PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Johansson R, Nyström M, Dewhurst R, Johansson M. 2022 Eye-movement replay supports episodic remembering. *Proc. R. Soc. B* **289**: 20220964. https://doi.org/10.1098/rspb.2022.0964

Received: 18 May 2022
Accepted: 25 May 2022

Subject Category:
Neuroscience and cognition

Subject Areas:
cognition, behaviour, neuroscience

Keywords:
eye movements, episodic memory, replay, scanpaths, reinstatement

Author for correspondence:
Roger Johansson
e-mail: roger.johansson@psy.lu.se

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.6035691.

THE ROYAL SOCIETY PUBLISHING

Eye-movement replay supports episodic remembering

Roger Johansson[1], Marcus Nyström[2], Richard Dewhurst[3] and Mikael Johansson[1]

[1]Department of Psychology, and [2]Humanities Lab, Lund University, Lund, Sweden
[3]Interacting Minds Centre, Århus University, Aarhus, Denmark

RJ, 0000-0003-3434-2538; MJ, 0000-0002-2903-6301

When we bring to mind something we have seen before, our eyes spontaneously unfold in a sequential pattern strikingly similar to that made during the original encounter, even in the absence of supporting visual input. Oculomotor movements of the eye may then serve the opposite purpose of acquiring new visual information; they may serve as self-generated cues, pointing to stored memories. Over 50 years ago Donald Hebb, the forefather of cognitive neuroscience, posited that such a sequential *replay* of eye movements supports our ability to mentally recreate visuospatial relations during episodic remembering. However, direct evidence for this influential claim is lacking. Here we isolate the sequential properties of spontaneous eye movements during encoding and retrieval in a pure recall memory task and capture their encoding-retrieval overlap. Critically, we show that the fidelity with which a series of consecutive eye movements from initial encoding is sequentially retained during subsequent retrieval predicts the quality of the recalled memory. Our findings provide direct evidence that such *scanpaths* are replayed to assemble and reconstruct spatio-temporal relations as we remember and further suggest that distinct scanpath properties differentially contribute depending on the nature of the goal-relevant memory.

1. Background

Episodic memory refers to our ability to mentally travel back in time to relive past experiences in vivid detail [1]. The formation of coherent episodic memories critically hinges upon the binding of spatio-temporal relationships into a context, which predominantly depends on how we visually 'sample' the world when we act upon it via eye movements [2,3]. Eye movements unfold in sequences of fixations and saccades, where fixations are the brief moments that allow us to sample visual information, and saccades are the rapid movements that occur from one fixation point to another. Although only a limited amount of information can be processed at each fixation point, a sequence of consecutive fixations and saccades can effectively bind multiple inter-related episodic details together, allowing us to encode a memory representation of the event as a whole [4–12]. Thus, our visual sampling of the world is highly predictive of the content and quality of episodic memory formation [6–12]. Spontaneous eye movements also occur during episodic recollections, i.e. when the previously encoded event information is internalized mentally in the *absence of supporting visual input*. These have been demonstrated to broadly reproduce the gaze patterns that were established during encoding [13–19]. A prominent view holds that such gaze reinstatements actively support episodic remembering [20–22] and that the unfolding sequence of connected eye movements may serve to reconstruct episodic memories across time and space [13,14,22–26]. This claim receives support from previous research showing that the extent of spatial encoding-retrieval overlaps in gaze patterns positively correlates with retrieval performance [14,27,28], and with neural

© 2022 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

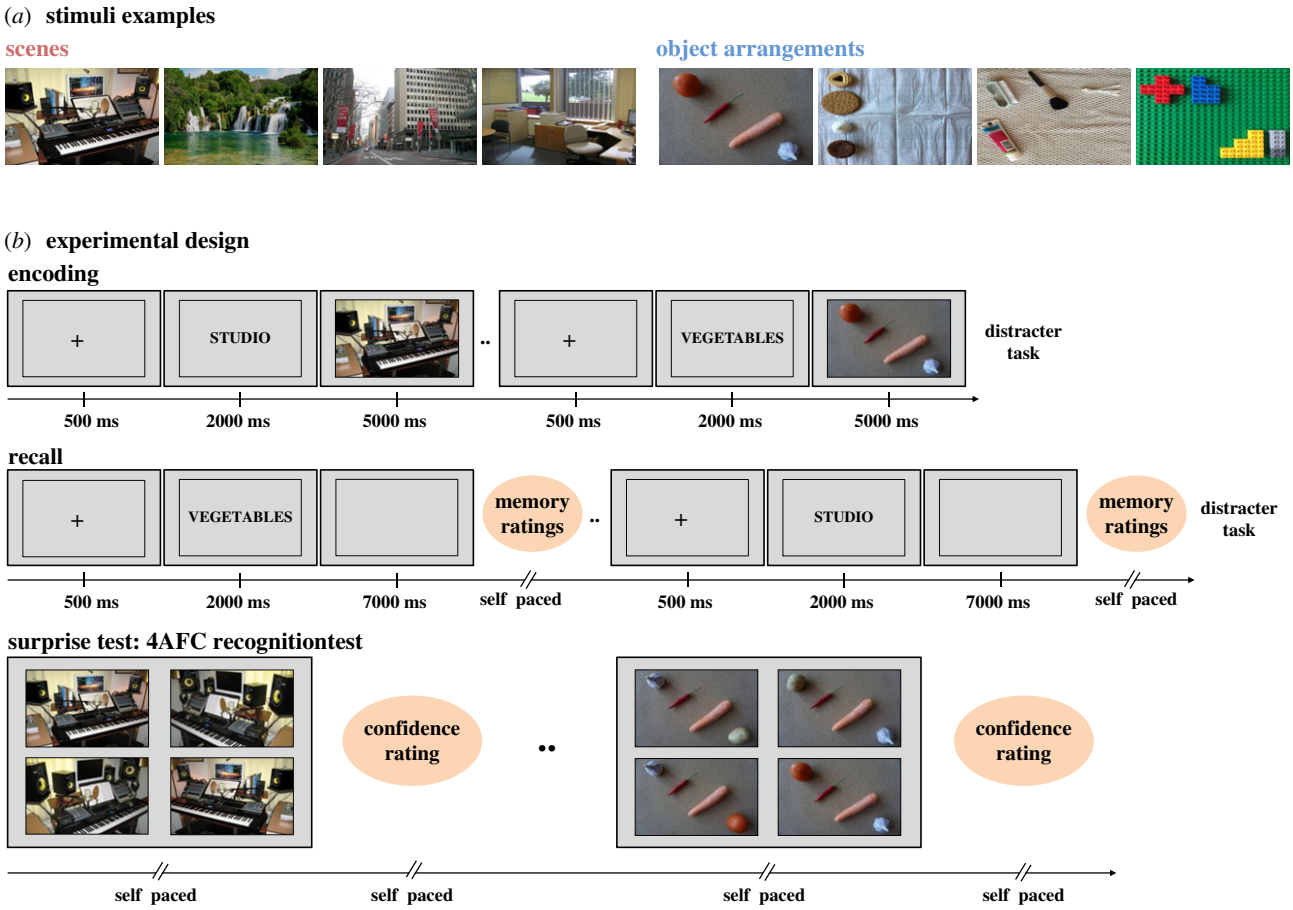

Figure 1. Encoding and recollection of scenes and object arrangements. (*a*) Example of stimuli images (scenes: studio, waterfalls, city street, office; object arrangements: vegetables, cookies, bathroom things, Lego). (*b*) Experimental design of the encoding phase, recall phase and surprise test. (Online version in colour.)

