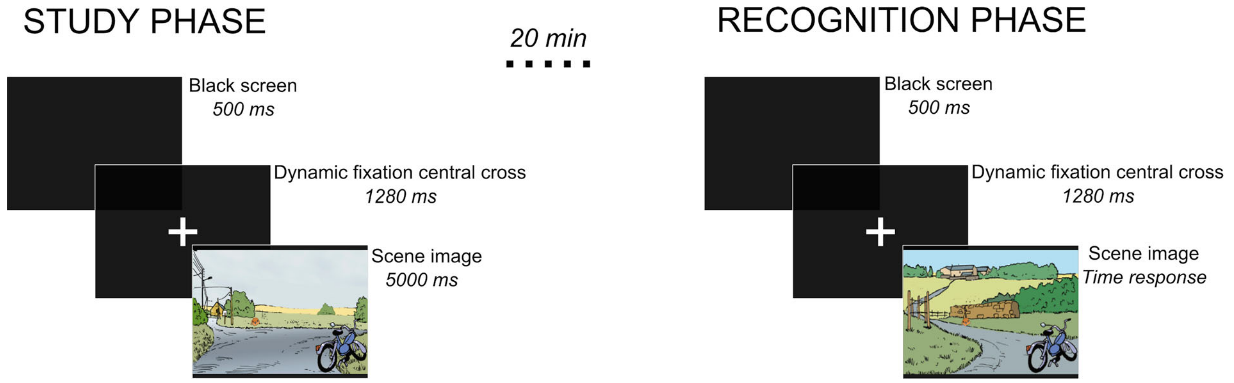

Figure 2. Experimental design. In the study phase (left), participants were presented with pictures and were asked to watch carefully. In the recognition phase (right), participants were presented with pictures again and were asked to decide whether they had seen them in the study phase or not. Pictures in the recognition phase were presented until the participant responded.

by means of a fixation cross, then pictures were presented on the screen until the participant responded, followed by a 500 ms black screen as in the study phase.

Eye-tracking analyses

Eye tracking recording and analyses were carried out separately for each picture, using data from the study phase only (Figure 3).

Data preprocessing

An I-VT filter (Velocity-Threshold Identification filter: classifier: 30°/s; velocity calculator window length: 20 ms) was applied to the eye-tracking data prior to exporting as recommended by the eye tracker manufacturer. The output was based on the average of both eyes. The data were then exported (Identifier, Scene Name, Gaze Event Duration, Fixation Point X, Fixation Point Y) into MATLAB. All data processing and analyses were performed using MATLAB (R2015a) software. The data were preprocessed before statistical analyses.

As the picture sizes were smaller than the screen size, fixations on the screen but outside of the pictures were removed from all analyses. For the same reason, the gaze point coordinates were recalculated to reflect this shift.

For each trial, the first fixation was discarded if it was central, as it was likely to be a reflection (or *remanence*) of the fixation cross presented just before the scene, and therefore was not informative (Hayes & Henderson, 2017; Ramey et al., 2020b; Wang et al., 2015). As a result, 37 first fixations (8.8% of all first fixations) which were not to the picture centre and thus were informative, were kept for analyses. All the remaining fixations were used for the analyses (see Table 1).

Fixation analysis

Gaze fixation location data were analysed using iMap4 (Caldara & Miellet, 2011; Lao et al., 2017; and see also Nicholls et al., 2019, with visual scenes), a freely available MATLAB open source toolbox for the statistical fixation mapping of eye movement data. iMap4 requires a two-step process: first, iMap4 generates fixation duration maps for each participant and every picture, creating individual fixation maps which are smoothed by convoluting them with a two-dimensional Gaussian Kernel function. These individual maps can be averaged to produce a group fixation map (heatmap). Second, these 3D fixation maps (dimensions are x, y, and fixation duration) are compared to produce statistically significant duration-weighted difference maps. The main advantage of this method is that it is data-driven, inspired by methods used in functional Magnetic Resonance Imaging, and no prior segmentation of pictures into ROIs is required.

Toolbox parameters were adapted for the scene stimuli. In order to clearly identify the items embedded in the scenes, the smoothing parameter was set to a 2.5° visual angle. To reduce memory usage and computational time, picture size was rescaled. We used the default scaler parameter of 0.24 (updated x size = 246 pixels, updated y size =

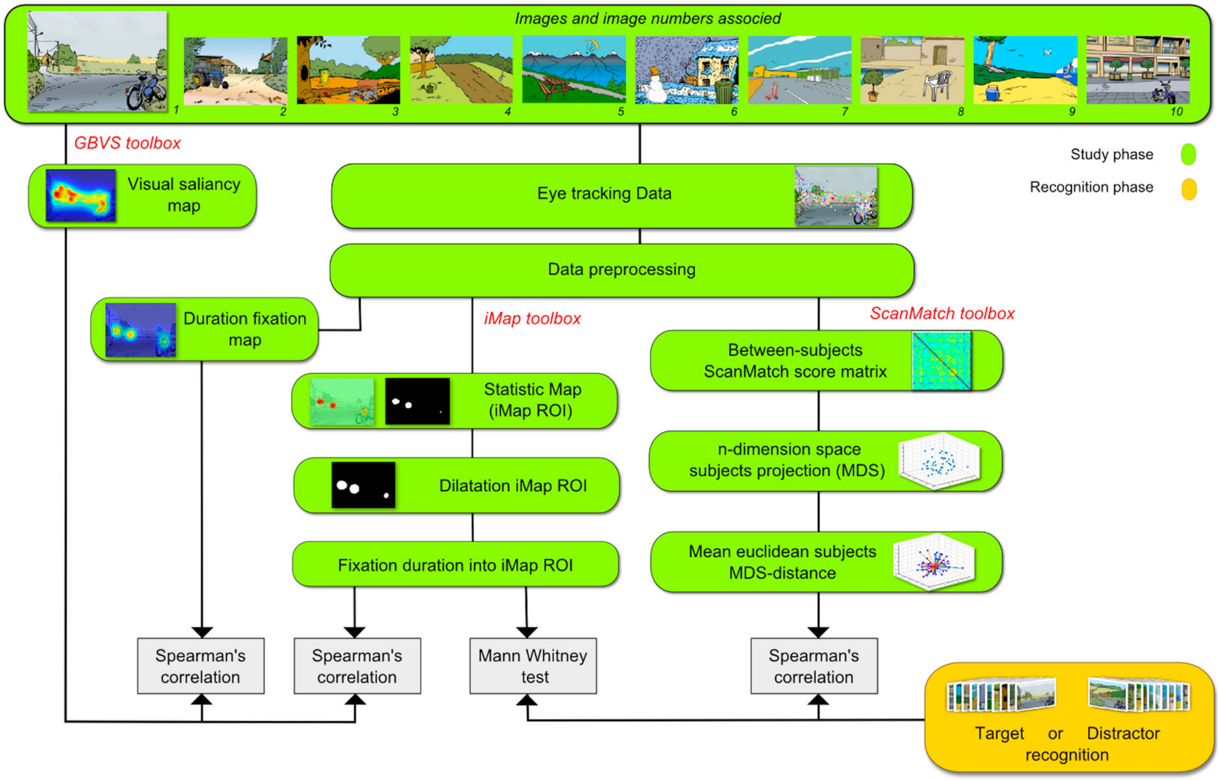

Figure 3. Schematic representation of the procedures for eye-tracking data analysis.

