Episodic memory formation in naturalistic viewing

Andrey R. Nikolaev[1,2]*, Inês Bramão[1], Roger Johansson[1], Mikael Johansson[1]

[1]Lund Memory Lab, Department of Psychology, Lund University, Lund, Sweden
[2] Brain & Cognition Research Unit, KU Leuven, Leuven, Belgium

* Corresponding author
Department of Psychology,
Lund University,
Box 213
22100 Lund
Sweden
andrey.nikolaev@psy.lu.se

Abstract

The brain systems of episodic memory and oculomotor control are tightly linked, suggesting a crucial role of eye movements in memory. But little is known about the neural mechanisms of memory formation across eye movements in naturalistic viewing behavior. Here, we leverage simultaneous recording and analysis of eye movements and EEG to examine the formation of episodic memory in free viewing. We designed a stimulus screen with multiple elements that together comprised an event. Participants were asked to memorize several of these events while their EEG and eye movements were concurrently recorded. A subsequent cued-recall test assessed participants' memory for the element associations that specified each event. We overcame the problem of overlapping EEG responses to sequential saccades in free viewing using a deconvolution approach. We segmented EEG relative to the fixation onsets and examined EEG power in the theta and alpha frequency bands, the putative oscillatory correlates of episodic encoding mechanisms. We found that high subsequent memory performance was predicted by three modulations of fixation-related EEG: 1) theta synchronization at fixations after saccades relevant to event integration, 2) theta synchronization at fixations after within-element scrutinizing saccades, 3) alpha desynchronization at fixations after saccades between elements that were incidental to the task. Thus, event encoding with unrestricted viewing behavior was characterized by three neural mechanisms, manifested in fixation-locked theta and alpha EEG activity that rapidly turned on and off during the unfolding eye movement sequences. These three distinct neural mechanisms may be the essential building blocks that subserve the buildup of coherent episodic memories during naturalistic viewing behavior.

Introduction

Life events consist of visual, semantic, contextual, and other features that are associated and stored as coherent representations in our episodic memory. Episodic memory allows us to mentally travel in time to relieve our personal past with a considerable level of detail and specificity (Tulving, 1983). A critical aspect when building such memories of the world around us is that the visual field of our perception is spatially limited. At any given point in time, we can apprehend information in full acuity only in a small spot, determined by the size of the fovea. We overcome this limitation by constantly shifting our visual focus with eye movements. In effect, these visual "samples" of the world constitute the pieces that we successively bind together into meaningful coherent episodic memories (Ryan et al., 2020; Voss et al., 2017). Nonetheless, in the neuroscience of human memory, memory encoding has primarily been studied in experimental paradigms where the study material is presented in a single location on the screen, and where eye movements typically are treated as artifacts. Thus, little is currently known about the neural mechanisms of episodic memory formation across eye movements. In the present study, we set out to capture the neural signatures of building coherent episodic memories through eye movements at the level of single gaze fixations. To this end, we apply a state-of-the-art approach of simultaneous eye-tracking and EEG recording and analysis in a free-viewing memory task.

Previous eye-tracking research has established that the amount of visual sampling (i.e., the cumulative number of fixations) on a visual stimulus predicts subsequent memory for that stimulus (Damiano and Walther, 2019; Loftus, 1972). However, eye movements not only supply memory with visual input from separated fixation locations, but also organize their relationships in space and time, thus acting themselves as a binding mechanism (Ryan et al., 2020; Voss et al., 2017). Moreover, a longstanding theory holds that the sequence of eye movements – the scanpath – established during memory formation is stored together with the sampled information (Noton & Stark 1971; Johansson et al. 2012; Foulsham & Kingstone 2013;Ferreira et al. 2008; Wynn et al. 2019; Johansson et al. 2021) to facilitate the binding of separate elements into coherent wholes. The clock frequency of this binding mechanism is the theta rhythm (4-7 Hz) (Herweg et al., 2020). In natural viewing, we typically make about 3-4 saccades per second, thus sampling visual information at the theta frequency (Amit et al., 2017; Otero-Millan et al., 2008). Attentional deployment during goal-directed behavior also fluctuates with the theta frequency (Busch and VanRullen, 2010; Hogendoorn, 2016; Landau and Fries, 2012). The theta oscillations are predominant in the hippocampus, which is the main dispatcher of relational memory formation (Cohen and Eichenbaum, 1993). The hippocampal theta orchestrates the binding of information coming from distributed brain regions into coherent episodic memories (Backus et al., 2016; Clouter et al., 2017). The hippocampus has tight anatomical connections with brain structures that control oculomotor behavior (Pierrot-Deseilligny et al., 2004; Shen et al., 2016). Recent research suggests that these connections can support the guiding of eye movements during memory formation on a fixation-by-fixation basis (Kragel et al., 2021; Voss et al., 2017). Thus, the theta cycle organizes sequential visual inputs and enables the binding of event elements into a coherent memory representation (Herweg et al., 2020).

The relevance of theta oscillations for memory processes has long been established in studies using stimulus-response paradigms with fixed stimulus presentation. For instance, the increased frontal theta power during episodic memory encoding predicts subsequent retrieval performance (Hsieh and Ranganath, 2014; Klimesch, 1996). The increase in theta power, i.e., theta synchronization, is often accompanied by a decrease in alpha (8-13 Hz) power, i.e., alpha desynchronization (Hanslmayr et al., 2016; Herweg et al., 2020; Waldhauser et al., 2012). Alpha desynchronization may represent the fidelity of memory representations for the constituent elements of episodes (Hanslmayr et al., 2016). Thus, theta and alpha oscillations serve important but distinct functions of episodic memory formation: binding individual elements together into a whole (theta) and representing elements

(alpha). It is currently unknown how these mechanisms engage during the sequential sampling of event elements across saccades to allow later episodic remembering of the event as a whole. In the current study we track how these neural mechanisms subserve the buildup of episodic memories during unrestricted viewing behavior.

To study neural activity across eye movements it is necessary to record and analyze brain activity and eye movements simultaneously. Over the last decade, the combination of EEG and eye-tracking has become a powerful tool to study perceptual and cognitive processes in free-viewing behavior (Coco et al., 2020; Devillez et al., 2015; Dimigen et al., 2011; Fudali-Czyz et al., 2018; Körner et al., 2014; Tyson-Carr et al., 2020), including attention and memory across saccades. For instance, presaccadic EEG activity has been found to predict successful encoding of individual scene elements and to reveal dissociation between attention and eye movements in a change blindness task, leading to detection failures (Nikolaev et al., 2011). However, despite the apparent benefit of utilizing EEG-eye movement coregistration, attempts to apply it are virtually absent in the episodic memory literature. One explanation may be the vast methodological challenges that come with these techniques, including the effects of overlapping saccades on EEG activity in free viewing (Dimigen and Ehinger, 2021; Dimigen et al., 2011; Nikolaev et al., 2016). Moreover, low-level oculomotor effects of eye movement characteristics can systematically differ between experimental conditions and distort the EEG (Thickbroom et al., 1991; Yagi, 1979), creating differences that are likely to be confounded with effects of the experimental conditions. To overcome these and other challenges as described below, the present study applied a regression-based deconvolution approach that recovers unknown isolated neural responses given the measured EEG and the latencies of experimental events (Dimigen and Ehinger, 2021; Ehinger and Dimigen, 2019). This ultimately allowed us to isolate and track the neural correlates of episodic memory formation across eye movements.

The present study introduces an associative memory task, in which participants encoded multiple events, each comprising three pairs of images from three different categories (faces, places, objects) in different locations on the screen (Fig. 1). Two exemplars from the same category were close together on the screen, forming a pair, while the distances between pairs, i.e., between categories, were much greater. This allowed us to distinguish gaze transitions between different categories (*between-category* saccades) and between two exemplars from the same category (*between-exemplar* saccades). In a subsequent cued-recall task, we tested episodic memory for between-category associations (i.e., associations between exemplars from different categories) within each event, the formation of which critically depends on the corresponding between-category saccades during encoding. The within category associations (i.e., associations between two exemplars from the same category) were not tested. Thus, between-category and between-exemplar saccades correspond to associations between event elements that differ in their task-relevance. Additionally, we examined scrutinizing saccades within an exemplar (*within-element*), as it is well-known that the cumulative number of fixations on a visual stimulus predicts subsequent memory of that stimulus (Loftus, 1972; Olsen et al., 2016). This allowed us to contrast the neural correlates of two types of memory-related mechanisms: the first related to eye movements that serve to sample visual information from individual elements, the second related to eye movements that serve to bind the elements into a coherent event.

Moreover, for between-category and between-exemplar saccades, we distinguished between saccades associated with the first “visit” to a category/exemplar and “revisits” to them. Revisits, i.e., refixations of previously visited locations, has been found to be of particular interest in memory-related gaze behavior. Particularly, refixations can increase sampling of information necessary for successful memory formation (Mirza et al., 2016; Turk-Browne, 2019; Voss et al., 2011) and recover information lost or missed during scanning (Gilchrist and Harvey, 2000; Körner and Gilchrist, 2008; Meghanathan et al., 2019; Tatler et al., 2005). Our previous coregistration studies have shown that

refixations and ordinary fixations differs in the allocation of attention during saccade planning (Nikolaev et al., 2018), and the task-relevance of re-fixated items affects refixation planning (Meghanathan et al., 2020). A recent study combining eye-tracking with intracranial EEG found that presaccadic hippocampal theta activity predicts whether an upcoming saccade will be a first visit or a revisit (Kragel et al., 2021). Based on these previous findings, we expected EEG activity during between-category revisits to be predictive of successful binding of the task-relevant event elements.

Methods

Participants

Thirty-six healthy adults took part in the experiment in exchange for a gift card in a shopping mall (approx. €10). Three participants were excluded because of system crashes during the recording and another three due to below-chance memory performance (53.2%). Finally, two participants were excluded due to ceiling performance (> 90%), i.e., too few incorrect trails for the planned analysis. The final group included 28 participants (20 females; mean age 23.3; age range 18-32).

The study was conducted in accordance with the Swedish Act concerning the Ethical Review of Research involving humans (2003:460). All participants gave written informed consent, and the study followed the local ethical guidelines at Lund University. Moreover, the study followed the Lund University safety protocol for conducting EEG, eye-tracking and behavioral experiments in the context of the COVID-19 pandemic.

Stimuli

A total of 324 photo images, from three different categories: faces, places, and objects, were used. Faces were selected from the Oslo Face Database (Chelnokova et al., 2014), which provided consent for the publication of the faces that were used in the figures. Objects were selected from the Bank of Standardized Stimuli (Brodeur et al., 2010). Places were selected from the Michael Tarr lab scene database (http://www.tarrlab.org/). The images were converted to black-and-white, and each image's pixel intensity was normalized to the average pixel intensity of all images. The images were combined into 54 events. Each event comprised three image pairs (two exemplars from each category) distributed across distinct locations on the display (Fig. 1). Exemplars from the same category were close to each other, whereas the different categories were more separated. Specifically, each image had a square shape with a side of 4.2 degrees of visual angle (dva). The width of a pair (two exemplars from one category) was 8.5 dva, and the distance between the centers of the three pairs was 17.9 dva. The three pairs were placed on an invisible circle with a radius of 9.2 dva centered at the center of the screen. The location of the pairs on the circle was random for each event, but the pairs were always separated by arcs of 120 degrees. The distribution of images and their pairing across events was random for each participant. The images were presented on a grey background.

Procedure

Participants sat in a dimly lit room, with their heads stabilized with a chin rest. Stimuli were presented at a viewing distance of 62 cm on an EIZO FlexScan EV2451 monitor with a resolution of 1920 x 1080 pixels and a refresh rate 60 Hz, which was part of the Tobii Pro Spectrum eye tracker.