reactivation during retrieval [29,30]. Extending such findings, it has been demonstrated that the likelihood of successful remembering increases when gaze locations during recall are directly manipulated to overlap with those from encoding [31–35]. While these findings clearly show that where you look has important consequences for what you remember, they are all related to *static gaze patterns in space* and do not take into account how *dynamic sequences* of connected eye-movements progress over space and time. Also, active manipulations of gaze direction eliminate the spatio-temporal dynamics that are essential properties of free viewing [36]. Although research on *recognition* memory has shown that a temporal reinstatement of gaze patterns can be beneficial when discriminating between novel and previously encountered stimuli [37,38] (but see [39]), there is to date no direct evidence for the influential claim that the sequential *replay* of eye movements serves to facilitate pure episodic reconstruction in the absence of visual input. To address this fundamental question, it is necessary to go beyond static gaze patterns in space and examine how the sequential replay of connected eye movements unfolding over time—*scanpaths*—may support episodic remembering in a free recall task.

In the present study, we tackle this issue head on by using state-of-the-art scanpath similarity techniques, capable of quantifying the sequential encoding-recollection similarity (SERS) of scanpaths, and decomposing it into distinct spatio-temporal scanpath properties. The degree of SERS can then be examined over these scanpath properties during episodic retrieval, allowing us to examine and stratify the critical concept of scanpath replay.

Participants encode images with associated verbal labels and subsequently recall each image while looking at a blank screen. In the recall phase, a particular image is cued by its verbal label, and the quality of the mnemonic content of each episodic reconstruction is subjectively rated with respect to vividness, spatial accuracy and overall recollection strength (figure 1*b*). To objectively assess the strength and quality of the mnemonic content during recall, our paradigm further combines those subjective ratings of memory quality with a subsequent surprise forced-choice recognition test (figure 1*b*). Finally, we use two types of image stimuli—scenes and object arrangements—to investigate if the SERS of scanpaths differentially contributes depending on the nature of the goal-relevant memory. The scenes comprise highly predictable spatial relations between meaningful scene elements (due to schematic and situational knowledge for the common scenarios), with relatively low demands on relational memory. The object arrangements comprise arbitrary spatial relations between their scene elements (objects), with relatively high demands on relational memory (figure 1*a*). These two types of stimuli allow us to examine the relative contribution of different scanpath reinstatements (such as fixation order and saccade direction) as a function of varying demands on relational memory [22].

2. Methods

(a) Participants

Sixty-two healthy adults participated in the experiment. Two were removed owing to extensive data loss and technical problems,

Table 1. Mean values for the performance data during the recall phase and the surprise test, with standard deviations within brackets. The mnemonic content score represents the mean of the three subjective ratings (recollection strength, vividness, spatial accuracy). Only correct trials were considered for confidence, response time and gaze transitions between options (a measure of choice certainty [45]).

	total	scenes	object arrangements
performance data			
recall phase			
mnemonic content score (%)	58 (30)	66 (26)	49 (31)
surprise test			
accuracy (%)	85 (36)	95 (22)	75 (43)
response time (ms)	7699 (6746)	5583 (5672)	9815 (7065)
confidence (%)	78 (29)	90 (18)	66 (33)
gaze transitions between options	10.8 (7.4)	8.5 (5.2)	13.2 (8.4)

leaving sixty participants (34 females; mean age 25.2, s.d. 6.4). Power analyses [40] for bivariate measures of association indicated that our sample size should be sufficient to detect a large correlation ($r = 0.5$) and difference between means ($d = 0.6$) with 90% power ($\alpha = 0.05$), which is in line with prior work on gaze reinstatement [17,27] and scanpath similarity [41]. All participants were fluent in Swedish and had normal or corrected-to-normal vision. Participants gave written informed consent and were compensated with a cinema voucher. All methods were conducted in accordance with the Swedish Act concerning the Ethical Review of Research involving Humans (2003:460) and the Code of Ethics of the World Medical Association (Declaration of Helsinki).

(b) Data acquisition and materials

Gaze data were recorded from both eyes individually, using a SensoMotoric Instruments (SMI RED-M) eye tracker, running iView X 2.7 software and sampling at 120 Hz. A Dell Optiplex 755 PC presented stimuli using PsychoPy [42] on a 22-inch monitor with a resolution of 1680 × 1050 pixels. Participants were seated with their heads in a chin and forehead rest, 65 cm away from the monitor. Calibration and validation of gaze data was conducted prior to each participant's experimental session. Fixations were detected with the I2MC algorithm [43]. Stimuli comprised 18 scenes and 18 object arrangements that were presented in a frame that covered 80% of the monitor (figure 1*b*). See electronic supplementary material for data quality and general oculomotor data during encoding and recall. The scenes were taken from the MIT test set Cat2000 (http://saliency.mit.edu/datasets.html) [44] and the object arrangements were photographs shot for this particular study (available at https://osf.io/d9zng/).

The content of the scene images comprised an equal amount of indoor and outdoor global scenes, where the spatial relations among scene elements were highly congruent in respect to the image's semantic structure. The content of the object arrangement images comprised four separate objects from the same category, where the spatial relations between the objects were completely arbitrary in respect to the image's semantic structure. The 18 object arrangements all comprised unique spatial configurations (figure 1*a*; complete set of images available at https://osf.io/d9zng/). By contrast to the scene images, we thus had full experimental control over the spatial relations among individual scene elements (objects) in the object arrangement images.

(c) Design and procedure

The experiment was divided into three phases: encoding, recall and surprise test (figure 1*b*). To conceal the true objective of the study, participants were told that the experiment concerned pupil dilation in relation to mental workload, and it was explained that their eyes were filmed for this matter.

(d) Encoding phase

Participants encoded 36 images accompanied by a verbal label that preceded each image. The verbal label described the semantic content of the succeeding image (e.g. 'studio', 'vegetables'). The images comprised two different types: scenes ($n = 18$) and object arrangements ($n = 18$) and were each presented for 5 s in randomized order. Participants were instructed to memorize each image as thoroughly as possible. When the encoding phase finished, participants engaged in a distracter task, where they were to count backward in steps of 3 from a randomly generated three-digit number for 12 s.

(e) Recall phase

Participants recalled all 36 encoded images while looking at a blank screen. Participants were cued by the associated verbal label (preceding the blank screen) in a randomized order and were then instructed to recall and visualize the corresponding image in as much detail as possible while looking at the blank screen. The blank screen remained for 7 s (compared to encoding, this time was increased to compensate for the increased fixation durations that occur when looking at 'nothing' [13]). After each recollection, participants were to rate the quality of their recollection based on overall recollection strength, vividness and spatial accuracy. Overall recollection strength was rated on a scale 0–100% in relation to the statement 'It was easy for me to remember the image'. Vividness was rated on a scale 0–100% in relation to the statement 'My mental image was clear, vivid and detailed—almost as if I could see the image in front of me' Spatial accuracy was rated on a scale of 0–100% in relation to the statement 'I could indicate, with high spatial accuracy, where different objects/scene elements were located in the image' (table 1). The three scores were highly correlated ($r > 0.89$, $p < 0.001$; see electronic supplementary material, figure S1) and based on a mean of those three ratings, a mnemonic content score was calculated for each image and participant.