Note: in order to represent the MDS space in the figure, the size of the n-dimensional space was chosen equal to 3 (but in the analysis, 8 was used).

173). An average "duration map" was computed for each picture. A one-tailed t-test was then performed against the "baseline activation" ("the mean fixation intensity within the iMap mask"). We used an alpha level of 0.01 Bonferroni-corrected for multiple comparisons. The output provided, for each scene, a map showing the significant above-chance fixation duration regions. From these maps, we extracted statistical iMap regions of interest (data-driven ROIs), where each data-driven ROI corresponded to a statistically significant cluster of voxels. These data-driven ROIs were binarized. To simulate the central foveal vision, binary ROIs were dilated by 25 pixels, which corresponded to the size, on the image, of half of a viewing angle of 1.5°.

Using these parameters, we calculated the relative fixation duration spent in data-driven ROIs at study phase for four subsequent response categories at test: hits (correctly recognized "targets"), misses (unrecognized "targets"), correct rejections (correctly rejected "distractors"), and false alarms ("distractors" incorrectly identified as "targets"). The relative fixation duration in ROIs was equal to the sum of fixation durations spent in ROI divided by the sum of all fixation durations spent in the picture. A Mann–Whitney test was used to compare relative time spent in ROIs at study phase for each category.

Visual saliency analysis

Visual saliency map. For each picture, a saliency map was created using the GBVS MATLAB toolbox (Graph-Based Visual Saliency, Harel et al., 2007). GBVS was selected over other saliency toolboxes because it offers the highest prediction level for saliency-based ocular fixations, especially for scenes (Borji et al., 2013). The GBVS process is divided into two stages:

Table 1. Total number of fixations and number of first fixations retained for analysis for each target picture.

	Picture										
	1	2	3	4	5	6	7	8	9	10	Total
Total no. of fixations retained	652	708	670	644	626	656	680	658	653	664	6611
No. of 1st fixation retained	4	3	3	4	4	2	4	5	2	6	37

activation maps are first created using specific feature channels, which are then normalized in a way that highlights conspicuity and admits combination with other maps, to generate a final saliency map. We used the default parameters of GBVS.

Fixation duration map. For each participant and for each image, an individual fixation duration map was constructed, by taking the set of locations where the eyes were fixated and the duration of the gaze point. A fixation duration map was computed for each image by adding up the fixation durations of all participants (O'Connell & Walther, 2015). To simulate the central foveal vision, the result was then convolved with a Gaussian kernel. The full width at half maximum of the Gaussian kernel was set to 1.5° of visual angle.

Correlation between saliency map and fixation duration map. For each picture, a Spearman's rank correlation was used to measure the similarity between saliency map and fixation duration map (Riche et al., 2013).

Correlation between visual saliency and fixation ROI. A Spearman's rank correlation was used to test the relation between saliency and data-driven ROI fixation duration found using iMap. In the saliency interaction analysis, pixel-level saliency for each ROI was selected as the maximum value of the object region in order to minimize the object size effect. This was because big objects tend to include uniform texture regions and thus have much smaller average pixel-level saliency, while fixations were normally attracted to the most salient region of an object. Thus, maximum saliency rather than average saliency was more representative of pixel-level saliency of an object (Wang et al., 2015).

Scanpath analysis

The similarity between participants' sequences of eye movements was analysed using the ScanMatch open source MATLAB toolbox (Cristino et al., 2010). This choice was guided by the Anderson et al. (2015) study which compared scanpaths methods and concluded that ScanMatch is the best adapted tool for analysing fixation sequences, since it can take into account spatial location, temporal duration, and sequential similarity between scanpaths.

For the spatial binning, each picture was divided with a grid composed of 96 (12 × 8) rectangular ROIs of 85 × 89 pixels. This grid size was determined without a priori: we used the same size bin as Cristino et al. (2010). Each rectangular ROI was labelled by a combination of two letters. Each fixation within the ROI was tagged with its name in the string sequence: a sequence contains a list of visited ROIs, respecting the order of fixations for a specific picture. For the temporal binning, the letters corresponding to a ROI were repeated in the sequence proportionally to the fixation duration: string sequences were divided into 100 ms bins.

Then, the ScanMatch string-edit distance methodology was used to find the best alignment over the whole string of two sequences by maximizing its score. This methodology was based on the Needleman–Wunsch algorithm which has been implemented to compare DNA sequence. To this end, the sequences were aligned based on a substitution matrix which provides a score for every alignment, based on the spatial relationship between ROIs. A similarity score of 1 corresponds to sequences being identical, while a similarity score of 0 indicates that there is no similarity.

We used a gap penalty of 0 which means that adding gaps decreased similarity scores (Frame et al., 2019). The "threshold value" was equal to 2 standard deviations of all the saccade amplitudes. This means that the alignment algorithm only aimed to align regions which were within the variability of the saccade amplitudes (Cristino et al., 2010).

The ScanMatch algorithm was used to calculate the similarity score between participant pairs, resulting in between-participant ScanMatch score matrix (matrix size: 42 × 42 participants).

ScanMatch statistical analysis. This matrix was transformed with MATLAB function "mdscale," allowing MultiDimensional Scaling (MDS) to be performed (Kruskal, 1964) and projection of each participant in an n-dimensional space. MDS then computed the participant coordinates in this n-dimensional space.

Beforehand, it was necessary to determine the number of dimensions of this space. The goodness-of-fit of such space was quantified using a residual sum of squares called the stress of the map. Stress values are positive and small values are better, e.g., a stress value of 10% indicates a fair fit (Kruskal, 1964).

The stress value depends on the space dimension: the greater the n-dimensional space, the smaller the stress value. We chose the size of the n-dimensional space so that the stress value of all the scenes was strictly less than 7.5%, which guaranteed a correct fit.