The experiment comprised six blocks with the possibility to take a short break between each block. Each block included two main stages: memory encoding and memory retrieval with a distractor task in-between (Fig. 1). Participants self-initiated each block by pressing the space bar. During encoding, nine events were presented for 10 s each. Before each event, a fixation cross was presented for a

random interval between 1 and 1.5 s. Participants were instructed to memorize the faces, places and objects that appeared together on the same event. The encoding stage was followed by a distractor task, where participants were asked to count backwards in steps of 7 from a random 3-digit number during 30 s.

Three seconds after the distractor task, a cued recall task followed, which assessed episodic memory by testing retrieval of the between-category associations of each event. A test screen with three images, one at the top and two at the bottom, was presented for 2 s (Fig. 1). Each test screen was preceded by the presentation of a fixation cross randomly jittered between 1 and 1.5 s. The image at the top served as the retrieval cue, and participants were asked to indicate which of the two images at the bottom that had appeared together with the cue during encoding, i.e., to indicate the target. The distractor was an exemplar from the same category as the target but selected from a different event studied no more than two events before or after the target event. Participants responded by pressing the left or right arrow keys of a standard keyboard. After response, confidence was measured. Participants were asked to indicate how certain they were in their response. The question "Are you sure?" was displayed in the screen and participants had three response options: "Sure", "Maybe", "Guess", which were given by pressing the left, down and right arrow keys, respectively.

Between-category associations for each image in each event were tested on four trials. For example, a face image was tested twice with place images (once as cue and once as target), and twice with object images (once as cue and once as target). The cue-target status of categories was counterbalanced across blocks, such that in the odd blocks the cue-target status was faces-places, places-objects, objects-faces, and in the even blocks the corresponding cue-target status was places-faces, objects-places, faces-objects. Thus, memory for each event was tested for 12 different associations. In total, 108 associations were tested in each block (12 x 9 = 108 tests for the nine events within a block) with randomized order of presentation.

Prior to the main experiment, two practice sessions were conducted in which participants received feedback on their performance after each test trial. The practice session ensured that participants were familiar with the task and had understood the instructions. The overall duration of an experiment was about two hours.

The experiment was programmed in PsychoPy v.2020.2.4 (Peirce et al., 2019).

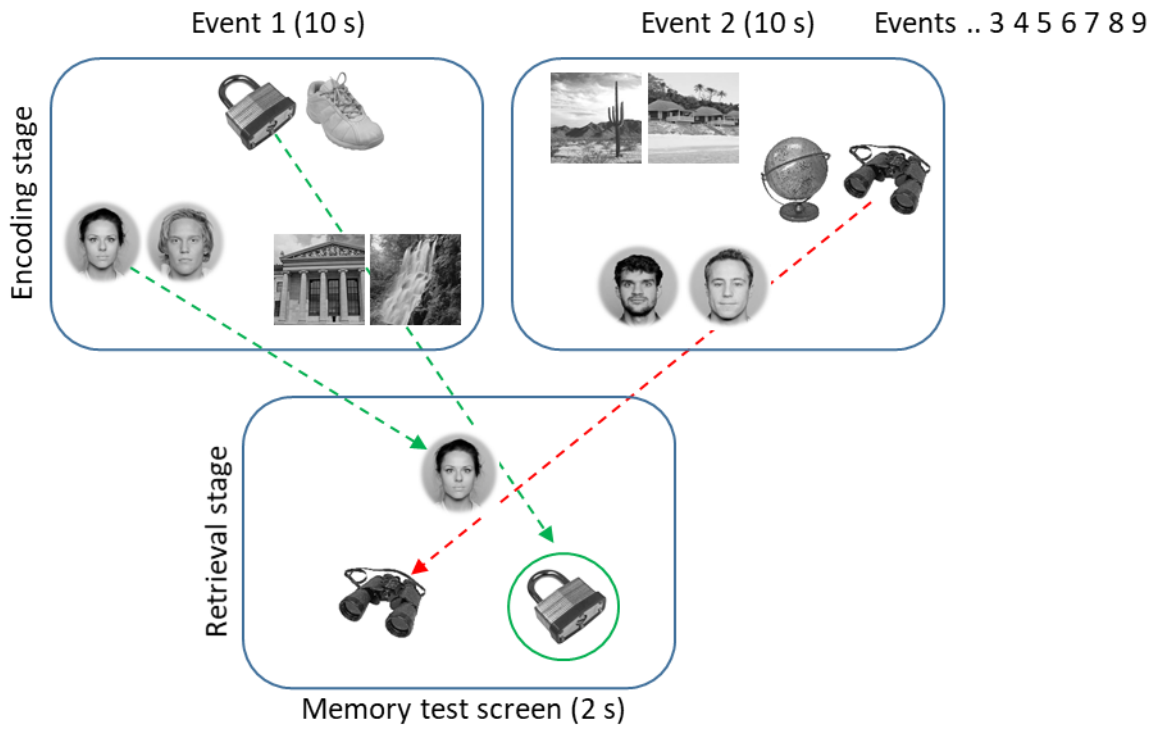

Fig. 1. At encoding, nine events were sequentially presented. Each event comprised two exemplar images from three different categories (faces, places, objects) and was presented for 10 s. After a 30-s distractor task, episodic memory was assessed by testing all between-category associations from each event. The test screen (presented for 2 s) included a cue image at the top with two images at the bottom from a different category. Participants were asked to indicate which of the two images that had appeared together with the cue in the previously encountered events. In this example the correct response is the lock to the right, as it appeared together with the face cue in Event 1, whereas the binoculars to the left appeared in Event 2. Note that on the actual encoding screen the angular distances between categories were much larger.

Simultaneous EEG and eye movement recording

EEG and eye movements were recorded concurrently throughout the experiment. EEG data were recorded using a Grael 4K EEG amplifier (Neuroscan, Compumedics Limited, Australia) at a sampling frequency of 2048 Hz using 31 Ag/AgCl electrodes positioned according to the extended 10-20 International system (Easycap GmbH, Germany). Two electrodes were mounted over the left and right mastoids, using the latter as recording reference. The ground electrode was AFz. Additional electrodes recorded the vertical and horizontal electrooculograms (VEOG and HEOG): the VEOG electrodes were placed above and below the right eye, the HEOG electrodes were placed at the left and right outer canthi of two eyes. Electrode impedances were kept below 5 kΩ.

The Tobii Pro Spectrum eye tracker (Tobii, Stockholm, Sweden) recorded movements of both eyes with the sampling frequency of 600 Hz. 9-point calibration and validation routines were conducted before the first block. An error of 1 dva between calibration and validation was tolerated: if it was larger, the calibration was repeated. The eye tracking was controlled from the PsychoPy experimental program via the open-source toolbox Titta (Niehorster et al., 2020). The experimental program ran on the same computer that controlled the eye tracker. To synchronize stimulus presentation and EEG recording, a transistor-transistor logic (TTL) signal was sent via a parallel port from the stimulus presentation computer to the EEG system at the beginning and end of each block, encoding event, and memory test.

Data analysis

Only data from the encoding stage were analyzed. Analysis was conducted with custom scripts for R (Team, 2020) as well as for MATLAB R2020a (The MathWorks Inc., Natick, Massachusetts), using the EEGLAB toolbox (Delorme and Makeig, 2004), the EYE-EEG extension (Dimigen et al., 2011) for EEGLAB, and the Unfold toolbox (Ehinger and Dimigen, 2019).

Behavioral analysis

Memory performance for each event was assessed by retrieval accuracy weighted by confidence, such that correct responses were scored as 3 for “Sure” responses, as 2 for “Maybe” responses, and as 1 for “Guess” responses. Each image from each event occurred either as a cue or a target in four memory tests trials. Thus, each image can in total receive a score ranging from 0 (incorrect responses in all four tests) to 12 (correct responses accompanied by four “sure” responses in all four tests). To calculate a memory score for the whole event, we then summed the scores for the six elements of each event (range from 0 to 72). For the EEG analysis, events were divided into the two categories “high memory” and ”low memory” according to the subsequent memory scores. For each participant, high and low memory was calculated by a median split of that participant’s memory scores for the whole events.

Eye movement analysis

Fixations and saccades were detected from participants’ right eye using the velocity-based algorithm for saccade detection (Engbert and Mergenthaler, 2006) implemented in the EYE-EEG extension. Fixations were considered to be on an image if they were located within 0.5 dva of the outer contour of the image. Since participants cannot adequately perform a memory test if they did not visually process (fixate) all images while event encoding, we excluded from all analyses the events where fixations were detected only on two of the three categories (on average 3±4 (mean ± SD) events per participant). Such fixations and fixations outside the images on different parts of the display were assigned to an 'other' saccade type, which were used as a reference in the EEG analysis (see the Deconvolution section below).

Saccades were categorized according to whether they occurred within an exemplar (within-element saccades), between the two exemplars of a category (between-exemplar saccades), and between categories (between-category saccades). In addition, we distinguished between the first visit to an image and revisits to it. Altogether, we defined five different types of saccades (Fig. 3A):

1. Between-category saccades: first visit
2. Between-category saccades: revisits
3. Between-exemplar saccades: first visit
4. Between-exemplar saccades: revisits
5. Within-element saccades

The fixation-related EEG analysis examined EEG activity in the fixation intervals that succeeded these five saccade types.

EEG preprocessing

The goal of our study required fixation-related analysis of EEG in short epochs time-locked to fixation onsets in the encoding interval. This analysis used deconvolution modelling to correct for the overlapping effects of saccades on EEG. Deconvolution modelling involves time regression, which requires continuous EEG, therefore, preprocessing was performed on continuous EEG. In addition, we analyzed EEG time-locked to the event onset during the entire 10-s encoding interval. The preprocessing for this analysis was the same as for the fixation-related analysis up to the segmentation step.

We down-sampled EEG signals to 256 Hz. The EEG and eye-tracking data were synchronized using the function *pop_importeyetracker* from the EYE-EEG extension. Saccades, fixations, event onsets, and bad eye-tracking intervals were inserted into the EEGLAB data structure.

The preprocessing pipeline included a series of cleaning functions from EEGLAB. The *pop_cleanline* function removed power line noise using multi-tapering and a Thompson F-statistic. The *clean_artifacts* function removed flat-line channels, low-frequency drifts, noisy channels, short-time bursts. This function is based on artifact subspace reconstruction (ASR), which compares the structure of the artifactual EEG to that of known artifact-free reference data (Kothe and Jung, 2016). The tradeoff between artifactual and retaining brain activities depends on the ASR parameter, which we set to 20 according to the recommendations by (Chang et al., 2020).

Ocular artifacts were removed with the *OPTICAT* function (Dimigen, 2020). First, EEG was high-pass filtered at 2 Hz in order to suppress large deviations from baseline due to summation of the corneo-retinal artifacts during sequential eye movements. Next, 30-ms segments around saccade onsets were obtained and re-appended to EEG in order to 'overweight' the contribution of the saccadic spike activity in the EEG input to independent component analysis (ICA). Then, the ICA weights obtained after ICA training on these filtered data were transferred to the unfiltered version of the same data. Finally, the ratio between the mean variance of independent components during saccade and fixation intervals was calculated (Plöchl et al., 2012). If this ratio was higher than 1.1, the corresponding independent components were considered as saccade-related and removed.

The automatic classifier (*pop_iclabel*) revealed independent components related to remaining artifacts; these components were removed. EEG was re-referenced to average reference. The removed channels (mean = 1.1, SD = 1 per participant) were interpolated with spherical spline interpolation.

Theta and alpha power were extracted by applying the procedure from (Ossandón et al., 2020). We filtered the continuous EEG with a low cut-off of 4 Hz and high cut-off of 7 for the theta band and with a low cut-off of 8 Hz and high cut-off of 13 for the alpha band using the *pop_eegfiltnew* function with default settings. We Hilbert-transformed the filtered signal and obtained the instantaneous power. We normalized the EEG power at each sampling point to its ratio with the respective channel mean power in a baseline interval. In the analysis of the encoding interval the baseline interval was set from -200 to 0 ms before the onset of an encoding event. In the fixation-related analysis, to avoid an interval of saccade execution just before fixation onset, the baseline interval was set from -200 to -100 ms before the fixation onset. The power values were log-transformed.