(f) Surprise test

Participants completed four-alternative forced-choice (4AFC) recognition tests that covered all 36 images presented during encoding. Participants were not aware that they would engage in this test beforehand. The target image and three different distracter images were presented in the four quadrants of the screen, and participants were instructed to select the image they had encoded and recalled as accurately and quickly as possible.

After each selection, participants rated how confident (0–100%) they were in their choice. For the scenes, the distracter images comprised: (1) a horizontal flip of the target image; (2) a similar lure image; and (3) a horizontal flip of the lure image. For the object arrangements, the distracter images comprised: (1) a lure image where two of the original objects had switched locations; (2) a lure image where one of the objects had been exchanged with another semantically congruent object; and (3) a lure image where one of the objects had been exchanged with the other semantically congruent object, and also the location of this new object had been switched with one of the other objects.

The present experimental design allowed us to collect *spontaneous eye movements* during recall, and to use the surprise test to evaluate the validity of the subjective ratings as a measure of recall success.

(g) Statistical analyses

All data were analysed using generalized linear mixed-effects models (GLMMs; glmer of package lme4 [46]) and linear mixed-effects models (LMEMs), where participants and images were modelled as random effects (intercepts) and with random slopes for image types. In order to describe the model-fit of an independent variable, the deviance of the proposed model was contrasted against an unconditional null model, including only the intercept and the random factors. When building models with several independent variables, we used a backward selection approach, starting with a maximal model, which included all variables and their interactions. We then used likelihood-ratio tests to compare models and then step-by-step removed non-significant effects until no further model changes resulted in a significant likelihood-ratio test ($p < 0.05$). Models were fitted with restricted maximum-likelihood (REML) and Satterthwaite approximations were used to assess the significance of individual predictors.

3. Results

(a) Behavioural results

We first sought to verify that the object arrangements were more demanding to recall than the scenes. As expected, mnemonic content was rated higher for scenes than object arrangements during recall, $\beta = 0.165$, (s.e. = 0.0317), $z = 5.22$, $p < 0.001$. For the surprise test, results revealed higher performance for scenes over four different measures of retrieval performance: response accuracy ($\chi^2_1 = 64.8$, $\beta = 2.17$, s.e. = 0.269, $z = 8.05$, $p < 0.001$), confidence ($\chi^2_1 = 51.0$, $\beta = 0.175$, s.e. = 0.0245, $z = 7.14$, $p < 0.001$), response time ($\chi^2_1 = 43.7$, $\beta = 4132$, s.e. = 625, $z = 6.61$, $p < 0.001$) and gaze transitions between options (a measure of choice certainty [45]; $\chi^2_1 = 53.6$, $\beta = 4.62$, s.e. = 0.631, $z = 7.32$, $p < 0.001$). Thus, these results demonstrate that the object arrangements, placing greater demand on relational memory, were indeed harder to recall than the scenes (table 1).

We next tested if the mnemonic content score during recall predicted subsequent retrieval performance during the surprise test. Results reveal that the mnemonic content score was a significant predictor of all four measures of retrieval performance: response accuracy ($\chi^2_1 = 6.17$, $\beta = 0.645$, s.e. = 0.260, $z = 2.48$, $p = 0.013$), confidence ($\chi^2_1 = 27.4$, $\beta = 0.101$, s.e. = 0.0192, $z = 5.24$, $p < 0.001$), response time ($\chi^2_1 = 6.53$, $\beta = -1339$, s.e. = 524, $z = -2.55$, $p = 0.011$) and choice certainty ($\chi^2_1 = 11.8$, $\beta = -1.98$, s.e. = 0.575, $z = -3.44$, $p < 0.001$; see electronic supplementary material, figure S2A–D). Higher subjective ratings during the recall phase thus predicted more accurate, confident, faster and certain responses during the subsequent surprise test. We therefore conclude that the mnemonic content score during recall is a representative index of the memory quality during episodic remembering.

(b) Spatial reinstatement of static gaze patterns

In the next step, we examined the spatial encoding-recollection overlap of omnibus static position-based gaze, i.e. without considering the dynamic path that connects a sequence of eye movements unfolding over time. Here, we divided the screen into four areas of interest (AOIs), corresponding to the four quadrants of the screen, and then the proportional number of fixations within each of those four AOIs was compared over encoding and recall, for each image and participant. We found that the spatial locations of participants' gaze patterns broadly overlapped during encoding and recall, and this effect was pronounced for the more demanding object arrangements (electronic supplementary material, figure S3). Importantly, we also found that the degree of this spatial gaze reinstatement was predictive of our index of memory quality during recall (electronic supplementary material, figure S4). This corroborates what previous research has found [14,27,28], and provides further evidence that a spatial encoding-recollection overlap in viewing patterns supports episodic remembering. An exhaustive presentation of these analyses is available in the electronic *supplementary material* (where we also conduct the analyses for 8, 12 and 16 AOIs, to test the sensitivity of these analyses in respect to 'spatial resolution').

(c) Sequential reinstatement of consecutive eye movements

To achieve the central goals of the present study, we next investigated SERS in scanpaths, and how the reinstatement of such ordered sequences of eye movements—scanpath replay—may support episodic remembering.

To measure SERS, we used MultiMatch (MM), which was introduced as a method for comparing scanpaths over different spatio-temporal dimensions [47], and has subsequently been validated, applied, and evaluated against other scanpath comparison tools [36,47,48,49]. The basic principle is that the MM algorithm simplifies the two scanpaths under comparison (in this case, one from encoding and the other from recall) into virtualized ordered sequences of connected saccadic vectors (see [36,47] for a full explanation of the scanpath simplification procedure). The MM algorithm then temporally aligns the two scanpaths, so that particular saccade and fixation pairings from the two ordered scanpaths can be compared. The temporal alignment is achieved by matching the sequence of saccadic vectors from the two scanpaths according to their vector shape (length and direction) using a comparison matrix in which costs are drawn from differences in shape similarity between potential pairings. The MM algorithm then uses the Dijkstra [50] algorithm to find the shortest (minimal cost) path through the comparison matrix (from the top left corner to the bottom right—taking all possible routes into account). Now the two scanpaths can be sequentially aligned according to this path. Temporal alignment is thus a relative procedure which is not dependent on an equal number of saccades in the two scanpaths (or on equal encoding and recall times). The process of temporal