At the end of this procedure, MDS provided relative locations for all participants. Participants with similar temporo-spatial gaze patterns were plotted close together near the centre of the space while those with atypical gaze patterns were plotted towards the periphery. The median location of the group, i.e., the MDS median point, was defined as the centre of the participants' group and represented the standard temporo-spatial gaze pattern. The relative location of participants was then computed by calculating the Euclidean distance (i.e., L2-norm) between the respective MDS coordinates participants and MDS median point. The mean MDS-distance of the 42 participants was calculated for each picture based on eye gaze patterns at study phase; higher MDS-distance values reflected dissimilar temporo-spatial gaze patterns.

To test the relation between dissimilarity at study phase and the subsequent recognition performance at test, we conducted Spearman correlations between the mean MDS-distance of 42 participants and accuracy scores, i.e., the number of targets correctly recognized and distractors correctly rejected.

Results

Behavioural performance

During the test phase, participants correctly recognized 79.52 ± 13.96% of target scenes, correctly rejected 70.24 ± 15.85% of distractor scenes and correctly rejected 97.86 ± 4.15% of new scenes (Figure 4). A Friedman's test (three conditions: target, distractor, new) revealed a main effect of condition ($\chi^2_F(2) = 64.78$, $p < 0.001$). A Wilcoxon signed-rank test indicated that all of these results were significantly different from each other: Target (*Mdn* = 80%) vs Distractor (*Mdn* = 70%): $T = 129$, $z = 2.71$, $p < 0.01$; Target vs New (*Mdn* = 100%): $T = 5.5$, $z = 5.07$, $p < 0.001$; Distractor vs New: $T = 0$, $z = 5.65$, $p < 0.001$.

Fixation duration in iMap data-driven ROIs

A total of 20 data-driven ROIs were identified, at study phase, in the 10 target scenes (Table 2, Figure 5).

There was no significant difference in the relative fixation time in data-driven ROIs at study phase, between subsequent correctly recognized targets (Hits) (*Mdn* = 30.22%) and missed targets (Misses) at test (*Mdn* = 31.55%), ($U(N_{Hits} = 334, N_{Misses} = 86) = 14{,}099$, $z = 0.40$, $p = 0.80$, Mann Whitney) (Figure 6). By contrast, we found a significant difference in the relative fixation time in data-driven ROIs at study phase, between false recognitions (false alarms) and correct rejection of distractors, with longer fixation durations for the former (*Mdn* = 38.46%) compared to correct rejections (*Mdn* = 29.98%), ($U(N_{Correct\ Rejections} = 295, N_{False\ Alarms} = 125) = 14191$, $z = 3.73$, $p < 0.001$, Mann Whitney) (Figure 6). This means that the more time participants spent looking in data-driven ROIs in the study phase, the less successful they were at rejecting "distractor" pictures in the test phase.

Visual saliency

We found a significant positive correlation between visual saliency and fixation duration for whole pictures at study phase (Spearman rho range: [0.526; 0.773], $p < 0.001$, $N = 1024 \times 717$ pixels) (Figure 7). By contrast, there was no significant correlation between the visual saliency of data-driven ROIs and the relative fixation duration time in these ROIs (Spearman rho = 0.332, $N = 840$), which confirmed that fixations were not driven by low-level scene features, i.e., saliency.

Scanmatch results

We chose a size of 8 for the n-dimensional space, for which all the stress values of images were strictly less than 7.5% (Figure 8). The stress values, found and used to determine the goodness of fit of an MDS solution, are in accordance with the literature. For instance, in an eye-tracking study consisting of sentence analysis in 44 students (von der Malsburg & Vasishth, 2011), the stress value was found to be 22% for a 2-dimensional space and decreased to 8.1% in a 7-dimensional space.

Using an 8-dimensional space, we determined the mean MDS-distance of 42 subjects for each picture seen during the study phase (L2-norm mean, Table 3), and calculated the Spearman correlation coefficients