Deconvolution

To overcome the problems associated with EEG-eye movement coregistration in free viewing behavior described in Introduction, we used deconvolution modelling implemented in the Unfold toolbox (Ehinger and Dimigen, 2019). The toolbox performs regression-based EEG analysis that includes massive univariate modeling, linear deconvolution modeling, and generalized additive modeling. The outputs of the toolbox are the partial effects (i.e., beta coefficients) for the predictors of interest, adjusted for the covariates included in the model.

The toolbox allowed us to correct overlapping EEG responses to sequential saccades in free viewing, as well as to account for several possible detrimental covariates. These covariates included the low-level oculomotor effects of eye movement characteristics and the ordering factors in time courses of an encoding trial, a block, and an entire experiment. Specifically, we considered the low-level oculomotor characteristics of fixation duration and saccade size and direction, which are known to

affect the fixation-related EEG (Dimigen et al., 2011; Nikolaev et al., 2016). Furthermore, we considered three ordering factors. Since time course of a trial with multiple eye movements may affect the fixation-related EEG activity (Fischer et al., 2013; Kamienkowski et al., 2018), we considered the rank of a fixation in relation to all other fixations that occurred during in the whole encoding interval of an event. The rank was assigned to each fixation as an ordinal number of this fixation in the encoding interval. Behavioral results (see below) revealed a nonlinear effect of event order within a block and a strong linear effect of block order on subsequent memory. We accounted for these factors by including them in the deconvolution model.

The Unfold toolbox involved the following procedures, which were sequentially applied to continuous EEG of each participant.

To account for the experimental conditions of interests and possible confounding factors in the same model we used the following Wilkinson notation of the model formula:

```
Fixation: y ~ 1 + cat(MemoryPerformance) + rank + iBlock + spl(iEventPerBlock,4) +
spl(duration,5) + spl(sac_amplitude,5) + circspl(sac_angle,5,-180,180)
Stimulus: y ~ 1
Levels of MemoryPerformance: 'other' (reference level), 'low memory', 'high memory'
```

This formula includes the multiple effects at fixations, as well as the onset of the event screen. Specifically, for fixations, the formula considers the fixation onset (y ~ 1, i.e., the intercept), a categorical predictor of the subsequent memory effect ('MemoryPerformance') and continuous predictors: fixation rank ('rank'), a block order in an experiment ('iBlock'), an event order in a block ('iEventPerBlock'), fixation duration ('duration'), size ('sac_amplitude') and direction ('sac_angle') of the preceding saccade. In this example formula, the 'MemoryPerformance' predictor includes three levels: 'other' fixations (see Eye movement analysis above), 'low memory' and 'high memory' (see Behavioral analysis above). We modified the formula by defining the categorical predictors depending on the analysis goals, as specified below in each section of Results. In all analyses we used treatment coding with 'other' fixations as the reference level. Since fixation duration, and saccade size and direction have nonlinear effects on EEG (Dimigen and Ehinger, 2021; Nikolaev et al., 2016), we modeled them with a basis set of five spline predictors (splines were circular for saccade direction). Since we found that the event order in the block has a nonlinear effect on the memory score (Fig. 2B), and we assumed its nonlinear effect on the EEG, we modeled it also with spline predictors. Spline knots were placed on the participant-specific percentiles of the covariates. For events, we considered only their onsets (y ~ 1), i.e., the intercept that described the overall shape of EEG evoked by the event screens.

To recover isolated EEG responses (betas) to each fixation and to each event screen, which best explained the continuous EEG, we created a design matrix and time-expanded it in a time window between −200 and +400 ms around fixation and event onsets. Time expansion involved duplicating and shifting the predictors of the mass univariate linear regression design matrix for each time lag. The time-expanded design matrix spanned the duration of the entire EEG recording. As a result, instead of estimating multiple linear models for all time points, we estimated one linear model, simultaneously estimating all fixation and event betas. The variable temporal distance between sequential fixations allowed separation of the overlapping effects.

Before fitting the model, we excluded irrelevant and artifactual intervals from modelling by setting the entire rows of the time-expanded design matrix to zero. Specifically, we removed intervals of the distractor task, memory tests, intervals between events, breaks between blocks, and bad eye-tracking intervals, leaving only intervals of free viewing exploration of event screens during encoding. We fitted the deconvolution model to each of the 31 electrodes using the iterative Least Squares Minimal Residual algorithm for sparse design matrices (Fong and Saunders, 2011).

To extract EEG power for the statistical analysis, we reconstructed the regression-based power waveforms from the beta coefficients of the model for various categorical conditions (specified in each section of Results) in epochs from −200 to +400 ms after the fixation onset. This reconstruction included adjusting the covariate effects of the eye movement characteristics and ordering factors (specified above) by adding the marginal effects of their means. Specifically, for each participant we calculated the average value for each covariate and evaluated the effect-power at this value. This value was then added to all other predictors, effectively adjusting all regression-based power waveforms to the same covariate values. The resulting regression-based power waveforms were equivalents to participant-level averages in a traditional EEG analysis.

Statistical analysis

We estimated theta and alpha oscillations at the first peak of their cycle about 100 ms after the fixation onset. In the statistical analysis, we used the average power between 60 and 140 ms after the fixation onset, that is, 40 ms before and after the peak. To capture effects with potentially different topographical distributions we divided all electrodes (except two temporal electrodes T7 and T8, which were typically noisy) into three broad groups:

1. 12 frontal electrodes: FP1,FP2,F7,F8,F3,FZ,F4,FC1,FCZ,FC2,FC5,FC6
2. 12 centro-parietal electrodes: C3,CZ,C4,CP5,CP1,CP2,CP6,P7,P3,PZ,P4,P8
3. 5 occipital electrodes: PO9,PO10,O1,IZ,O2

In each group, the electrodes were symmetrically distributed over the scalp relatively to the sagittal axis. Since regional averaging of electrodes provides a more reliable estimate of activity in a region than a single measurement (Dien and Santuzzi, 2005), we averaged power waveforms across the electrodes of each group. The statistical analysis was done for these areas separately.

Unless indicated otherwise, for the statistical analysis we used a repeated measures ANOVA. To handle violation of sphericity we applied the Huynh-Feldt correction for p-values associated with two or more degrees of freedom. Fisher's LSD (least significant difference) test was used for post-hoc analyses. The statistical analyses were performed with STATISTICA 10 (StatSoft Inc., Tulsa, OK, USA) and R (R Core Team, 2020).

Results

Behavioral results

Average memory performance across 28 participants was 68.6 ± 8.33 % (mean ± SD). The memory score weighted by confidence for the low memory condition was 27.8 +/- 6.8 and for the high memory condition 48.8 +/- 9 (mean +/- SD). Memory performance changed in the time course of a block and an experiment (Fig. 2). We evaluated these changes with a repeated measures ANOVA on the memory scores weighted by confidence with the factors of Event (nine levels) and Block (six levels). There were significant main effects of Event ($F(8, 216) = 3.7$, $p = .004$, $\varepsilon = .62$) and Block ($F(5, 135) = 9.8$, $p < .001$, $\varepsilon = .95$)), but no interaction between them. To characterize the effects of Event and Block, we used planned comparisons with the linear and quadratic contrasts. The contrast analysis revealed significant effects for the linear contrast for Block ($F(1, 27) = 31.2$, $p < .001$) and for the quadratic contrast for Event ($F(1, 27) = 16.8$, $p < .001$), but not for the linear contrast for Event ($F(1, 27) = 0.7$, $p = .4$). These results indicate a prominent learning effect during an experiment, as well as possible primacy and recency effects in time course of a block.

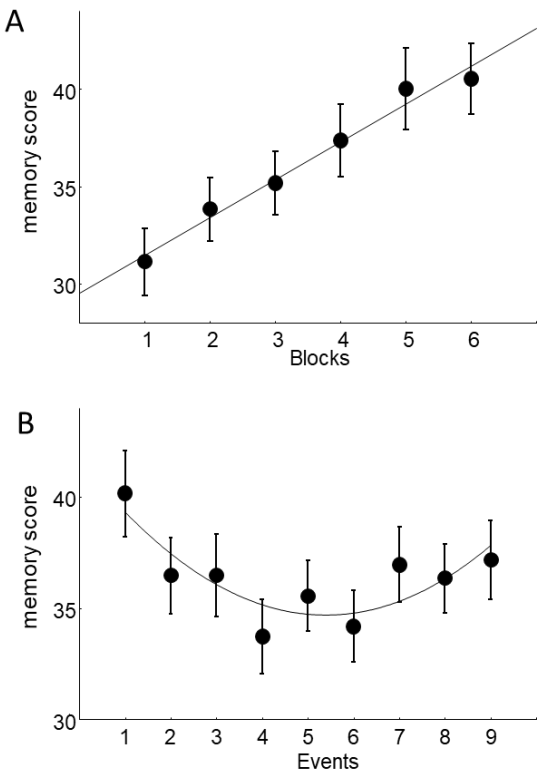

Fig. 2. Memory performance weighted by confidence in the time course of an experiment (A) and a block (B). The thin lines indicate significant linear (A) and quadratic (B) fits.

Eye movement results

The dynamics of eye movement behavior during the 10-s encoding interval is illustrated in Fig. 3B by the fixation rank probability density for each type of preceding saccade. At first, participants tended to visit all categories quickly with between-category saccades. Then, between-exemplar saccades became predominant. The number of within-element saccades was relatively stable throughout the whole encoding interval. Towards the latter part of the interval, between-category revisits became more frequent.

To determine if, and to what degree, the saccade categories predicted subsequent memory performance, we analyzed the relationship between the cumulative number of saccades (in each category) within an event and subsequent memory performance for that event. The between-category saccades for first visits were not analyzed since there were always 3 per event. We computed the number of remaining four saccade types for each event and pooled these numbers with the corresponding event memory scores over 28 participants. We applied a one-way ANOVA on the event memory score with the number of the saccades as a categorical predictor (bins with less than 20 values at the left and right ends of the distributions were excluded). The results showed that memory for the event significantly increased with the cumulative number of between-category revisits ($F(8, 1392) = 3.8$, $p < .001$) and with the cumulative number of within-element saccades ($F(14, 1360) = 3.1$, $p < .001$) (Fig. 3C).

To be successful in the retrieval task, association from all three categories need to be considered during event encoding. This can be achieved by two different gaze patterns: (1) engaging in active gaze transitions between each of the three categories, or (2) engaging in active gaze transitions between only two of the categories, i.e., when the association between the first visited category and the last visited category is inferred without an active gaze transition between the categories. In the present data set, gaze transitions between one of the two categories were absent in approximately

one-third of the events. To evaluate whether the two eye movement patterns are associated with different subsequent memory performance, we compared event memory scores between these pattern types with a paired t-test. We found no difference in memory between the two patterns ($t(27) = 0.05$, $p = .96$), suggesting that event memory does not require gaze transitions between all three categories.