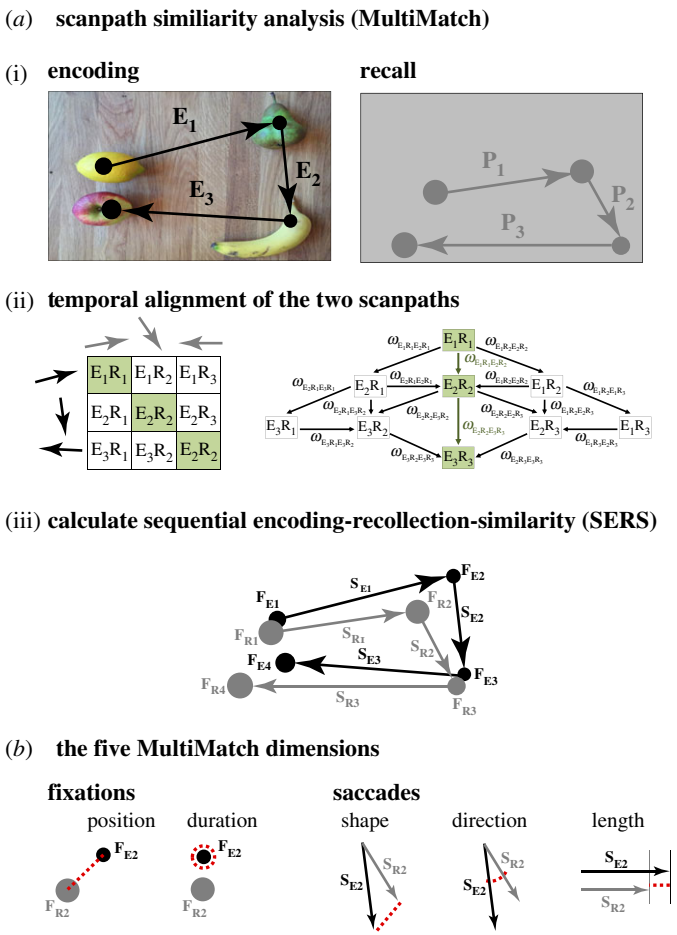

Figure 2. Illustrations of the method to capture sequential encoding-recollection similarity. (*a*) Overview of the MultiMatch scanpath similarity analysis. In the first panel (i), the scanpaths from encoding and recall to be compared are shown, where fixations are represented as dots and saccades as arrows between the fixations, and where larger dots represent longer fixation duration. In the second panel (ii), the basic principle behind the temporal alignment of the two scanpaths is illustrated. In the matrix to the left, each saccadic vector during encoding (E1–E3) is compared to each saccadic vector during recall (R1–R3) according to their shape. Using the Dijkstra algorithm [50], the optimal temporal alignment between the two scanpaths is then computed as the minimal cost—shortest path—from the upper left corner to the bottom right corner of the comparison matrix. All possible paths along with the cost for each transition (ω) between the matrix elements is outlined in the right figure. The minimal cost—shortest path—in this example (E_1R_1 to E_2R_2 to E_3R_3) is highlighted. Finally, the temporally aligned scanpaths from encoding and recall are shown superimposed in Euclidean space (iii) to illustrate how sequential encoding-recollection similarity (SERS) can be calculated for each individual fixation and saccade pairings over the five MultiMatch-dimensions. SERS for the complete scanpaths is then quantified as the average similarity over all the temporally aligned saccade (encoding: E_{S1}–E_{S4}; recall: R_{S1}–R_{S4}) or fixation (encoding: E_{F1}–E_{F4}; recall: R_{F1}–R_{F4}) pairings. (*b*) An illustration of the five MulitMatch dimensions of fixation position, fixation duration, saccade shape, saccade direction and saccade length for the temporally aligned fixation pairs $F_{E3} - F_{R3}$ and saccade pairs $S_{E3} - S_{R3}$. The numeric difference in SERS between each dimension is illustrated with a dotted line for each dimension separately. The fixation dimension of position relies on spatial coordinates in absolute space and quantifies how similar temporally aligned fixations are in respect to Euclidean distances, thus representing a similarity measure of fixation order. In contrast, the saccade dimensions of shape, direction and length rely on differences in relative space. The shape dimension quantifies how similar temporally aligned saccadic vectors are in overall geometric shape. The direction dimension quantifies how similar temporally aligned saccadic vectors are in geometric angle, thus representing a similarity measure of the particular heading of eye movements. The length dimension quantifies how similar temporally aligned saccades are in their absolute amplitude, irrespective of shape and direction. The fixation dimension duration does not rely on any spatial coordinates and quantifies how similar temporally aligned fixations are in their duration. (*c*) Examples of varying SERS: (i) complete SERS in respect to all five MM dimensions; (ii) relatively high SERS in shape, but during recall there is a dislocation in absolute space, large dissimilarities in saccadic angles, overall shorter saccades and overall longer fixation durations—therefore the SERS in all other MM dimensions are relatively low; (iii) high SERS in direction, but during recall there are dislocations in absolute space, disproportional saccadic lengths and overall longer fixation durations—therefore the SERS in all other MM dimensions are relatively low; (iv) low SERS over all five MM dimensions. (Online version in colour.)

alignment is indicated in figure 2*a* and is explained in full in [47]. SERS between the temporally aligned scanpaths can then be separately determined over five MM dimensions: (1) fixation position, (2) fixation duration, (3) saccade shape, (4) saccade direction and (5) saccade length (figure 2*b*). The five MM dimensions of SERS thus reflect different embodied connections between eye movements and memory, allowing us to not only investigate *if* a sequential reinstatement of consecutive eye movements supports episodic remembering, but also to scrutinize *how* qualitatively distinct aspects of such scanpaths influence this interplay—such as the sequential reinstatement of fixation order (MM dimension: fixation position) and pathway from one consecutive fixation to the next (MM dimension: saccade direction). As most of the MM dimensions do not rely on a reference frame in absolute space, the method is relatively unaffected by the documented spatial 'offsets' and 're-scaling' of gaze patterns during recall [13,15–17,29], which are known to dilute the results when investigating scanpaths during a blank screen paradigm [36,47].