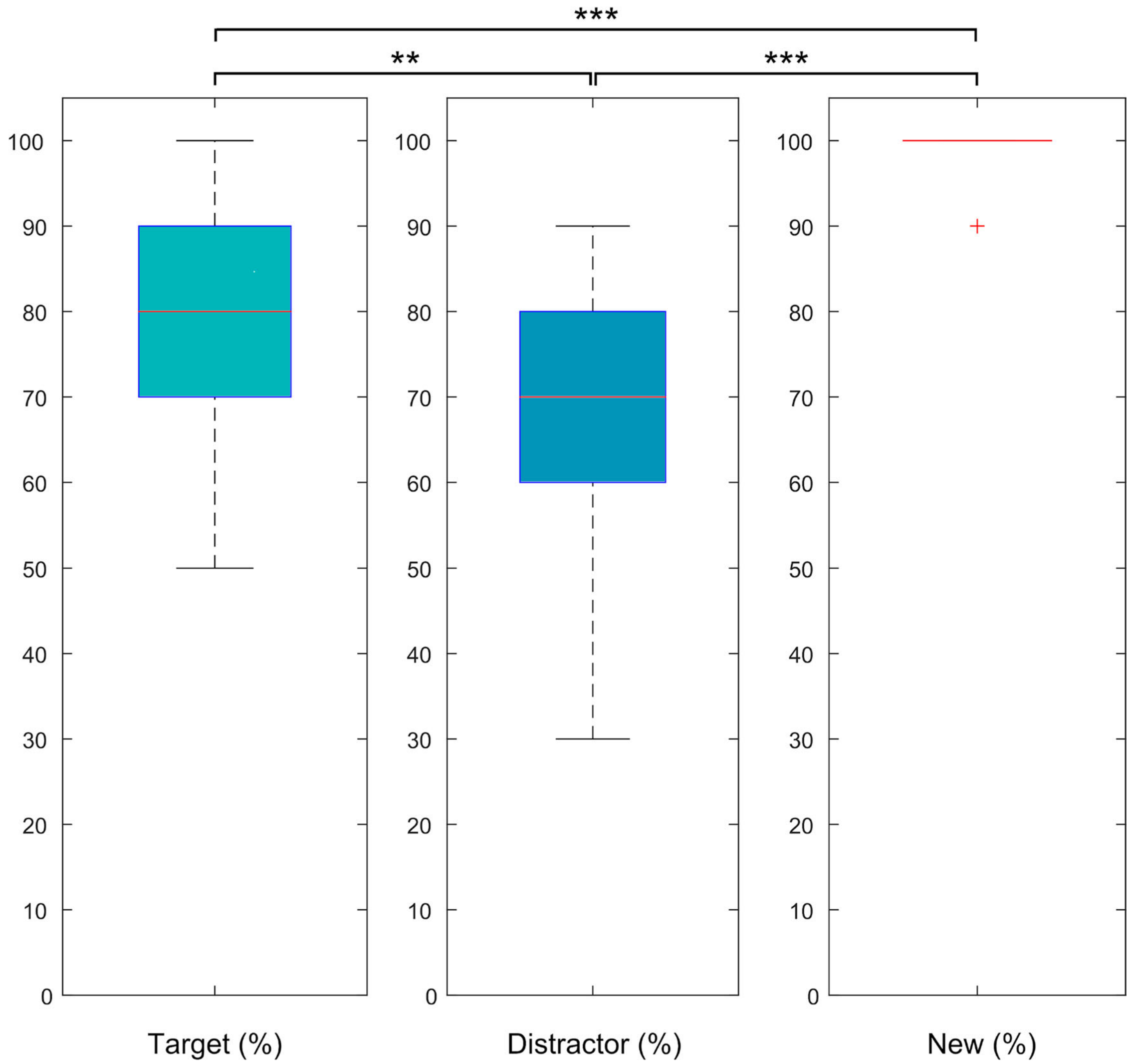

Figure 4. Recognition accuracy: percentage of pictures correctly recognized (targets) or correctly rejected (distractors and new) in recognition phase (** $p < 0.01$, *** $p < 0.001$).

Note: The red line indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol.

between recognition performances (Hits for "targets" and Correct Rejection of "distractors") and this average distance.

There was no significant correlation between the proportion of correctly recognized targets and average MDS-distance (Spearman rho = −0.092, $p = 0.80$, $N = 10$) (Figure 9). By contrast, we identified a significant negative correlation between the proportion of correctly rejected distractors and average MDS-distance (Spearman rho = −0.659, $p = 0.04$, $N = 10$) (Figure 9). This result indicates that the more similar the picture exploration was between subjects in the study phase, the more correctly they rejected "distractor" pictures at test.

Table 2. Number of data-driven regions of interest identified in each target picture during study phase.

	Picture										
	1	2	3	4	5	6	7	8	9	10	Total
Number of ROI	3	2	2	1	2	1	2	2	3	2	20

Discussion

This study aimed to provide a better understanding of how scene exploration strategies during free viewing at study phase are associated with subsequent memory discrimination between targets, distractors, and new scenes at test phase. Eye movement data

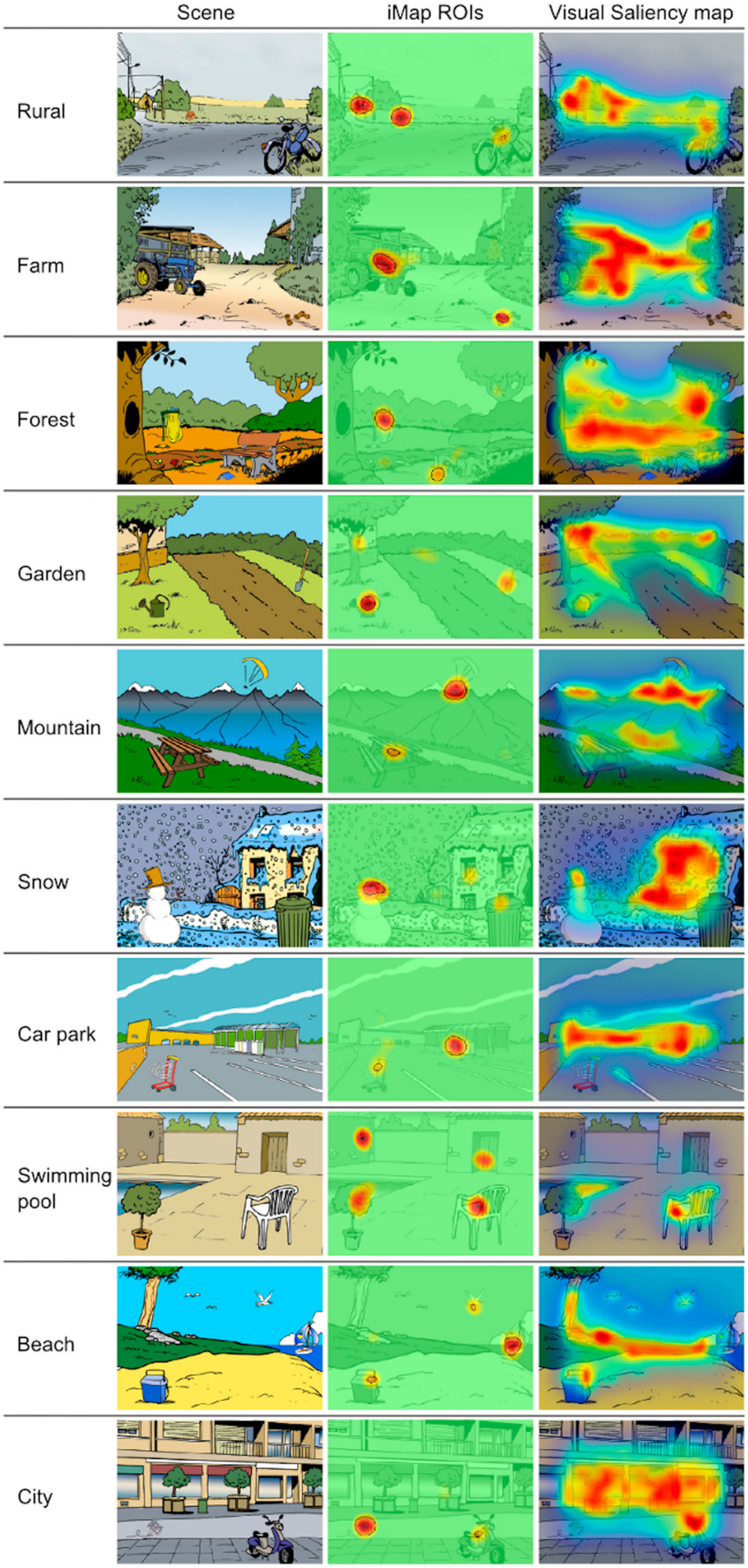

Figure 5. Data-driven ROIs ($p < 0.01$ Bonferroni corrected, significant area marked by dark line) and visual saliency (obtained using GBVS) for scene viewing during study phase.