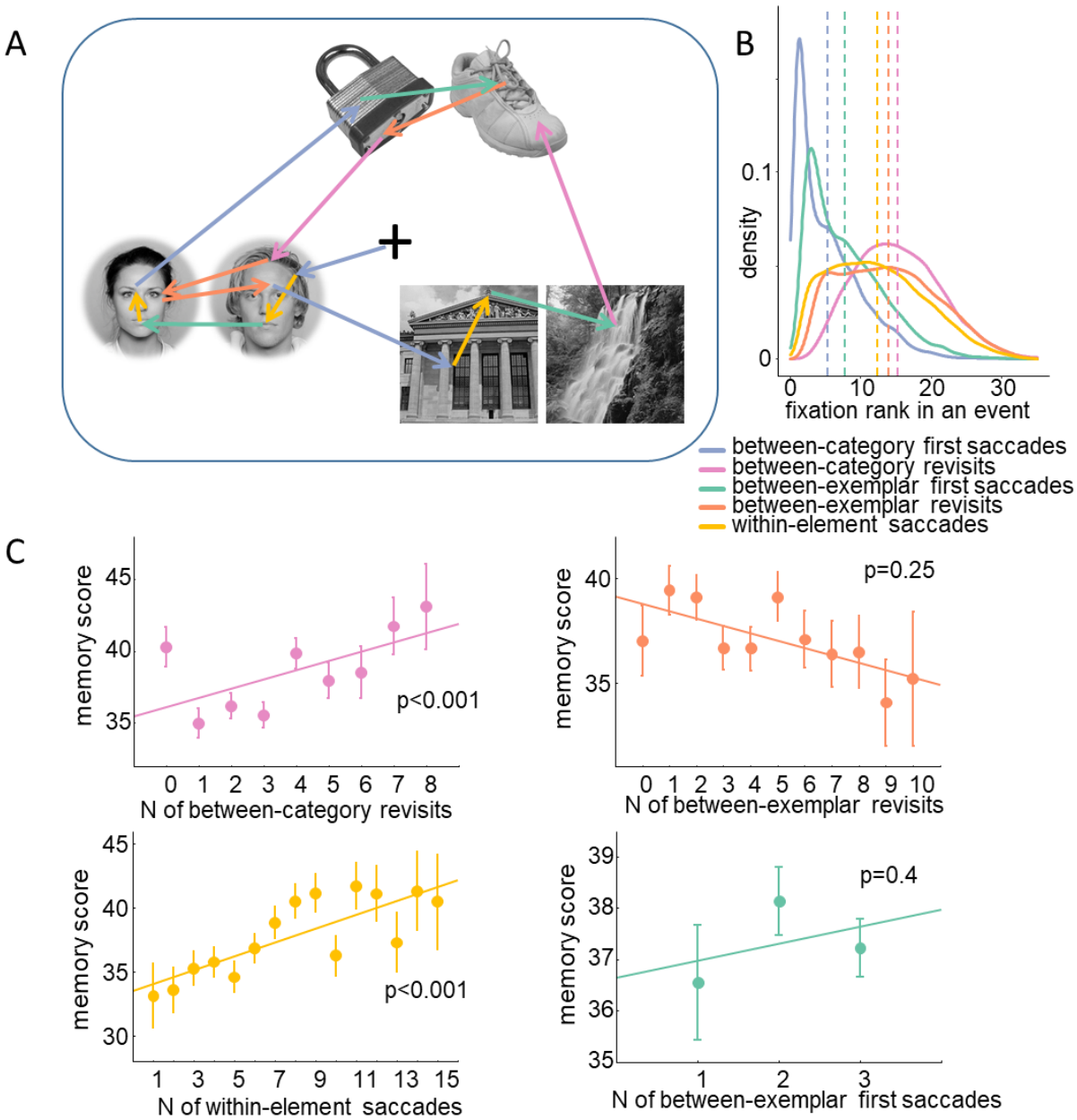

Fig. 3. A: an example scanpath illustrating fives types of saccades. B: Probability density estimation of the fixation rank in a 10-s event interval pooled across 28 participants. The vertical lines indicate mean ranks of the corresponding fixation type. C: Dependence between the occurrence frequency of saccade types and the memory score. Significance of the dependence obtained with one-way ANOVA is indicated at each panel. Error bars indicate standard errors of means across 28 participants.

In sum, the eye movement results show that between-category saccades (revisits) and within-element saccades are predictive of subsequent memory performance, whereas between-exemplar saccades are not. These dependencies were determined by our memory task, which required strong associations between the categories whereas associations between the exemplars of a category

were irrelevant to the task. Within-element saccades are relevant for acquiring sufficient visual information about each image, a prerequisite for establishing a strong representation of each element and for recognizing them during the retrieval task.

EEG results at the event level

To assess how theta and alpha power at encoding predicts subsequent memory at the time scale of the entire encoding interval we segmented the (not deconvolved) EEG power from -0.2 to 10 s relatively to the event onset and baseline corrected it to the interval from – 0.2 to 0 s before the event onset. We smoothed the individual power waveforms with a Savitzky-Golay filter (R package 'signal'). Events were split between low and high memory performance (the number of low memory events was 28.5 (1.35); the number of high memory events was 25.1 (1.74) (mean (SD) per participant). To investigate the modulation across a 10-s encoding interval, we created ten time-bins of 1 s each and averaged the power values in these bins. We applied a repeated-measures ANOVA on the power values with the factors of Memory (low vs. high) and Time bins (10 levels). For theta power, we found interactions between Memory and Time bins for the frontal ($F(9, 243) = 2.4$, $p = .02$, $\varepsilon = .95$) and centro-parietal ($F(9, 243) = 2.1$, $p = .04$, $\varepsilon = .92$) areas (Fig. 4A). Post-hoc tests revealed that these interactions occurred due to higher power for the high than low memory events starting for the frontal areas from the 7th time bin (all $p < .05$) and for the centro-parietal areas from 8th time bin (all $p < .03$) until the end of the encoding interval. For alpha power, there was a clear trend towards lower power in high-memory versus low-memory events, most noticeably over the occipital areas (Fig. 4B), which, however, did not reach significance.

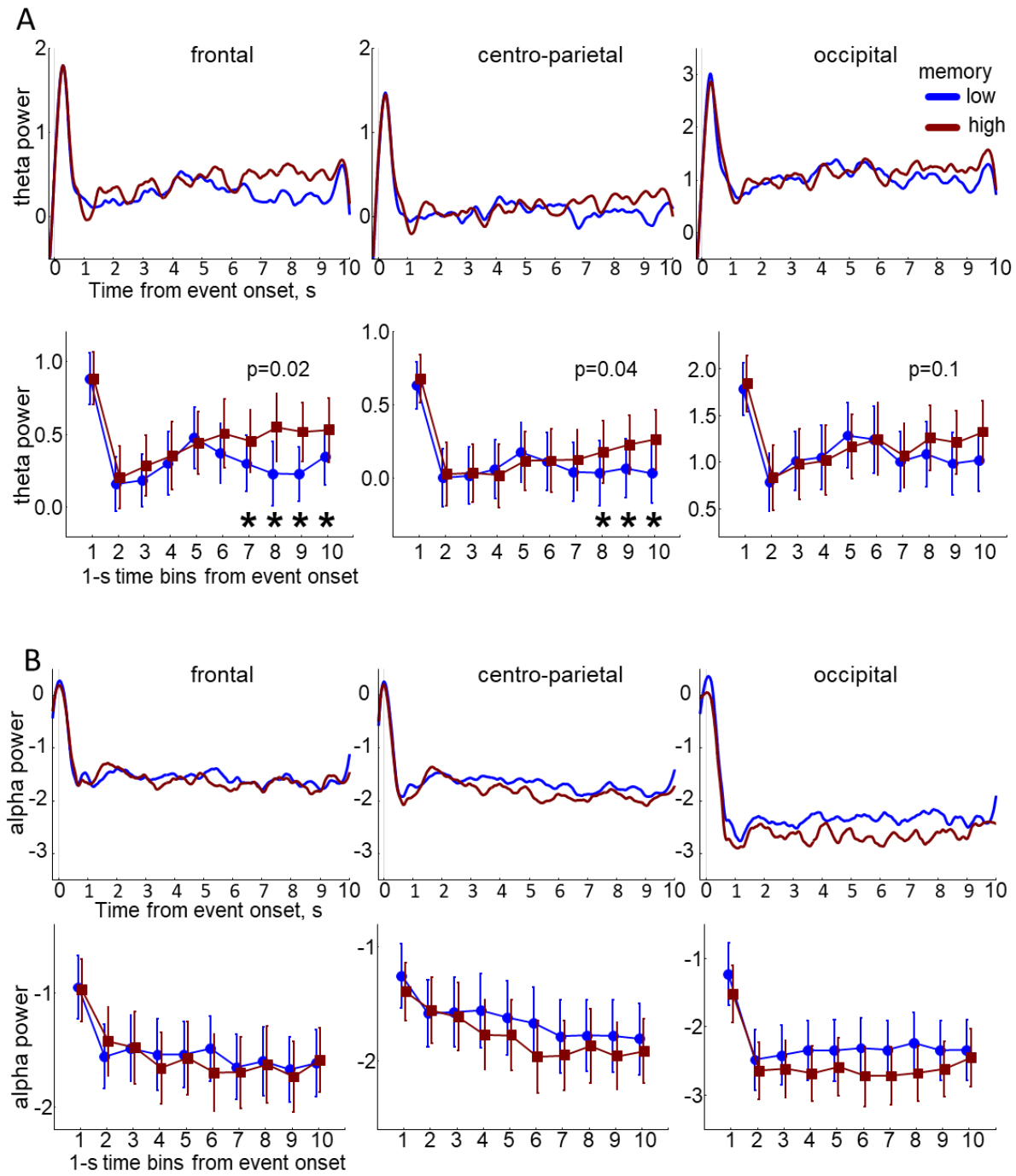

Fig. 4. Theta (A) and alpha (B) power time-locked to the event onset during a 10-s encoding interval for low and high subsequent memory performance. The time course of the grand-averaged power waveforms and the mean-error plots (mean +/- SEM) for ten 1-s time bins are shown. In (A) the p-levels of the interactions between Memory and Time bins and significance of the post-hoc tests (with the asterisks) are indicated.

In sum, EEG analysis of the entire encoding interval revealed the expected neural signatures, where higher subsequent memory performance was associated with increased theta synchronization and alpha desynchronization during encoding (Hanslmayr et al., 2016). Theta synchronization was more prominent in the second half of the encoding interval over the frontal and centro-parietal areas. Theta synchronization was accompanied by a tendency for alpha desynchronization, which was most prominent over the occipital areas (Fig. 4B), although there was no significant difference between high and low memory.

EEG results at the fixation level

Since the deconvolution framework is a novel approach to free-viewing behavior in tasks of episodic memory formation, we first evaluated how it changes the power waveforms (see Supporting Information for details).

The general memory effects of theta and alpha activity

Before analyzing the data for the different saccade categories, we wanted to assess if the EEG findings at the event level could be reproduced at the fixation level. Thus, we tested the subsequent memory effect on fixation-related theta and alpha power by contrasting the fixation-related EEG of high memory events with low memory events over all saccades within an event. Since the event analysis showed a significant memory effect only in the second part of the event (Fig. 4), we divided the fixations into two bins according to fixation rank, which roughly correspond to the first and second half of the 10-s interval and restricted our analysis to the second bin (the second half of the event). To accurately account for the overlapping effects of eye movement sequences on EEG, the deconvolution formula should include fixations of the entire encoding interval. Therefore, in the formula for deconvolution modelling we used five levels of categorical predictors as follows: Other, First half/Low memory, First half/High memory, Second half/Low memory, Second half/High memory. In the second half, the number of the fixation-related epochs per participant was 258 (65.2) in the low memory condition and 222 (58.6) in the high memory condition (mean (SD)).

In a statistical analysis of power averaged in the window from 60 to 140 ms after the fixation onset, we found higher theta power over the frontal areas for encoding leading to high than low memory performance ($t(27) = 2.4$, $p = .03$) (Fig. 5A). No significant effects for other areas were found. Additionally, we observed lower alpha power for encoding leading to high than low memory performance over the frontal areas ($t(27) = 3.3$, $p = .003$) and the occipital areas ($t(27) = 3.0$, $p = .006$) (Fig. 5B).