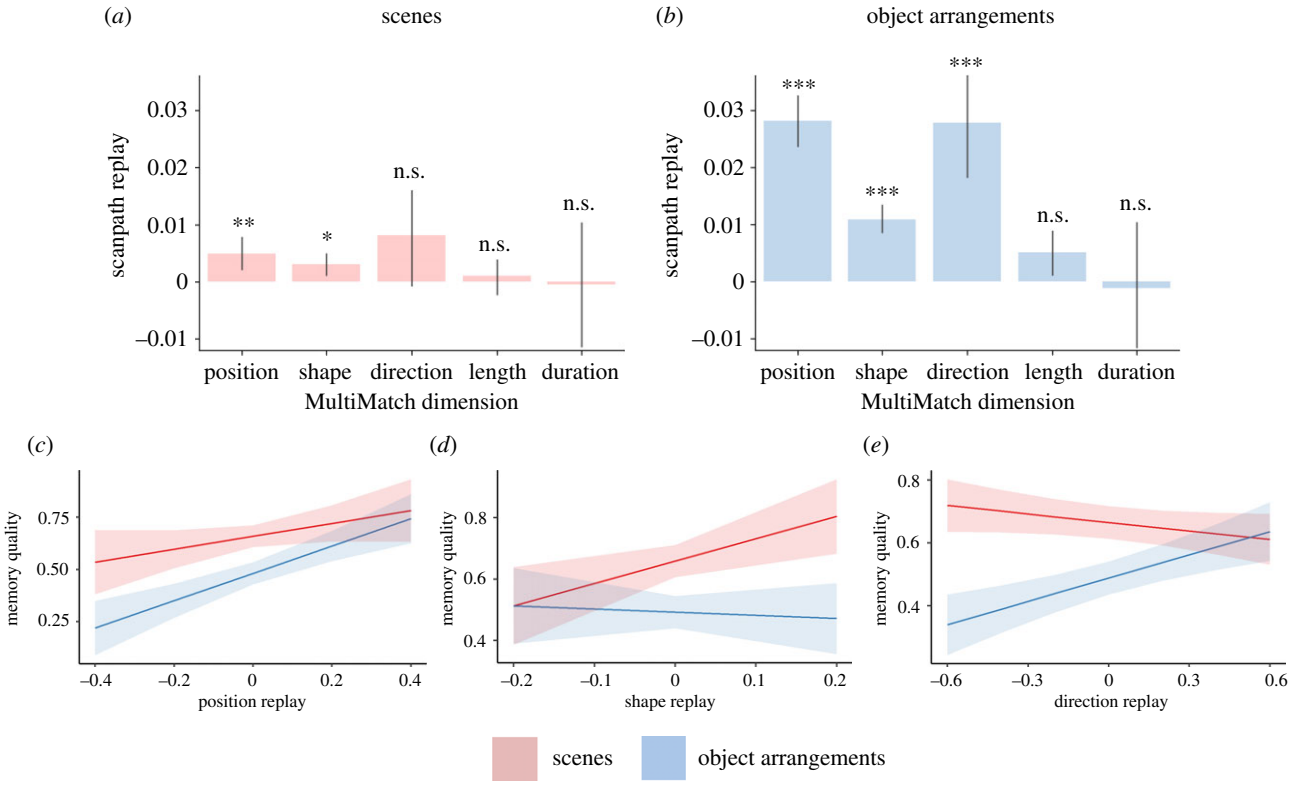

Figure 3. Scanpath replay over the five MM dimensions for the (*a*) scenes and (*b*) object arrangements. The measure of scanpath replay represents the difference between SERS and baseline similarity (a value greater than zero indicates scanpath replay). (*c*) The relationship between position replay and memory quality. (*d*) The relationship between shape replay and memory quality. (*e*) The relationship between direction replay and memory quality. Memory quality corresponds to the mnemonic content score during recall. Error bars and shaded areas denote 95% confidence intervals, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. (Online version in colour.)

To determine if and how scanpaths from encoding become sequentially reinstated during recall, we first calculated SERS scores over all five MM dimensions (fixation: position, duration; saccades; shape, direction, length) by comparing the scanpaths produced by a participant during the encoding of an image, with those produced during recall of the same image. figure 2*c* for examples of varying SERS over the five MM dimensions. Second, by comparing the scanpaths produced by a participant during the encoding of an image with those produced during recall of all the other images from the same type (scenes, object arrangements), and then averaging across the scores calculated for all images, we acquired a baseline similarity score. This represents the average sequential similarity in scanpaths for each participant across images.

For the scenes, results revealed that the SERS score was significantly higher than the baseline similarity score for position ($\chi^2_1 = 6.83$, $\beta = 0.0048$, s.e. = 0.0018, $z = 2.61$, $p = 0.009$) and shape ($\chi^2_1 = 5.70$, $\beta = 0.0030$, s.e. = 0.0013, $z = 2.39$, $p = 0.017$). For the object arrangements, results revealed that the SERS score was significantly higher than the baseline similarity score for position ($\chi^2_1 = 148$, $\beta = 0.0278$, s.e. = 0.0023, $z = 12.2$, $p < 0.001$), shape ($\chi^2_1 = 55.8$, $\beta = 0.0106$, s.e. = 0.0014, $z = 7.47$, $p < 0.001$) and direction ($\chi^2_1 = 33.0$, $\beta = 0.0269$, s.e. = 0.0047, $z = 5.74$, $p < 0.001$). Length and duration were not significant predictors for either scenes or object arrangements ($\chi^2_1 < 2.30$, p-values > 0.13; figure 3*a*,*b*). These results thus show that position, shape and direction are scanpath properties that were sequentially reinstated during recall. However, significant SERS in direction was only observed for the object arrangements. The absence of SERS in length and duration demonstrates that amplitude of saccades (i.e. independent of direction and shape), and fixation durations are not sequentially retained during recall. However, as these properties typically depend on visual features from the image [51,52], which are absent during recall, there is no particular reason to expect SERS for either. Moreover, it is well known that fixation durations become atypically long when there are no external features to look at [13] and that saccades are frequently 'contracted' in size during recall [13,15,17,29]. See electronic supplementary material (electronic supplementary material, figure S5) for follow-up analyses that validate our methods of capturing sequential reinstatement and that scrutinize the results in relation to idiosyncratic behaviours.

(d) Scanpath replay and episodic remembering

In the final step, we then tested whether the degree of scanpath replay for the significant MM dimensions (position, shape and direction) predicted our index of memory quality during the recall phase (the mnemonic content score). To quantify the degree of scanpath replay, we used a similar rationale as Wynn *et al.* [53], and subtracted the baseline similarity score from the SERS score for each MM dimension over each participant and image. The best model fit of memory quality revealed significant effects of image type ($\chi^2_1 = 26.48$, $\beta = 0.169$, s.e. = 0.033, $z = 26.11$, $p < 0.001$), position replay ($\chi^2_1 = 16.95$, $\beta = 0.472$, s.e. = 0.115, $z = 4.12$, $p < 0.001$), direction replay ($\chi^2_1 = 4.04$, $\beta = 0.085$, s.e. = 0.042, $z = 2.01$, $p = 0.045$) and significant interaction effects between image type and shape replay ($\chi^2_1 = 5.28$, $\beta = 0.896$, s.e. = 0.390, $z = 2.30$, $p = 0.022$), and between image type and direction replay ($\chi^2_1 = 16.05$, $\beta = 0.337$, s.e. = 0.084, $z = 4.01$, $p < 0.001$; see electronic supplementary material, table S2). Greater position

replay predicted better memory quality independent of image type (figure 3*c*), whereas greater shape replay only predicted better memory quality for the scenes (figure 3*d*), and greater direction replay only predicted better memory quality for the object arrangements (figure 3*e*). See electronic supplementary material (electronic supplementary material, table S4, figure S6) for a follow-up analysis testing the specificity of this replay effect in relation to the temporal order of eye movements.

The results thus show that independent of the mnemonic content, looking at locations in a similar sequential order (i.e. position replay) during recall as during the original encoding predicts the quality of episodic remembering. This extends previous findings that encoding-recollection overlaps in fixation locations are relevant for episodic remembering [14,27,28] and indicates that the sequential order in which such fixations unfold also plays an active role in the recollection process. The interaction effects show that shape and direction replay contributed differently depending on the mnemonic content. Retaining the general spatial structure of consecutive saccades (i.e. shape replay) appears to be more important when reconstructing the more predictable spatial relations in the scene images. Conversely, retaining the particular heading of consecutive saccades (i.e. direction replay) appears to be more important when reconstructing the arbitrary spatial relations in the object arrangements.