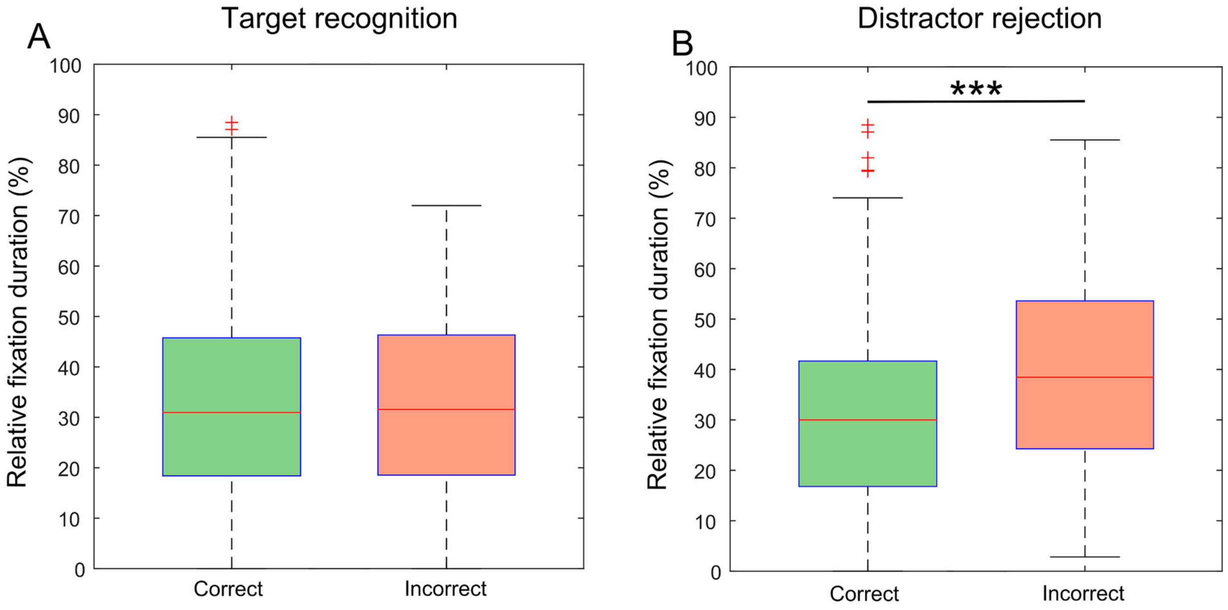

Figure 6. Relative fixation duration in study phase corresponding to subsequent target hits or misses (A), correct rejections and false alarms (B) in recognition phase (*** $p < 0.001$).

Note: The red line indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol.

were analysed at study phase using two data-driven methods, namely a fixation density map (using iMap4) associated with a saliency map (using GBVS), and scanpath analyses without a priori (using the ScanMatch toolbox). First, we identified longer fixation durations in data-driven ROIs for subsequent false alarms over rejection of distractors, while there was no significant difference in the relative fixation time in data-driven ROIs between subsequent hits and misses, indicating that fixation maps were more an indicator of memory discrimination accuracy than of recognition performance. Second, we identified a negative correlation between average MDS-distance scanpaths and the correct rejection of distractors, and no significant correlation between average MDS-distance and target recognition performance, which suggests that scanpath consistency (or inter-observer congruency of scanpaths) at study phase was a factor of subsequent memory discrimination abilities, rather than of memory performance. Taken together, eye-tracking can provide insights into the pattern separation process, suggesting that

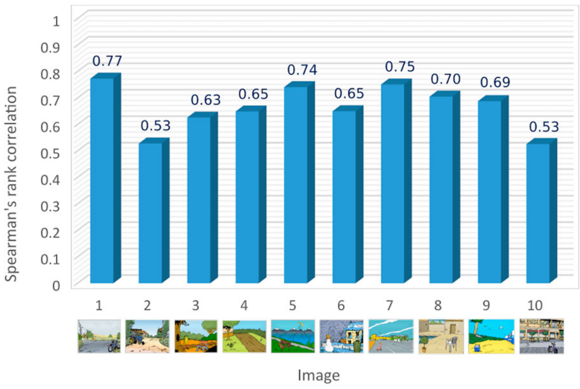

Figure 7. Spearman's rank correlation between visual saliency and fixation duration, during the study phase, in the whole picture for each target picture. The numbers above each bar correspond to Spearman's rank coefficient.

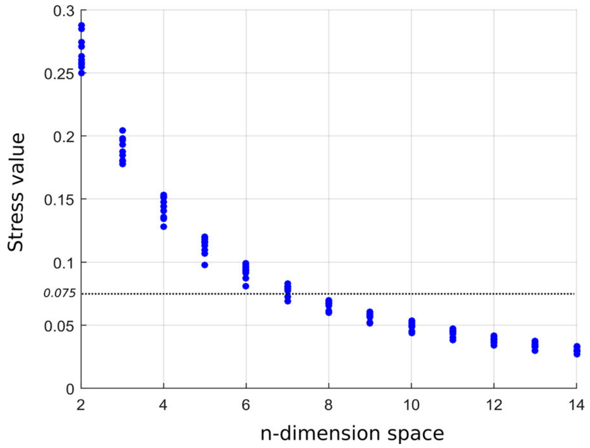

Figure 8. Stress value and dimensional-space for the 10 scenes; 0.075 corresponds to the maximum accepted value for the stress value.

Table 3. For each target picture, % of participants who recognized target correctly, rejected distractor correctly, and L2-norm mean.

Picture	% of target pictures correctly recognized	% of distractor pictures correctly rejected	L2-norm mean
1	78.57	52.38	0.3619
2	69.05	64.29	0.40743
3	73.81	85.71	0.35782
4	80.95	78.57	0.35135
5	88.10	40.48	0.40763
6	85.71	83.33	0.34745
7	80.95	69.05	0.37707
8	88.10	83.33	0.38137
9	61.90	59.52	0.39328
10	88.10	85.71	0.36134
Total	79.52	70.24	

a broad and consistent exploration during encoding increases subsequent memory discrimination.

First, longer fixation durations in data-driven ROIs (using iMap4) were associated with subsequent false alarms over rejection of distractors. As fixation locations are a marker of visual attention (Vidal et al., 2012), this result implies that the more the participants focused their attention on these ROIs at encoding, the less they explored the rest of the visual scene, to be able subsequently to reject distractors. Using single items, Molitor et al. (2014) have previously suggested that low abilities in memory discrimination – i.e., false alarms – results from insufficient number of fixations at study – i.e., the poor encoding hypothesis. The current results go further, by showing that the relative fixation duration in data-driven ROIs is a marker of subsequent memory discrimination. A high relative fixation duration in some ROIs implies a low relative fixation duration outside these ROIs that accounts for a reduced global exploration and is predictive of false alarms. Memory discrimination requires high-resolution mnemonic representations of studied items supported by the *pattern separation* process, and the intra-hippocampal binding of features that constitute an event (Hunsaker & Kesner, 2013). This suggests that encoding details with high specificity to be subsequently discriminated from other similar lures requires a broad exploration of scenes during encoding. Interestingly, the relative fixation time in data-driven ROIs at study phase did not differ for subsequent hits and misses, indicating that fixation time is not coupled with recognition performance. In line with this finding, several studies have shown that recognition performance is better when fixation times are shorter (see Choe et al., 2017, in a search task, and Parag & Vakil, 2018, with faces), or that they are independent parameters (Schomaker & Wittmann, 2017).