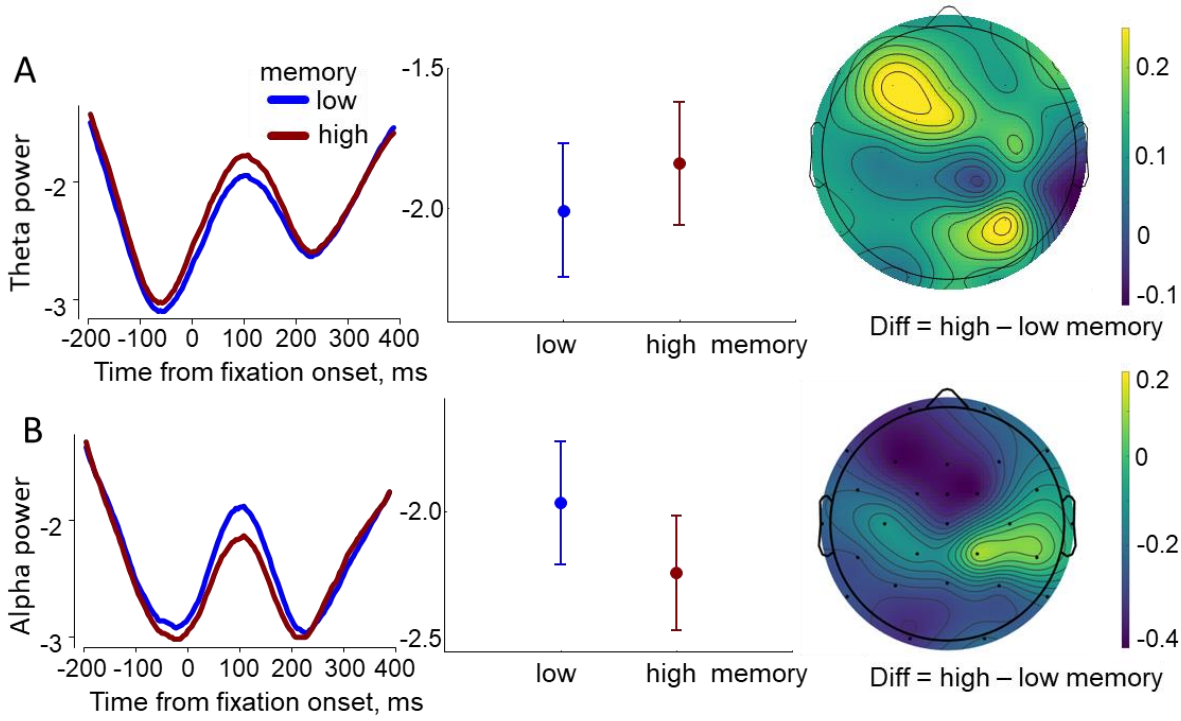

Fig. 5. Theta (A) and alpha (B) power over the frontal areas for fixations in the second half of the encoding interval. The grand-averaged (N=28) power in low and high memory conditions. The mean-error (mean +/-

SEM) plot of the power in the interval from 60 to 140 ms after the fixation onset. The difference power map: high minus low memory.

To summarize, the fixation-related EEG results corroborate and extend what was found at the event level, i.e., theta synchronization and alpha desynchronization during encoding covary with subsequent memory performance. Next, we analyzed EEG power at fixations succeeding the five saccade types.

The EEG results across saccade types

Here, we examined EEG power in the fixation intervals succeeding the five saccade types (Fig. 3A). For this purpose, we specified ten levels of categorical predictors in the deconvolution model formula: the five preceding saccade types x memory performance (low, high). "Other" fixations were the reference level, as in the previous model. The number of the obtained fixation-related epochs is presented in Table 1.

Table 1. The number of fixation-related epochs succeeding each saccade type per participant (mean and SD across 28 participants)

	low memory	high memory
Between-category saccades: first visit	85 (9.8)	67 (8.8)
Between-category saccades: revisits	83 (36.8)	72 (31.3)
Between-exemplar saccades: first visit	67 (12.4)	52 (12.5)
Between-exemplar revisits: revisits	120 (55.2)	90 (40.5)
Within-element fixations	195 (73.2)	155 (58.5)

Between-category saccades

Here, we examined EEG power in the fixation intervals succeeding the between-category saccades that we predicted to support task-relevant associations between categories. We applied a repeated measures ANOVA with the 2x2 design: Gaze transition status (First visits vs. Revisit) x Memory performance (Low vs. High) on theta and alpha power (Fig. 6A). For theta power, we found a significant main effect of Memory performance over the *centro-parietal* areas: theta power was higher for high subsequent memory performance, $F(1,27) = 5.6$, $p = .03$ (Fig. 6B-D). In addition, we found a significant main effect of Gaze transition status over the frontal areas ($F(1,27) = 5.7$, $p = .02$): theta power was higher during the first visit. For alpha power, no significant effects were found.

Thus, theta power over the centro-parietal areas, after between-category saccades, predicts subsequent event memory performance. This theta memory effect is consistent with the eye-movement analyses, where subsequent memory performance increased with the cumulative number of between-category revisits during encoding (Fig. 3C).

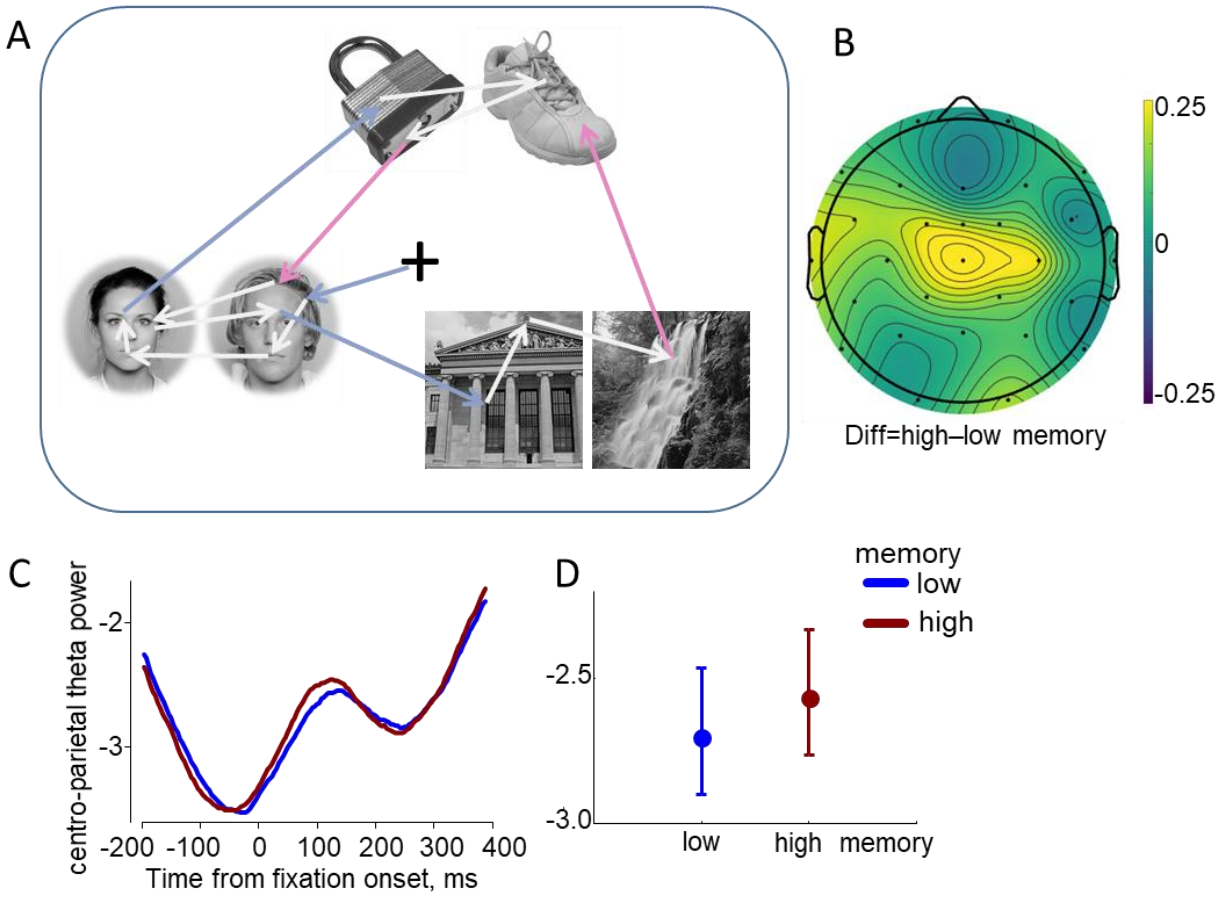

Fig. 6. A: An example scanpath illustrating a subset of between-category (colored) saccades preceding the fixation intervals used in the EEG analysis. The grey arrows are saccades not included in this analysis. B: The difference power maps: high minus low memory. C: The grand-averaged (N=28) theta power over the centro-parietal areas for low and high memory for fixations after the between-category saccades. D: The mean-error (mean +/- SEM) plot of the power in the interval from 60 to 140 ms after the fixation onset.

Between-exemplar saccades

Next, we examined EEG power in the fixation intervals succeeding between-exemplar saccades, i.e., gaze transitions that may be considered irrelevant when building the task-relevant associations between categories (Fig. 7A). As in the previous analysis, we used a repeated measures ANOVA with a 2x2 design: Gaze transition status (First visit vs. Revisit) x Memory performance (Low vs. High). For theta power, no significant effects were found. But we found lower alpha power for high than low memory performance over the centro-parietal ($F(1,27) = 5.0$, $p = .03$) and occipital ($F(1,27) = 6.4$, $p = .02$) areas (Fig. 7B-D).

Thus, as expected, the theta power after task-irrelevant between-exemplar saccades did not predict subsequent event memory performance. This is consistent with the eye-movement result, where the cumulative number of between-exemplar saccades did not influence subsequent memory performance (Fig. 3C). The alpha desynchronization effect over posterior areas was unexpected, but may reflect other mechanisms, apart from binding per se, that are important for optimal memory formation.

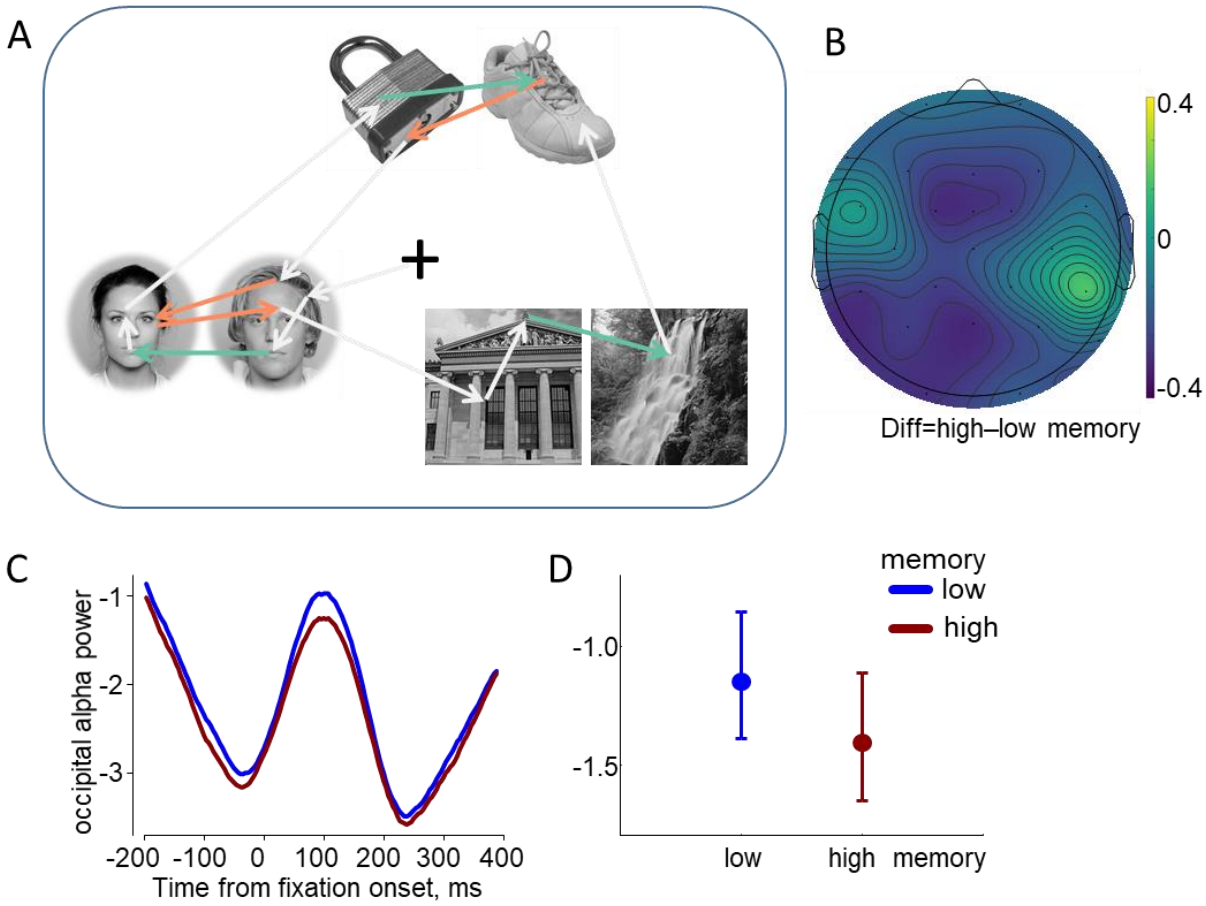

Fig. 7. A: An example scanpath illustrating a subset of between-exemplar (colored) saccades preceding the fixation intervals used in the EEG analysis. The grey arrows are saccades not included in this analysis. B: The difference power maps: high minus low memory. C: The grand-averaged (N=28) theta power over the occipital areas for low and high memory for fixations after between-exemplar saccades. D: The mean-error (mean +/- SEM) plot of the power in the interval from 60 to 140 ms after the fixation onset.