(e) Facilitating remembering or a consequence of the memory representation?

While the results indicate a functional role of scanpath replay in episodic reconstruction, an alternative explanation is that greater replay is a consequence of having a strong memory representation. To disambiguate between those two explanations, we sought to determine whether scanpath replay depends on how well the images are initially encoded into memory. Previous research has established that the cumulative number of fixations during encoding is a reliable index of how well a visual stimulus is encoded into memory [6–12]. Thus, to account for effects of encoding strength, we added the cumulative number of study fixations on each image into our model. The model was significantly improved ($\chi^2_1 = 24.90$, $p < 0.001$). All previous effects and interaction effects were significant also in this model, and as expected, the cumulative number of study fixations significantly predicted memory quality during recall ($\beta = 0.008$, s.e. = 0.002, $z = 4.99$, $p < 0.001$), which confirms the validity of this measure as an index of encoding strength. Interactions of number of study fixations and the other predictors (position replay, shape replay, direction replay, image type), were then stepwise added to the model. Importantly, none of those interactions significantly improved the model (p-values > 0.19). See electronic supplementary material, table S3. Thus, the reported effects of scanpath replay upon memory quality do not reduce to how well the images were originally encoded into memory. Instead, our results substantiate the idea that the replay of eye movements supports the reconstruction of inter-related and task-relevant mnemonic features into a spatio-temporal context.

4. Discussion

Corroborating previous research [14,27,28,31,32], we found that spatial reinstatement of gaze patterns predicts the fidelity of episodic remembering. State-of-the-art scanpath-similarity techniques [36] allowed us to move beyond omnibus static position-based gaze to examine sequential reinstatement of qualitatively different scanpath properties during episodic reconstructions. Of central importance, we provide evidence that the reinstatement of an ordered sequence of eye movements supports episodic remembering and that different spatio-temporal properties of the unfolding scanpaths differentially contribute depending on the nature of the goal-relevant memory. Finally, we provide evidence that these sequential reinstatement effects on episodic remembering are not simply a consequence of how well the event was originally encoded.

The claim that gaze behaviour supports episodic recollection dates back to (at least) Ulric Neisser and Donald Hebb. Neisser argued that eye movements are actively associated with memory reconstruction [24], and Hebb claimed that eye movements are necessary to assemble and organize 'part images' into a whole visualized image [23]. Such ideas were further developed in Noton and Stark's Scanpath Theory, which states that memories are stored in a 'feature ring', comprising visual features and the sequence of eye movements (scanpaths) linking them together [25,26]. Scanpath Theory holds that (a) eye movements during recall play out in the same sequential order as during encoding, and (b) such scanpath replay serves a functional role in episodic remembering. While a strong interpretation of Scanpath Theory, where memory is accompanied by an exact and full scanpath recapitulation, has been refuted in more recent research [17,22,39], there is extensive evidence that episodic remembering involves eye movements that broadly reproduce gaze patterns at encoding [13–17,27,29]. Thus, the first tenet of scanpath theory, that episodic remembering is accompanied by scanpath replay, is established in the literature, and gets further explicated in the present study. However, there is to date virtually no evidence for the second tenet, that scanpath replay serves a functional role when episodic information is recalled from memory in the absence of supporting visual input. Recent work has demonstrated that gaze reinstatements promote cortical reconstruction [29,30] and successful remembering [14,27,28,31–35]. However, these findings are all related to spatial reinstatement of static position-based gaze and shed no light upon the proposition that the sequential reinstatement of consecutive eye movements serve to reconstruct and bind spatio-temporal information into a coherent memory (but see ref. [54] for different temporal gaze dynamics during perception and mental imagery). Interesting recent research on recognition memory shows that a temporal replay of gaze patterns can support recognition performance [37,38,53]. Still, recognition tasks involve situations where all (or some) encoded information is available as a 'copy cue', and where visual information accumulates until sufficient evidence is available to solve the old/new-discrimination task. The purpose of visual exploration under such conditions is thus considerably different from a recall task, where the rememberer needs to mentally reconstruct the complete spatio-temporal properties of a prior event without supporting visual input (see ref. [6]). Here, we present direct evidence that a replay of sequentially ordered eye movements plays an active role in episodic reconstruction in a pure recall task.

Recently, Wynn and colleagues proposed that eye movements support active memory retrieval by broadly reinstating the spatio-temporal context based on current task demands

and available cognitive resources [22]. The idea that the facilitatory role of gaze reinstatements increases with task demands has been supported by their studies on older adults [28] and task difficulty [53]. Further support comes from the present findings where more extensive spatial and sequential gaze reinstatement was seen during the reconstruction of the more demanding object arrangements as compared to the less demanding scenes, and that different spatio-temporal scanpath properties contributed differentially to successful recollection of those two types of images.

A vast amount of research has demonstrated that the hippocampus plays a critical role in binding information in space and time during memory formation, and it is assumed that the hippocampus stores an event index, pointing to the cortical representations of each aspect of the whole event (for a review see [55]). Remembering depends on a retrieval cue matching the hippocampal index, which in turn triggers pattern completion and the reactivation of the cortical traces (and thus the experience of remembering the original event). Based on this cortical-hippocampal interplay, a prominent view holds that the reconstruction of past events fundamentally relies on mental simulations that reinstate approximations of the sensorimotor processes that characterized the original event [56,57]. Successful episodic remembering is considered to depend on the overlap between available retrieval cues and stored memory traces [58]. Compatibility between the processes triggered by a retrieval cue and those engaged during encoding increases the likelihood of successful retrieval [58,59]. As previous research has demonstrated that gaze behaviour during encoding and retrieval are linked to activity in the hippocampus [3,60–65], and that gaze reinstatement correlates with neural reactivation [29,30], it is conceivable that the scanpath replay reported here is responsible for promoting cortical-hippocampal reactivation of visuospatial relations during episodic memory reconstruction. The successive reinstatement of consecutive eye movements may act as internally generated retrieval cues that continuously updates in an iterative fashion as the rememberer moves his/her eyes during the pattern completion process. In situations with high demands on relational memory, as for the recollection of the object arrangements, more specified retrieval cues would be required during the pattern completion process, as compared to the less demanding scenes, where the pattern completion process would be supported by more predictable scene semantics.

5. Conclusion

Spatial reinstatement of gaze patterns has proven important for episodic remembering. The present study extends this finding and provides direct evidence that the actual replay of an ordered sequence of eye movements unfolding over time facilitates episodic remembering, and that specific spatio-temporal scanpath properties differentially contribute depending on the nature of the goal-relevant memory.

Ethics. Participants gave written informed consents prior to participation. All methods were conducted in accordance with the Swedish Act concerning the Ethical Review of Research involving Humans (2003:460) and the Code of Ethics of the World Medical Association (Declaration of Helsinki). As established by Swedish authorities and specified in the Swedish Act concerning the Ethical Review of Research involving Humans (2003:460), the present study does not require specific ethical review by the Swedish Ethical Review Authority due to the following reasons: (1) it does not deal with sensitive personal data, (2) it does not use methods that involve a physical intervention, (3) it does not use methods that pose a risk of mental or physical harm, (4) it does not study biological material taken from a living or dead human that can be traced back to that person.