Second, the positive correlation between the visual saliency and fixation duration for the full picture highlights the role of visual saliency on full scene exploration, as previously demonstrated in landscape photographs (Dupont et al., 2016). Interestingly, the absence of significant correlations between the visual saliency of data-driven ROIs and the fixation duration time in these ROIs confirmed that data-driven fixation maps were not related to low-level scene features, i.e., saliency.

Third, scanpath analyses revealed no significant correlation between the average MDS-distance and target recognition performance which appeared to be independent factors. This result is consistent with other studies focusing on scanpaths and scene recognition, which showed that performance is related to scanpath idiosyncrasy (i.e., greater within- than between-participants similarity when compared at encoding and at retrieval, Foulsham et al., 2012), or that recognition performance is relatively dissociable from scanpath analyses (Foulsham & Kingstone, 2013). By contrast, we identified a negative correlation between the average MDS-distance and the correct rejection of distractors, which implies that scanpath consistency across observers during scene exploration at study phase is a significant factor in the ability to discriminate distractors from targets at test. This hypothesis is congruent with Lyu et al.'s (2019) study which identified that the *consistency of fixation maps* across viewers was predictive of scene memory. A few studies have highlighted scanpath consistency, using various paradigms. Shakespeare et al. (2015) observed higher scanpath consistency between healthy participants compared to neurological patients in search and memory tasks with scenes, although not considering the order or duration of fixations in their analyses, and suggested a *disorganized approach* when patients viewed scenes. Using a deep neural network, Wei et al. (2017) developed a method to predict which features will capture the most attention in a visual scene (i.e., *gaze agreement*), and the most consistent scanpath across viewers (i.e., *scanpath agreement*). More recently, Frame et al. (2019), using dynamic surveillance videos and a guided search task, showed that an effective search strategy was associated with consistent scanpaths

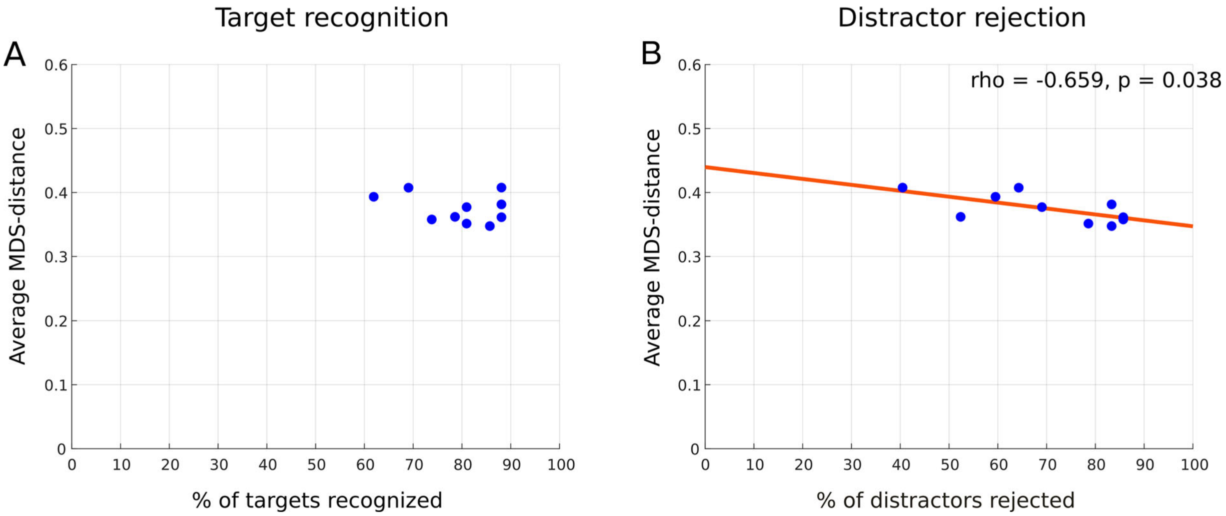

Figure 9. Relationship between average MDS-distance and percentage of: targets correctly recognized (A), distractors correctly rejected (B) (see Table 3).

across observers. Together, our results suggest that there may be an optimal scene exploration strategy during free viewing at study phase to encode foreground and background scene features with high specificity, leading to a better memory discrimination at test phase.

Although this study provides interesting results, there are some limitations that open up opportunities for further research. First, to further study visual exploration strategies during scene encoding, we need to evaluate how the visualization of items of different sizes may influence fixation density maps and behavioural performance. Because participants viewed large items at different points in space, iMap4 did not detect their fixations, which led us to use a 2.5° smoothing. In addition, statistical data-driven ROIs comparisons are more difficult to realize, and their interpretation is mainly visual. Second, we used a small number of images, which may limit the statistical power to assess memory discrimination. In addition, using same items in both target and distractor supported a certain degree of visual similarity necessary to assess memory discrimination, but did not allow control of whether the proportion of eye movements in these items versus the background influenced the rejection of the distractors. Third, we suggested that the consistency of scanpath at scene encoding affects the subsequent identification of “distractors” at test. The next step would be to identify scene characteristics that may facilitate scanpath consistency and increase memorization. This would open up new possibilities in terms of care and support for people with memory difficulties.

Conclusions

To conclude, eye-tracking analysis methods without a priori are particularly suitable to study encoding in memory for visual scenes, given their size and the amount of information they contain. In particular, we showed that iMap4 and ScanMatch toolboxes are valuable methods to study gaze movements associated with scene memory discrimination. The current study contributes to our growing understanding of memory strategies at encoding. Most importantly, we identified that scene memory discrimination is associated with a wide gaze exploration at study phase, and a consistent scanpath. Future studies may further explore the inter-observer scanpath consistency and their relationship to memory and cognitive parameters.