Within-element saccades

The analysis of the fixation intervals after within-element saccades revealed higher theta power for high than low memory performance over the frontal areas ($t(27) = 3.02$, $p = .006$) (Fig. 8). No significant effects were found for other areas or for alpha power.

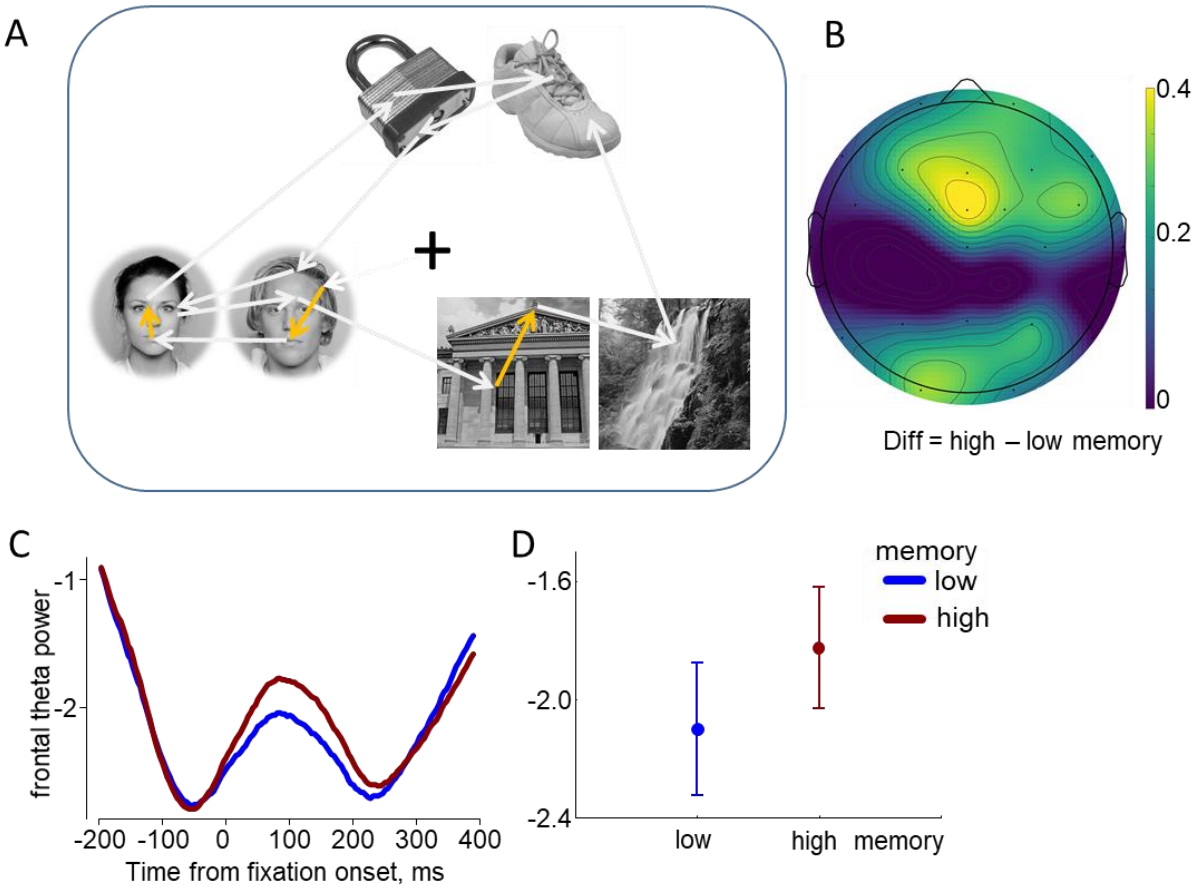

Fig. 8. A: An example scanpath illustrating a subset of within-element (colored) saccades preceding the fixation intervals used in the EEG analysis. The grey arrows are saccades not included in this analysis. B: The difference power map: high minus low memory levels. C: The grand-averaged (N=28) theta power over the frontal areas for low and high memory for fixations after the within-element saccades. D: The mean-error (mean +/- SEM) plot of the power in the interval from 60 to 140 ms after the fixation onset.

Thus, the theta power related to within-element saccades predicts subsequent memory performance similarly as for the between-category saccades. Moreover, for both saccade types, subsequent memory increased as a function of the cumulative number of saccades during encoding (Fig. 3C). Importantly, however, the topographies of the theta memory effects for the two saccade types were *distinct*: a centro-parietal maximum for between-category saccades (Fig. 6B) and a frontal maximum for within-element saccades (Fig. 8B), suggesting dissociable neural mechanisms.

To test the reliability of the topographical dissociation of the two theta effects we extracted the 'within-element effect' as the difference between theta power in high and low memory performance for fixations after within-element saccades. We extracted the 'between-category effect' as the difference between theta power in high and low memory performance for fixations after between-category saccades (averaged across first visits and revisits). We scaled the power values across the frontal, centro-parietal and occipital areas by dividing the value of each area by the square root of the sum of squares for all areas (McCarthy and Wood, 1985). A repeated-measures ANOVA on the scaled theta power with factors of Theta effect ('within-element' vs. 'between-category') and Areas (frontal, centro-parietal and occipital) revealed an interaction between these factors ($F(2, 54) = 4.8$, $p = .02$, $\varepsilon = .81$). The post-hoc test revealed higher power for the between-category effect over the centro-parietal area ($p = .01$) and a reversed tendency for the frontal area ($p = .08$) (Fig. 9). This finding indicates topographical differences for the two theta effects and suggests that they are associated with distinct neural mechanisms, operating in the same frequency band.

In sum, the correspondence of theta and eye movement results for between-category saccades and within-element saccades, which may be based on distinct mechanisms, indicates a diversity of relationships between theta oscillations, eye movements, and episodic memory formation, as will be discuss below.

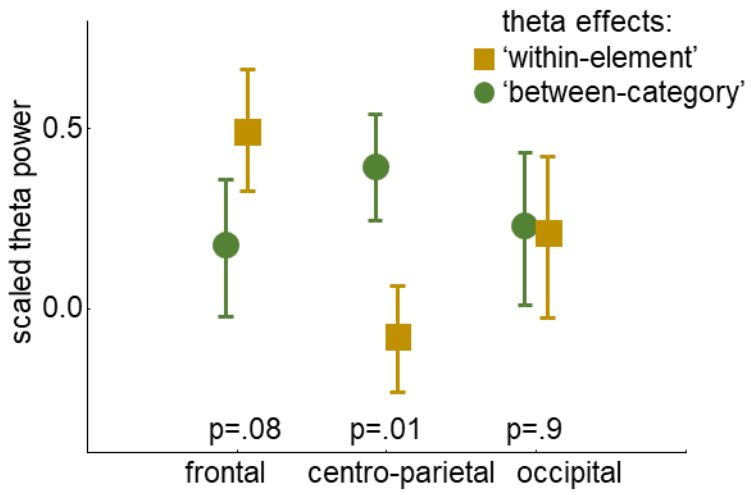

Fig. 9. The scaled theta power for the 'within-element' and 'between-category' theta memory effects. The mean-error (mean ± SEM) plot of the power in the interval from 60 to 140 ms after the fixation onset. The p-levels indicate the results of post-hoc comparisons between low and high memory performance for each area.

Discussion

The present study set out to advance our current understanding of the relationship between eye movements and episodic memory encoding. In particular, we aimed at elucidating the neural mechanisms that mediate the formation of coherent episodic memories as a function of saccades to event elements during unrestricted viewing behavior. To this end, we applied a state-of-the-art analytical approach to EEG coregistered with eye movements during a free-viewing episodic memory task. Our approach provided a unique possibility to capture neural encoding mechanisms across eye movements at the level of gaze fixations while overcoming confounding effects of sequential saccades on brain activity. We identified neural signatures that indicate the buildup of coherent episodic memories as a function of visual sampling behavior with varying relevance for the integration of event elements. The encoding of task-relevant associations between event elements was accompanied by simultaneous modulation of eye movements and theta oscillations that were predictive of subsequent memory. This modulation may reflect the binding of event elements into a coherent representation across eye movements. In addition, a subsequent memory effect was also observed for theta oscillations related to small scrutinizing saccades within elements. The topographies of these theta effects were distinct, revealing the contributions of dissociable neural mechanisms to memory formation. Although association between category exemplars was task-irrelevant, alpha oscillations related to between-exemplar saccades also proved predictive of later episodic remembering. Overall, we have identified the engagement of three distinct neural mechanisms that may be the essential building blocks in formation of episodic memory during naturalistic viewing.

Thus, theta and alpha oscillations serve important but distinct functions of episodic memory formation: binding individual elements together into a whole (theta) and representing elements

(alpha) (Hanslmayr et al., 2016). But it has not yet been established how these mechanisms engage during sequential sampling of event elements with saccades for subsequent episodic remembering of the event as a whole. In our study, we tracked how these neural mechanisms subserve the buildup of episodic memories during unrestricted viewing behavior.

Saccade types that predict subsequent memory

It is well known that eye movement behavior during free viewing is important for successful formation of episodic memory (Wynn et al., 2019). We replicate and extend this finding by showing that the number of within-element saccades and between-category revisits positively correlated with performance in a relational memory task (Fig. 3C). Short within-element saccades scrutinize the visual features of single exemplars, increasing the visual sampling of different details, which leads to better subsequent memory of the stimuli (Loftus, 1972; Olsen et al., 2016). The role of these saccades in memory formation is clearly different from the role of between-category revisits that create associations between categories, which was crucial for successful performance in the memory task. As expected, the cumulative number of gaze transitions between categories during encoding increased memory performance. The between-category saccades are likely involved in binding separate elements into a coherent representation of the event. The number of gaze fixations (Liu et al., 2017) and especially the number of revisits increases hippocampal activity and memory performance (Voss et al., 2011), which is consistent with the results of the between-category revisits. In contrast, the associations between between-exemplar saccades were not relevant for successful task performance, which may explain why between-exemplar saccades were not predictive of memory performance.

Theta and alpha activity throughout the encoding interval predicts subsequent memory

The EEG analysis first confirmed the presence of theta and alpha oscillations as predicted from previous work on episodic memory where to-be-remembered material typically is presented in a central location on the screen with reduced need for visual exploration via eye movements (Hanslmayr and Staudigl, 2014; Klimesch, 1996). In line with the extensive evidence from such studies, theta power was higher for high than low subsequent event memory during the 10-s encoding interval (Fig. 4A). This difference in theta power increased over the time course of encoding and became significant towards the end of the encoding interval, about 7-10 s after the event onset. The predominantly anterior location of the theta effect suggests the contribution of frontal midline theta oscillations, which are related to intrinsic episodic memory processes (Hsieh and Ranganath, 2014).