Data accessibility. Data and materials are available on the OSF: https://osf.io/d9zng/.

The data are provided in electronic supplementary material [66].

Authors' contributions. R.J.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, software, supervision, validation, visualization, writing—original draft, writing—review and editing; M.N.: data curation, formal analysis, methodology, resources, software, writing—review and editing; R.D.: investigation, writing—review and editing; M.J.: conceptualization, supervision, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. The authors declare no competing interests.

Funding. The work was funded by the Swedish Research Council grant no. 2015-01206 (Roger Johansson) and the Marcus and Amalia Wallenberg Foundation award MAW2015.0043 (Mikael Johansson).

Acknowledgements. We gratefully acknowledge Lund University Humanities Laboratory.

References

1. Tulving E. 1983 *Elements of episodic memory*. New York, NY: Oxford University Press.
2. Voss JL, Bridge DJ, Cohen NJ, Walker JA. 2017 A closer look at the hippocampus and memory. *Trends Cogn. Sci.* **21**, 577–588. (doi:10.1016/j.tics.2017.05.008)
3. Ryan JD, Shen K, Liu ZX. 2020 The intersection between the oculomotor and hippocampal memory systems: empirical developments and clinical implications. *Ann. NY Acad. Sci.* **1464**, 115–141. (doi:10.1111/nyas.14256)
4. Findlay JM, Gilchrist ID. 2003 *Active vision: the psychology of looking and seeing*. Oxford, UK: Oxford University Press.
5. Henderson JM. 2003 Human gaze control during real-world scene perception. *Trends Cogn. Sci.* **7**, 498–504. (doi:10.1016/j.tics.2003.09.006)
6. Damiano C, Walther DB. 2019 Distinct roles of eye movements during memory encoding and retrieval. *Cognition* **184**, 119–129. (doi:10.1016/j.cognition.2018.12.014)
7. Fehlmann B, Coynel D, Schicktanz N, Milnik A, Gschwind L, Hofmann P, Papassotiropoulos A, de Quervain DJ. 2020 Visual exploration at higher fixation frequency increases subsequent memory recall. *Cereb. Cortex Commun.* **1**, tgaa032. (doi:10.1093/texcom/tgaa032)
8. Kafkas A, Montaldi D. 2011 Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *Q. J. Exp. Psychol.* **64**, 1971–1989. (doi:10.1080/17470218.2011.588335)
9. Liu ZX, Shen K, Olsen RK, Ryan JD. 2017 Visual sampling predicts hippocampal activity. *J. Neurosci.* **37**, 599–609. (doi:10.1523/JNEUROSCI.2610-16.2016)
10. Loftus GR. 1972 Eye fixations and recognition memory for pictures. *Cogn. Psychol.* **3**, 525–551. (doi:10.1016/0010-0285(72)90021-7)
11. Molitor RJ, Ko PC, Hussey EP, Ally BA. 2014 Memory-related eye movements challenge behavioral measures of pattern completion and pattern separation. *Hippocampus* **24**, 666–672. (doi:10.1002/hipo.22256)
12. Olsen RK, Sebanayagam V, Lee Y, Moscovitch M, Grady CL, Rosenbaum RS, Ryan JD. 2016 The relationship between eye movements and subsequent recognition: evidence from individual

differences and amnesia. *Cortex* **85**, 182–193. (doi:10.1016/j.cortex.2016.10.007)