Acknowledgements

Authors are grateful to L. Raoult, C. Piet and L. Legentil for their help with recruitment, cognitive testing and eye-tracking examination. We thank M. Lalevée and J. Spiess for the creation of drawings. We also thank F. Wallois, C. Mille and J.M. Guilé as well as the clinical team who participated into inclusions at the Amiens hospital.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by the French *Conseil Régional de Normandie* and *Programme Hospitalier de Recherche Clinique*.

ORCID

B. Guillery-Girard http://orcid.org/0000-0002-2316-4716

References

Anderson, N. C., Anderson, F., Kingstone, A., & Bischof, W. F. (2015). A comparison of scanpath comparison methods. *Behavior Research Methods*, *47*(4), 1377–1392. https://doi.org/10.3758/s13428-014-0550-3

Arizpe, J. M., Noles, D. L., Tsao, J. W., & Chan, A. W. Y. (2019). Eye movement dynamics differ between encoding and recognition of faces. *Vision*, *3*(1), 9. https://doi.org/10.3390/vision3010009

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. In *Clinical psychology Review*, *32*(8), 704–723. https://doi.org/10.1016/j.cpr.2012.09.004.

Borji, A., Sihite, D. N., & Itti, L. (2013). What stands out in a scene? A study of human explicit saliency judgment. *Vision Research*, *91*, 62–77. https://doi.org/10.1016/j.visres.2013.07.016

Caldara, R., & Miellet, S. (2011). iMap: A novel method for statistical fixation mapping of eye movement data. *Behavior Research Methods*, *43*(3), 864–878. https://doi.org/10.3758/s13428-011-0092-x

Castelhano, M. S., & Krzyś, K. (2020). Rethinking space: A review of perception, attention, and memory in scene processing. *Annual Review of Vision Science*, *6*(1), 563–586. https://doi.org/10.1146/annurev-vision-121219-081745

Choe, K. W., Kardan, O., Kotabe, H. P., Henderson, J. M., & Berman, M. G. (2017). To search or to like: Mapping fixations to differentiate two forms of incidental scene memory. *Journal of Vision*, *17*(12), 8. https://doi.org/10.1167/17.12.8

Cristino, F., Mathôt, S., Theeuwes, J., & Gilchrist, I. D. (2010). Scanmatch: A novel method for comparing fixation sequences. *Behavior Research Methods*, *42*(3), 692–700. https://doi.org/10.3758/BRM.42.3.692

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

De Agostini, M., & Dellatolas, G. (1988). Une épreuve simple pour évaluer la préférence manuelle chez l'enfant à partir de 3 ans. *Enfance; Psychologie, Pedagogie, Neuropsychiatrie, Sociologie*, *41*(3), 139–147. https://doi.org/10.3406/enfan.1988.2161

Declaration of Helsinki. (2008, octobre 22). *WMA declaration of Helsinki—Ethical principles for medical research involving human subjects*. https://doi.org/https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/

Dupont, L., Ooms, K., Antrop, M., & Van Eetvelde, V. (2016). Comparing saliency maps and eye-tracking focus maps: The potential use in visual impact assessment based on landscape photographs. *Landscape and Urban Planning*, *148*, 17–26. https://doi.org/10.1016/j.landurbplan.2015.12.007

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

Foulsham, T., Dewhurst, R., Nyström, M., Jarodzka, H., Johansson, R., Underwood, G., & Holmqvist, K. (2012). Comparing scanpaths during scene encoding and recognition: A multi-dimensional approach. *Journal of Eye Movement Research*, *5*(4), 3. https://doi.org/10.16910/jemr.5.4.3

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

Foulsham, T., & Underwood, G. (2011). If visual saliency predicts search, then why? Evidence from normal and gaze-contingent search tasks in natural scenes. *Cognitive Computation*, *3*(1), 48–63. https://doi.org/10.1007/s12559-010-9069-9

Frame, M. E., Warren, R., & Maresca, A. M. (2019). Scanpath comparisons for complex visual search in a naturalistic environment. *Behavior Research Methods*, *51*(3), 1454–1470. https://doi.org/10.3758/s13428-018-1154-0

French, R. M., Glady, Y., & Thibaut, J.-P. (2017). An evaluation of scanpath-comparison and machine-learning classification algorithms used to study the dynamics of analogy making. *Behavior Research Methods*, *49*(4), 1291–1302. https://doi.org/10.3758/s13428-016-0788-z

Hainmueller, T., & Bartos, M. (2020). Dentate gyrus circuits for encoding, retrieval and discrimination of episodic memories. *Nature Reviews Neuroscience*, *21*(3), 153–168. https://doi.org/10.1038/s41583-019-0260-z

Harel, J., Koch, C., & Perona, P. (2007). Graph-based visual saliency. In B. Scholkopf, J. Platt, & T. Hoffman (Eds.), *Advance in neural information processing systems (NIPS)* (pp. 545–552). MIT Press.

Hayes, T. R., & Henderson, J. M. (2017). Scan patterns during real-world scene viewing predict individual differences in cognitive capacity. *Journal of Vision*, *17*(5), 23. https://doi.org/10.1167/17.5.23

Henderson, J. M., & Hayes, T. R. (2018). Meaning guides attention in real-world scene images: Evidence from eye movements and meaning maps. *Journal of Vision*, *18*(6), 10. https://doi.org/10.1167/18.6.10

Hoffman, J. (1998). Visual attention and eye movements. In H. Pashler (Ed.), *Attention* (pp. 119–153). Psychology Press/Erlbaum (UK) Taylor & Francis. https://psycnet.apa.org/record/1998-07791-003

Hunsaker, M. R., & Kesner, R. P. (2013). The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neuroscience & Biobehavioral Reviews*, *37*(1), 36–58. https://doi.org/10.1016/j.neubiorev.2012.09.014

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Kafkas, A., & Montaldi, D. (2012). Familiarity and recollection produce distinct eye movement, pupil and medial temporal lobe responses when memory strength is matched. *Neuropsychologia*, *50*(13), 3080–3093. https://doi.org/10.1016/j.neuropsychologia.2012.08.001

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*(1), 1–27. https://doi.org/10.1007/BF02289565

Lao, J., Miellet, S., Pernet, C., Sokhn, N., & Caldara, R. (2017). iMap4: An open source toolbox for the statistical fixation mapping of eye movement data with linear mixed modeling. *Behavior Research Methods*, *49*(2), 559–575. https://doi.org/10.3758/s13428-016-0737-x

Liu, K. Y., Gould, R. L., Coulson, M. C., Ward, E. V., & Howard, R. J. (2016). Tests of pattern separation and pattern completion in humans-A systematic review. *Hippocampus*, *26*(6), 705–717. https://doi.org/10.1002/hipo.22561