The alpha power in the encoding interval was not statistically different between high and low memory. However, Fig. 4B shows a tendency for this difference, mostly noticeable over the occipital areas (characteristic of alpha activity). As expected, the relationship with subsequent memory is opposite for alpha and theta activity: higher subsequent memory is associated with lower alpha power. This is consistent with the well-documented finding that a decrease in alpha power reflects successful memory formation (Griffiths et al., 2019; Hanslmayr and Staudigl, 2014; Klimesch et al., 1996).

Overall, in good agreement with the expected role of theta and alpha oscillations in episodic memory formation, we observe opposite trends in theta and alpha activity throughout the encoding interval that are predictive of subsequent memory. These theta synchronization and alpha desynchronization are likely complementing each other, reflecting the functioning of the hippocampal and cortical memory systems, respectively (Hanslmayr et al., 2016).

Fixation-related EEG predicts subsequent memory

In the initial analysis of the fixation-related EEG, we found higher theta and lower alpha power for high than low memory performance (Fig. 5), similar to the theta and alpha effects at the event level. However, at the fixation level, both theta and alpha effects were statistically significant. The significance of the effects at the fixation level, but not at the event level, suggest that segmentation of the EEG relative to fixations isolates its features characteristic of the subsequent memory effect. This refining property of fixation-based segmentation may be a consequence of normalization of the fixation-related EEG relatively to the -200-100 ms interval before fixation onset. Normalization reduces nonspecific baseline trends in the EEG throughout the encoding interval but accentuates the acquisition of memory-related information at fixations.

The time window in the fixation-related EEG analysis was centered on the first power peak about 100 ms after the fixation onset. The latency of this peak coincides with the latency of lambda activity, which is associated with early perceptual processes at gaze fixation (Kazai and Yagi, 2003; Ossandón et al., 2010; Ries et al., 2016; Thickbroom et al., 1991). However, it seems unlikely that the observed memory effects are related to early perception because they have a quite distinct topography from that of lambda activity, over occipital regions (Ossandón et al., 2010; Thickbroom et al., 1991). The widespread topographies of theta and alpha memory effects indicate that the time window we used captures the cycle of ongoing brain oscillations that carry out the rhythmic sampling of external information. Indeed, oscillations at the theta frequency drive perceptual sampling by organizing sequential inputs (Herweg et al., 2020). In free-viewing behavior, perceptual sampling is performed by saccadic eye movements generated at theta frequency (Amit et al., 2017; Otero-Millan et al., 2008). Eye movements are accompanied with phase resetting in a wide frequency range of oscillations observed in the visual cortex of monkeys (Ito et al., 2011; Rajkai et al., 2008), as well as in the scalp EEG in humans (Nikolaev et al., 2016). Since phase resetting of hippocampal theta oscillations supports successful memory formation (Jutras et al., 2013; Kota et al., 2020), rhythmic gaze sampling may synchronize brain areas and optimize the processing and encoding of visual information (Jutras and Buffalo, 2010; Voss et al., 2017). Importantly, theta oscillations represent several items within the same cycle, where the current item is strongly represented and previous items concurrently represented in a weaker form (Herweg et al., 2020). Transferred to our paradigm, theta activity may represent not only the fixated information but also previously visited exemplars (as well as previously presented events). Our time window thus captures individual perceptual moments at fixations as well as the sum of active and latent representations accumulated over a series of eye movements.

The observed memory effects indicate that theta synchronization and alpha desynchronization contribute to formation of episodic memory at each fixation location in free-viewing encoding. This is consistent with the view that memory formation is intimately tied with exploratory viewing (Voss et al., 2017), and extends this view by elucidating the role of theta and alpha activity. Theta synchronization may support binding episodes and alpha desynchronization may represent the content of these episodes, as it has been established in studies where stimuli were presented in a single location on the screen (Hanslmayr et al., 2016).

EEG memory effects for saccades between event elements

Perhaps the most fascinating results were obtained when analyzing EEG in fixation intervals after different saccade types during the 10-s encoding interval. We analyzed EEG power at fixations that followed between-category, between-exemplar, and within-element saccades. In our paradigm, between-category saccades supported task-relevant associations between event exemplars. The theta power was higher for high than low subsequent memory at fixations after between-category saccades (Fig. 6), and the cumulative number of such gaze transitions predicted subsequent memory (Fig. 3C). The alpha power was lower for high than low memory at fixations after between-exemplar

saccades, i.e., task-irrelevant associations (Fig. 7). The cumulative number of such gaze transitions did not predict subsequent memory (Fig. 3C).

Thus, it is conceivable that gaze transitions between categories link the event elements together, creating a holistic representation of the entire event. These gaze transitions are thus of critical importance to episodic memory as they support the binding of information experienced separately in space and time (Cohen and Eichenbaum, 1993; Voss et al., 2017). The binding mechanism relies on the transfer of information between distant brain regions and is promoted by the synchronization of their activity (Clouter et al., 2017; Fell and Axmacher, 2011). In the binding of episodic memory, the key role belongs to the interaction between the hippocampus and the neocortex (Hanslmayr et al., 2016; Herweg et al., 2020), including the cortical regions of oculomotor control that drive visual exploration (Voss et al., 2017). The cortical-hippocampal interaction underlying memory integration involves theta oscillations (Backus et al., 2016; Kota et al., 2020; Staudigl and Hanslmayr, 2013). The theta rhythm also underlies viewing behavior optimal for memory formation (Herweg et al., 2020; Jutras and Buffalo, 2010; Voss et al., 2017). The high theta power observed after between-category saccades, predictive of high event memory, may reflect optimized oculomotor control for the binding of task-relevant elements into a coherent episodic representation.

The key role of the hippocampal theta oscillations in memory-related gaze behavior has recently been demonstrated using intracranial EEG recording combined with eye tracking in free-viewing memory tasks (Kragel et al., 2021; Kragel et al., 2020). These studies have shown that the hippocampal theta-gamma phase amplitude coupling predicts memory updates when encoding new information. Moreover, the hippocampus guides visual exploration by phase-locking the fixation onset to the peak or trough of theta oscillations (Kragel et al., 2020). Furthermore, Kragel and colleagues (2021) described how revisits support the efficiency of visual sampling in the formation of memory for scenes. Hippocampal theta power decreased before revisits and increased after revisits, indicating rapid switches between memory-based saccade guidance and perceptual processing at fixation. Thus, the theta increase after revisits may reflect the mechanism by which contents attended to at multiple fixations become bound into a coherent representation. In the present study, the number of between-category *revisits* was predictive of event memory, whereas the EEG analysis contrasting between-category first visits vs. revisits failed to reveal any interaction with memory. This suggests that the between-category saccades for first visits and revisits are comparably important for the buildup of a coherent event representation in our paradigm. Alternatively, the EEG difference between first visits and revisits may have been attenuated by the deconvolution method, which eliminates the overlapping effects of sequential saccades on EEG and corrects the confounding effects of eye movement characteristics. In our analysis, the EEG in fixation intervals was also corrected for ordering factors, including fixation rank, which distinguishes between first visits and revisits and may affect the fixation-related EEG (Fischer et al., 2013; Kamienkowski et al., 2018). It remains unclear to what extent such potentially confounding factors have been controlled in the previous work.

The cumulative number of between-exemplar saccades that were irrelevant to the task did not predict memory. Nevertheless, alpha power at fixations after these saccades was lower for high than low subsequent memory over the centro-parietal and occipital areas (Fig. 7). Although this alpha effect thus reflects a mechanism that supports successful memory formation, its exact nature remains unclear. For example, in episodic memory encoding, alpha activity indicates cortical information processing during the perception of an event (Jensen and Mazaheri, 2010; Klimesch, 2012), and its decrease predicts successful formation of episodic memory (Hanslmayr et al., 2016). Specifically, alpha desynchronization is thought to increase the capacity of local cell ensembles to process information by reducing local synchronization (Hanslmayr et al., 2016). Thus, the occipital predominance of memory-related alpha desynchronization in our visual memory task may reflect an increase in the capacity of cell ensembles in the visual system, which is beneficial for category-

specific processing. Visual perception in our task involves discriminating between two similar exemplars from the same category and the between-exemplar saccades may facilitate a comparison of the two exemplars. Alpha activity has been shown to predict behavioral performance in visual discrimination tasks (Hanslmayr et al., 2005; Nelli et al., 2017). Thus, the here observed alpha desynchronization after between-exemplar saccades may indicate efficient discrimination of two similar exemplars that promote the successful encoding of the task-relevant associations.

Theta memory effects for within-element saccades

Another theta memory effect was observed at fixations after within-element saccades (Fig. 8), the number of which predicted subsequent memory (Fig. 3C). Remarkably, the theta memory effect for within-element saccades had a frontal maximum, whereas the theta effect for between-category saccades had a centro-parietal maximum (Fig. 9). The different topographies suggest distinct neural mechanisms underlying these theta effects.

Short within-element saccades, which serve to scrutinize visual features of images, increase visual sampling, which leads to better subsequent memory (Loftus, 1972; Olsen et al., 2016). Given this, we would have predicted within-element saccades to be associated with alpha desynchronization. This was however not found; instead, we found a frontal theta synchronization to be predictive of high event memory. Increased visual sampling results in high memory load, which is reflected in the high frontal theta power (Jensen and Tesche, 2002; Sederberg et al., 2006). In our paradigm, an increased memory load across encoding time corresponds to the accumulation of more visual details about event exemplars, which is favorable for later recall. Alternatively, frontal theta power has also been associated with the detection of interference during selective retrieval in episodic memory (Ferreira et al., 2014; Kerrén et al., 2021; Staudigl et al., 2010) and during multiple-list learning in working memory (Kliegl and Bauml, 2021). In the present study, encoding exemplars from multiple events may be accompanied by interference from exemplars of the same categories in previously acquired events. Therefore, detection of interference may be crucial for encoding distinct memories of each event and may trigger adaptive inhibition of exemplars from previous events, ensuring that each event is stored separately to succeed in the memory test. It is also conceivable that the frontal theta reflects the complementary effects of interference detection and memory load, as high memory load can lead to increased interference.

Conclusions

We used a novel method of EEG-eye movement coregistration in ecologically valid experiments, where participants were asked to remember a series of events by moving their eyes freely. In this way, we elucidated the mechanisms of episodic memory construction with a precision unavailable before. Our findings confirm a crucial role of eye movements in the formation of episodic memory (Voss et al., 2017). We demonstrate that the functional specialization of eye movements manifested in various types of saccades is associated with different perceptual and cognitive processes, as evidenced by fixation-locked theta and alpha EEG power. Specifically, two distinct processes reflected in theta synchronization and one process reflected in alpha desynchronization predict the success of episodic memory formation. Episodic memory about an event is constructed through rapid switching between these processes in a sequence of fixations. These processes may indicate the complementary functioning of the hippocampal theta system for binding episodes and the cortical alpha system for representing the content of these episodes (Hanslmayr et al., 2016). Our study extends previous literature by elucidating how these systems engage across eye movements in the service of building coherent episodic memories for personal experiences.

Acknowledgments

This work was supported by a grant from the Marcus and Amelia Wallenberg Foundation (MAW2015.0043) to MJ. We thank Benedikt Ehinger for his help with deconvolution and Axel Ekström for his assistance with data collection. We are grateful to Marcus Nyström and Diederick Niehorster for their help in developing the EEG-eye movement coregistration system.

References

Amit, R., Abeles, D., Bar-Gad, I., Yuval-Greenberg, S., 2017. Temporal dynamics of saccades explained by a self-paced process. Sci Rep 7, 886.

Backus, A.R., Schoffelen, J.M., Szebenyi, S., Hanslmayr, S., Doeller, C.F., 2016. Hippocampal-Prefrontal Theta Oscillations Support Memory Integration. Curr Biol 26, 450-457.

Brodeur, M.B., Dionne-Dostie, E., Montreuil, T., Lepage, M., 2010. The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research. Plos One 5.