13. Brandt SA, Stark LW. 1997 Spontaneous eye movements during visual imagery reflect the content of the visual scene. *J. Cogn. Neurosci.* **9**, 27–38. (doi:10.1162/jocn.1997.9.1.27)
14. Laeng B, Teodorescu D-S. 2002 Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cogn. Sci.* **26**, 207–231. (doi:10.1207/s15516709cog2602_3)
15. Gbadamosi J, Zangemeister WH. 2001 Visual imagery in hemianopic patients. *J. Cogn. Neurosci.* **13**, 855–866. (doi:10.1162/089892901753165782)
16. Johansson R, Holsanova J, Holmqvist K. 2006 Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cogn. Sci* **30**, 1053–1079. (doi:10.1207/s15516709cog0000_86)
17. Johansson R, Holsanova J, Dewhurst R, Holmqvist K. 2012 Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 1289–1314. (doi:10.1037/a0026585)
18. Richardson DC, Spivey MJ. 2000 Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition* **76**, 269–295. (doi:10.1016/S0010-0277(00)00084-6)
19. Ryan JD, Althoff RR, Whitlow S, Cohen NJ. 2000 Amnesia is a deficit in relational memory. *Psychol. Sci.* **11**, 454–461. (doi:10.1111/1467-9280.00288)
20. Ferreira F, Apel A, Henderson JM. 2008 Taking a new look at looking at nothing. *Trends Cogn. Sci.* **12**, 405–410. (doi:10.1016/j.tics.2008.07.007)
21. Richardson DC, Altmann GTM, Spivey MJ, Hoover MA. 2009 Much ado about eye movements to nothing: a response to Ferreira *et al.*: Taking a new look at looking at nothing. *Trends Cogn. Sci.* **13**, 235–236. (doi:10.1016/j.tics.2009.02.006)
22. Wynn JS, Shen K, Ryan JD. 2019 Eye movements actively reinstate spatiotemporal mnemonic content. *Vision* **3**, 21. (doi:10.3390/vision3020021)
23. Hebb DO. 1968 Concerning imagery. *Psychol. Rev.* **75**, 466–477. (doi:10.1037/h0026771)
24. Neisser U. 1967 *Cognitive psychology*. New York, NY: Appleton-Century-Crofts.
25. Noton D, Stark LW. 1971 Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vis. Res.* **11**, 929–942. (doi:10.1016/0042-6989(71)90213-6)
26. Noton D, Stark LW. 1971 Scanpaths in eye movements during perception. *Science* **171**, 308–311. (doi:10.1126/science.171.3968.308)
27. Laeng B, Bloem IM, D'Ascenzo S, Tommasi L. 2014 Scrutinizing visual images: the role of gaze in mental imagery and memory. *Cognition* **131**, 263–283. (doi:10.1016/j.cognition.2014.01.003)
28. Wynn JS, Olsen RK, Binns MA, Buchsbaum BR, Ryan JD. 2018 Fixation reinstatement supports visuospatial memory in older adults. *J. Exp. Psychol. Hum. Percept. Perform.* **44**, 1119–1127. (doi:10.1037/xhp0000522)
29. Bone MB, St-Laurent M, Dang C, McQuiggan DA, Ryan JD, Buchsbaum BR. 2018 Eye movement reinstatement and neural reactivation during mental imagery. *Cereb. Cortex* **29**, 1075–1089. (doi:10.1093/cercor/bhy014)
30. Johansson R, Bramao I, Dewhurst R, Johansson M. 2018 *Neurocognitive bases for the functional role of gaze direction during episodic memory retrieval*. In *The 25th Annual Meeting of the Cognitive Neuroscience Society, Boston, US, 24–27 March*.
31. Johansson R, Johansson M. 2014 Look here, eye movements play a functional role in memory retrieval. *Psychol. Sci.* **25**, 236–242. (doi:10.1177/0956797613498260)
32. Johansson R, Johansson M. 2020 Gaze position regulates memory accessibility during competitive memory retrieval. *Cognition* **197**, 104169. (doi:10.1016/j.cognition.2019.104169)
33. Scholz A, Mehlhorn K, Krems JF. 2016 Listen up, eye movements play a role in verbal memory retrieval. *Psychol. Res.* **80**, 149–158. (doi:10.1007/s00426-014-0639-4)
34. Scholz A, Klichowicz A, Krems JF. 2018 Covert shifts of attention can account for the functional role of 'eye movements to nothing'. *Mem. Cogn.* **46**, 230–243. (doi:10.3758/s13421-017-0760-x)
35. Van Ede F, Chekroud SR, Nobre AC. 2019 Human gaze tracks attentional focusing in memorized visual space. *Nat. Hum. Behav.* **3**, 462–470. (doi:10.1038/s41562-019-0549-y)
36. Dewhurst R, Nyström M, Jarodzka H, Foulsham T, Johansson R, Holmqvist K. 2012 It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behav. Res. Methods* **44**, 1079–1100. (doi:10.3758/s13428-012-0212-2)
37. Bochynska A, Laeng B. 2015 Tracking down the path of memory: eye scanpaths facilitate retrieval of visuospatial information. *Cogn. Process.* **16**, 159–163. (doi:10.1007/s10339-015-0690-0)
38. Kragel JE, Voss JL. 2020 Temporal context guides visual exploration during scene recognition. *J. Exp. Psychol. Gen. Advance* **150**, 873. (doi:10.1037/xge0000827)
39. Foulsham T, Kingstone A. 2013 Fixation-dependent memory for natural scenes: an experimental test of scanpath theory. *J. Exp. Psychol. Gen.* **142**, 41–56. (doi:10.1037/a0028227)
40. Champely S. 2016 pwr: Basic functions for power analysis. R package version 1.1-3. 2015.
41. Dewhurst R, Foulsham T, Jarodzka H, Johansson R, Holmqvist K, Nyström M. 2018 How task demands influence scanpath similarity in a sequential number-search task. *Vis. Res.* **149**, 9–23. (doi:10.1016/j.visres.2018.05.006)
42. Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, Kastman E. JE, Lindeløv JK. 2019 PsychoPy2: experiments in behavior made easy. *Behav. Res. Methods* **51**, 195–203. (doi:10.3758/s13428-018-01193-y)
43. Hessels RS, Niehorster DC, Kemner C, Hooge IT. 2017 Noise-robust fixation detection in eye movement data: Identification by two-means clustering (I2MC). *Behav. Res. Methods* **49**, 1802–1823. (doi:10.3758/s13428-016-0822-1)
44. Borji A, Itti L. 2015 Cat2000: A large scale fixation dataset for boosting saliency research. Preprint at See https://arxiv.org/abs/1505.03581.
45. Pärnamets P, Johansson R, Gidlöf K, Wallin A. 2016 How information availability interacts with visual attention during judgment and decision tasks. *J. Behav. Decis. Mak.* **29**, 218–231. (doi:10.1002/bdm.1902)
46. Bates D, Mächler M, Bolker B, Walker S. 2015 Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* **67**, 1–48. (doi:10.18637/jss.v067.i01)
47. Jarodzka H, Holmqvist K, Nyström M. 2010 A vector-based, multidimensional scanpath similarity measure. In *Proceedings of the 2010 symposium on eye-tracking research & applications,* pp. 211–218. New York, NY: Association for Computing Machinery.
48. Anderson NC, Anderson F, Kingstone A, Bischof WF. 2015 A comparison of scanpath comparison methods. *Behav. Res. Methods* **47**, 1377–1392. (doi:10.3758/s13428-014-0550-3)
49. Frame ME, Warren R, Maresca AM. 2019 Scanpath comparisons for complex visual search in a naturalistic environment. *Behav. Res. Methods* **51**, 1454–1470. (doi:10.3758/s13428-018-1154-0)
50. Dijkstra EW. 1959 A note on two problems in connexion with graphs. *Numer. Math.* **1**, 269–271. (doi:10.1007/BF01386390)
51. 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. *J. Eye Mov. Res.* **5**, 1–14. (doi:10.16910/jemr.5.4.3)
52. Foulsham T, Underwood G. 2008 What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *J. Vis.* **8**, 1–17. (doi:10.1167/8.2.6)
53. Wynn JS, Ryan JD, Buchsbaum BR. 2020 Eye movements support behavioral pattern completion. *Proc. Natl Acad. Sci. USA* **117**, 6246–6254. (doi:10.1073/pnas.1917586117)
54. Gurtner LM, Hartmann M, Mast FW. 2021 Eye movements during visual imagery and perception show spatial correspondence but have unique temporal signatures. *Cognition* **210**, 104597. (doi:10.1016/j.cognition.2021.104597)
55. Eichenbaum H, Yonelinas AP, Ranganath C. 2007 The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* **30**, 123–152. (doi:10.1146/annurev.neuro.30.051606.094328)
56. Danker JF, Anderson JR. 2010 The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychol. Bull.* **136**, 87–102. (doi:10.1037/a0017937)
57. Kent C, Lamberts K. 2008 The encoding-retrieval relationship: retrieval as mental simulation. *Trends Cogn. Sci.* **12**, 92–98. (doi:10.1016/j.tics.2007.12.004)
58. Tulving E, Thomson DM. 1973 Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* **80**, 352–373. (doi:10.1037/h0020071)

59. Roediger III HL, Guynn MJ. 1996 Retrieval processes. In *Memory: handbook of perception and cognition* (eds EL Bjork, RA Bjork), pp. 237–313. Academic Press.
60. Bridge DJ, Cohen NJ, Voss JL. 2017 Distinct hippocampal versus frontoparietal network contributions to retrieval and memory-guided exploration. *J. Cogn. Neurosci.* **29**, 1324–1338. (doi:10.1162/jocn_a_01143)
61. Hannula DE, Ranganath C. 2009 The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron* **63**, 592–599. (doi:10.1016/j.neuron.2009.08.025)
62. Ryals AJ, Wang JX, Polnaszek KL, Voss JL. 2015 Hippocampal contribution to implicit configuration memory expressed via eye movements during scene exploration. *Hippocampus* **25**, 1028–1041. (doi:10.1002/hipo.22425)
63. Smith CN, Squire LR. 2008 Experience-dependent eye movements reflect hippocampus-dependent (aware) memory. *J. Neurosci.* **28**, 12 825–12 833. (doi:10.1523/JNEUROSCI.4542-08.2008)
64. Kragel JE, VanHaerents S, Templer JW, Schuele S, Rosenow JM, Nilakantan AS, Bridge DJ. 2020 Hippocampal theta coordinates memory processing during visual exploration. *Elife* **9**, e52108. (doi:10.7554/eLife.52108)
65. Kragel JE, Voss JL. 2022 Looking for the neural basis of memory. *Trends Cogn. Sci.* **26**, 53–65. (doi:10.1016/j.tics.2021.10.010)
66. Johansson R, Nyström M, Dewhurst R, Johansson M. 2022 Data from: Eye-movement replay supports episodic remembering. Figshare. (doi:10.6084/m9.figshare.c.6035691)