Loiotile, R. E., & Courtney, S. M. (2015). A signal detection theory analysis of behavioral pattern separation paradigms. *Learning & Memory*, *22*(8), 364–369. https://doi.org/10.1101/lm.038141.115

Lyu, M., Choe, K. W., Kardan, O., Kotabe, H. P., Henderson, J. M., & Berman, M. G. (2019). Scenes that produce more consistent fixation maps are more memorable. *PsyArXiv Preprints*. https://pdfs.semanticscholar.org/50b7/6aa18ed9ba66e0cca3d5f630307981b2512d.pdf

Meghanathan, R. N., Nikolaev, A. R., & van Leeuwen, C. (2019). Refixation patterns reveal memory-encoding strategies in free viewing. *Attention, Perception, & Psychophysics*, *81*(7), 2499–2516. https://doi.org/10.3758/s13414-019-01735-2

Molitor, R. J., Ko, P. C., Hussey, E. P., & Ally, B. A. (2014). Memory-related eye movements challenge behavioral measures of pattern completion and pattern separation. *Hippocampus*, *24*(6), 666–672. https://doi.org/10.1002/hipo.22256

Nicholls, V. I., Jean-Charles, G., Lao, J., de Lissa, P., Caldara, R., & Miellet, S. (2019). Developing attentional control in naturalistic dynamic road crossing situations. *Scientific Reports*, *9*(1), 4176. https://doi.org/10.1038/s41598-019-39737-7

Noton, D., & Stark, L. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, *11*(9), 929–IN8. https://doi.org/10.1016/0042-6989(71)90213-6

O'Connell, T. P., & Walther, D. B. (2015). Dissociation of salience-driven and content-driven spatial attention to scene category with predictive decoding of gaze patterns. *Journal of Vision*, *15*(5), 20. https://doi.org/10.1167/15.5.20

Parag, O., & Vakil, E. (2018). Distinct eye movements for different cognitive processes as expressed in the face recognition task. *Memory (Hove, England)*, *26*(4), 524–534. https://doi.org/10.1080/09658211.2017.1381265

Park, J., Jin, Y., Ahn, S., & Lee, S. (2019). The impact of design representation on visual perception: Comparing eye-tracking data of architectural scenes between photography and line drawing. *Archives of Design Research*, *32*(1), 5–29. https://doi.org/10.15187/adr.2019.02.32.1.5

Ramey, M. M., Henderson, J. M., & Yonelinas, A. P. (2020a). The spatial distribution of attention predicts familiarity strength during encoding and retrieval. *Journal of Experimental Psychology: General*, *149*(11), 2046–2062. https://doi.org/10.1037/xge0000758

Ramey, M. M., Yonelinas, A. P., & Henderson, J. M. (2020b). Why do we retrace our visual steps? Semantic and episodic memory in gaze reinstatement. *Learning & Memory*, *27*(7), 275–283. https://doi.org/10.1101/lm.051227.119

Riche, N., Duvinage, M., Mancas, M., Gosselin, B., & Dutoit, T.. (2013). Saliency and human fixations: State-of-the-art and study of comparison metrics. In *2013 IEEE International Conference on Computer Vision, Sydney, NSW, Australia* (pp. 1153–1160). IEEE. https://doi.org/10.1109/ICCV.2013.147.

Rolls, E. T. (2016). Pattern separation, completion, and categorisation in the hippocampus and neocortex. *Neurobiology of Learning and Memory*, *129*, 4–28. https://doi.org/10.1016/j.nlm.2015.07.008

Schomaker, J., & Wittmann, B. C. (2017). Memory performance for everyday motivational and neutral objects is dissociable from attention. *Frontiers in Behavioral Neuroscience*, *11*(June), 1–13. https://doi.org/10.3389/fnbeh.2017.00121

Schurgin, M. W. (2018). Visual memory, the long and the short of it: A review of visual working memory and long-term memory. *Attention, Perception, & Psychophysics*, *80*(5), 1035–1056. https://doi.org/10.3758/s13414-018-1522-y

Shakespeare, T. J., Pertzov, Y., Yong, K. X. X., Nicholas, J., & Crutch, S. J. (2015). Reduced modulation of scanpaths in response to task demands in posterior cortical atrophy. *Neuropsychologia*, *68*, 190–200. https://doi.org/10.1016/j.neuropsychologia.2015.01.020

Sharot, T., Davidson, M. L., Carson, M. M., & Phelps, E. A. (2008). Eye movements predict recollective experience. *PLoS ONE*, *3*(8), e2884. https://doi.org/10.1371/journal.pone.0002884

Toh, W. L., Rossell, S. L., & Castle, D. J. (2011). Current visual scanpath research: A review of investigations into the psychotic, anxiety, and mood disorders. *Comprehensive Psychiatry*, *52*(6), 567–579. https://doi.org/10.1016/j.comppsych.2010.12.005

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, *59*(11), 1931–1949. https://doi.org/10.1080/17470210500416342

Vidal, M., Turner, J., Bulling, A., & Gellersen, H. (2012). Wearable eye tracking for mental health monitoring. *Computer Communications*, *35*(11), 1306–1311. https://doi.org/10.1016/j.comcom.2011.11.002

von der Malsburg, T., & Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language*, *65*(2), 109–127. https://doi.org/10.1016/j.jml.2011.02.004

Wang, S., Jiang, M., Duchesne, X. M., Laugeson, E. A., Kennedy, D. P., Adolphs, R., & Zhao, Q. (2015). Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. *Neuron*, *88*(3), 604–616. https://doi.org/10.1016/j.neuron.2015.09.042

Wei, Z., Adeli, H., Hoai, M., Zelinsky, G., & Samaras, D. (2017). Predicting scanpath agreement during scene viewing using deep neural networks. *Journal of Vision*, *17*(10), 749. https://doi.org/10.1167/17.10.749

Williams, C. C., & Castelhano, M. S. (2019). The changing landscape: High-level influences on eye movement guidance in scenes. *Vision*, *3*(3), 33. https://doi.org/10.3390/vision3030033

Wynn, J. S., Bone, M. B., Dragan, M. C., Hoffman, K. L., Buchsbaum, B. R., & Ryan, J. D. (2016). Selective scanpath repetition during memory-guided visual search. *Visual Cognition*, *24*(1), 15–37. https://doi.org/10.1080/13506285.2016.1175531

Wynn, J. S., Ryan, J. D., & Buchsbaum, B. R. (2020). Eye movements support behavioral pattern completion. *Proceedings of the National Academy of Sciences*, *117*(11), 6246–6254. https://doi.org/10.1073/pnas.1917586117