Busch, N.A., VanRullen, R., 2010. Spontaneous EEG oscillations reveal periodic sampling of visual attention. Proc Natl Acad Sci U S A 107, 16048-16053.

Chang, C.Y., Hsu, S.H., Pion-Tonachini, L., Jung, T.P., 2020. Evaluation of Artifact Subspace Reconstruction for Automatic Artifact Components Removal in Multi-Channel EEG Recordings. Ieee Transactions on Biomedical Engineering 67, 1114-1121.

Chelnokova, O., Laeng, B., Eikemo, M., Riegels, J., Loseth, G., Maurud, H., Willoch, F., Leknes, S., 2014. Rewards of beauty: the opioid system mediates social motivation in humans. Molecular Psychiatry 19, 746-747.

Clouter, A., Shapiro, K.L., Hanslmayr, S., 2017. Theta Phase Synchronization Is the Glue that Binds Human Associative Memory. Current Biology 27, 3143-+.

Coco, M.I., Nuthmann, A., Dimigen, O., 2020. Fixation-related Brain Potentials during Semantic Integration of Object-Scene Information. Journal of Cognitive Neuroscience 32, 571-589.

Cohen, N.J., Eichenbaum, H., 1993. Memory, Amnesia, and The Hippocampal System. MIT Press, Cambridge, MA.

Damiano, C., Walther, D.B., 2019. Distinct roles of eye movements during memory encoding and retrieval. Cognition 184, 119-129.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134, 9-21.

Devillez, H., Guyader, N., Guerin-Dugue, A., 2015. An eye fixation-related potentials analysis of the P300 potential for fixations onto a target object when exploring natural scenes. Journal of vision 15, 20.

Dien, J., Santuzzi, A.M., 2005. Application of repeated measures ANOVA to high density ERP datasets: A review and tutorial. In: Handy, T.C. (Ed.), Event-related potentials: A methods handbook MIT Press, Cambridge, MA, pp. 57– 82.

Dimigen, O., 2020. Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments. Neuroimage 207, 116117.

Dimigen, O., Ehinger, B.V., 2021. Regression-based analysis of combined EEG and eye-tracking data: Theory and applications. Journal of vision 21.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A.M., Kliegl, R., 2011. Coregistration of eye movements and EEG in natural reading: analyses and review. Journal of Experimental Psychology: General 140, 552-572.

Ehinger, B.V., Dimigen, O., 2019. Unfold: an integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. PeerJ 7, e7838.

Engbert, R., Mergenthaler, K., 2006. Microsaccades are triggered by low retinal image slip. Proceedings of the National Academy of Sciences of the United States of America 103, 7192-7197.

Fell, J., Axmacher, N., 2011. The role of phase synchronization in memory processes. Nat Rev Neurosci 12, 105-118.

Ferreira, C.S., Marful, A., Staudigl, T., Bajo, T., Hanslmayr, S., 2014. Medial prefrontal theta oscillations track the time course of interference during selective memory retrieval. J Cogn Neurosci 26, 777-791.

Fischer, T., Graupner, S.T., Velichkovsky, B.M., Pannasch, S., 2013. Attentional dynamics during free picture viewing: evidence from oculomotor behavior and electrocortical activity. Frontiers in systems neuroscience 7, 17.

Fong, D.C.L., Saunders, M., 2011. Lsmr: An Iterative Algorithm for Sparse Least-Squares Problems. Siam Journal on Scientific Computing 33, 2950-2971.

Fudali-Czyz, A., Francuz, P., Augustynowicz, P., 2018. The Effect of Art Expertise on Eye Fixation-Related Potentials During Aesthetic Judgment Task in Focal and Ambient Modes. Frontiers in Psychology 9, 1972.

Gilchrist, I.D., Harvey, M., 2000. Refixation frequency and memory mechanisms in visual search. Current Biology 10, 1209-1212.

Griffiths, B.J., Parish, G., Roux, F., Michelmann, S., van der Plas, M., Kolibius, L.D., Chelvarajah, R., Rollings, D.T., Sawlani, V., Hamer, H., Gollwitzer, S., Kreiselmeyer, G., Staresina, B., Wimber, M., Hanslmayr, S., 2019. Directional coupling of slow and fast hippocampal gamma with neocortical alpha/beta oscillations in human episodic memory. Proc Natl Acad Sci U S A 116, 21834-21842.

Hanslmayr, S., Klimesch, W., Sauseng, P., Gruber, W., Doppelmayr, M., Freunberger, R., Pecherstorfer, T., 2005. Visual discrimination performance is related to decreased alpha amplitude but increased phase locking. Neurosci Lett 375, 64-68.

Hanslmayr, S., Staresina, B.P., Bowman, H., 2016. Oscillations and episodic memory: addressing the synchronization/desynchronization conundrum. Trends Neurosci 39, 16-25.

Hanslmayr, S., Staudigl, T., 2014. How brain oscillations form memories--a processing based perspective on oscillatory subsequent memory effects. Neuroimage 85 Pt 2, 648-655.

Herweg, N.A., Solomon, E.A., Kahana, M.J., 2020. Theta oscillations in human memory. Trends Cogn Sci 24, 208-227.

Hogendoorn, H., 2016. Voluntary saccadic eye movements ride the attentional rhythm. J Cogn Neurosci 28, 1625-1635.

Hsieh, L.T., Ranganath, C., 2014. Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. Neuroimage 85 Pt 2, 721-729.

Ito, J., Maldonado, P., Singer, W., Grün, S., 2011. Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. Cerebral Cortex 21, 2482-2497.

Jensen, O., Mazaheri, A., 2010. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front Hum Neurosci 4, 186.

Jensen, O., Tesche, C.D., 2002. Frontal theta activity in humans increases with memory load in a working memory task. Eur J Neurosci 15, 1395-1399.

Jutras, M.J., Buffalo, E.A., 2010. Synchronous neural activity and memory formation. Current Opinion in Neurobiology 20, 150-155.

Jutras, M.J., Fries, P., Buffalo, E.A., 2013. Oscillatory activity in the monkey hippocampus during visual exploration and memory formation. Proc Natl Acad Sci U S A 110, 13144-13149.

Kamienkowski, J.E., Varatharajah, A., Sigman, M., Ison, M.J., 2018. Parsing a mental program: Fixation-related brain signatures of unitary operations and routines in natural visual search. Neuroimage 183, 73-86.

Kazai, K., Yagi, A., 2003. Comparison between the lambda response of eye-fixation-related potentials and the P100 component of pattern-reversal visual evoked potentials. Cognitive, affective & behavioral neuroscience 3, 46-56.

Kerrén, C., Bramão, I., Hellerstedt, R., Johansson, M., 2021. Strategic retrieval prevents memory interference: The temporal dynamics of retrieval orientation. Neuropsychologia 154, 107776.

Kliegl, O., Bauml, K.H.T., 2021. The Mechanisms Underlying Interference and Inhibition: A Review of Current Behavioral and Neuroimaging Research. Brain Sciences 11.

Klimesch, W., 1996. Memory processes, brain oscillations and EEG synchronization. Int J Psychophysiol 24, 61-100.

Klimesch, W., 2012. alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn Sci 16, 606-617.

Klimesch, W., Schimke, H., Doppelmayr, M., Ripper, B., Schwaiger, J., Pfurtscheller, G., 1996. Event-related desynchronization (ERD) and the Dm effect: does alpha desynchronization during encoding predict later recall performance? Int J Psychophysiol 24, 47-60.

Körner, C., Braunstein, V., Stangl, M., Schlogl, A., Neuper, C., Ischebeck, A., 2014. Sequential effects in continued visual search: using fixation-related potentials to compare distractor processing before and after target detection. Psychophysiology 51, 385-395.

Körner, C., Gilchrist, I.D., 2008. Memory processes in multiple-target visual search. Psychological Research 72, 99-105.

Kota, S., Rugg, M.D., Lega, B.C., 2020. Hippocampal Theta Oscillations Support Successful Associative Memory Formation. J Neurosci 40, 9507-9518.

Kothe, C.A.E., Jung, T.P., 2016. Artifact removal techniques with signal reconstruction. U.S. Patent Application. 14/895,440.

Kragel, J.E., Schuele, S., VanHaerents, S., Rosenow, J.M., Voss, J.L., 2021. Rapid coordination of effective learning by the human hippocampus. Sci Adv 7.

Kragel, J.E., VanHaerents, S., Templer, J.W., Schuele, S., Rosenow, J.M., Nilakantan, A.S., Bridge, D.J., 2020. Hippocampal theta coordinates memory processing during visual exploration. Elife 9.

Landau, A.N., Fries, P., 2012. Attention samples stimuli rhythmically. Curr Biol 22, 1000-1004.

Liu, Z.X., Shen, K., Olsen, R.K., Ryan, J.D., 2017. Visual Sampling Predicts Hippocampal Activity. J Neurosci 37, 599-609.

Loftus, G.R., 1972. Eye fixations and recognition memory for pictures. Cognitive Psychology 3, 525-551.

McCarthy, G., Wood, C.C., 1985. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. Electroencephalography and Clinical Neurophysiology 62, 203-208.

Meghanathan, R.N., Nikolaev, A.R., van Leeuwen, C., 2019. Refixation patterns reveal memory-encoding strategies in free viewing. Atten Percept Psychophys 81, 2499-2516.

Meghanathan, R.N., van Leeuwen, C., Giannini, M., Nikolaev, A.R., 2020. Neural correlates of task-related refixation behavior. Vision Research 175, 90-101.

Mirza, M.B., Adams, R.A., Mathys, C.D., Friston, K.J., 2016. Scene Construction, Visual Foraging, and Active Inference. Frontiers in Computational Neuroscience 10.

Nelli, S., Itthipuripat, S., Srinivasan, R., Serences, J.T., 2017. Fluctuations in instantaneous frequency predict alpha amplitude during visual perception. Nat Commun 8, 2071.

Niehorster, D.C., Andersson, R., Nystrom, M., 2020. Titta: A toolbox for creating PsychToolbox and Psychopy experiments with Tobii eye trackers. Behavior Research Methods 52, 1970-1979.

Nikolaev, A.R., Meghanathan, R.N., van Leeuwen, C., 2016. Combining EEG and eye movement recording in free viewing: pitfalls and possibilities. Brain and Cognition 107, 55-83.

Nikolaev, A.R., Meghanathan, R.N., van Leeuwen, C., 2018. Refixation control in free viewing: a specialized mechanism divulged by eye-movement related brain activity. Journal of Neurophysiology 120, 2311-2324.

Nikolaev, A.R., Nakatani, C., Plomp, G., Jurica, P., van Leeuwen, C., 2011. Eye fixation-related potentials in free viewing identify encoding failures in change detection. Neuroimage 56, 1598-1607.

Olsen, R.K., Sebanayagam, V., Lee, Y., Moscovitch, M., Grady, C.L., Rosenbaum, R.S., Ryan, J.D., 2016. The relationship between eye movements and subsequent recognition: Evidence from individual differences and amnesia. Cortex 85, 182-193.

Ossandón, J.P., Helo, A.V., Montefusco-Siegmund, R., Maldonado, P.E., 2010. Superposition model predicts EEG occipital activity during free viewing of natural scenes. Journal of Neuroscience 30, 4787-4795.

Ossandón, J.P., König, P., Heed, T., 2020. No Evidence for a Role of Spatially Modulated alpha-Band Activity in Tactile Remapping and Short-Latency, Overt Orienting Behavior. Journal of Neuroscience 40, 9088-9102.

Otero-Millan, J., Troncoso, X.G., Macknik, S.L., Serrano-Pedraza, I., Martinez-Conde, S., 2008. Saccades and microsaccades during visual fixation, exploration, and search: foundations for a common saccadic generator. J Vis 8, 21 21-18.

Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Hochenberger, R., Sogo, H., Kastman, E., Lindelov, J.K., 2019. PsychoPy2: Experiments in behavior made easy. Behavior Research Methods 51, 195-203.

Pierrot-Deseilligny, C., Milea, D., Muri, R.M., 2004. Eye movement control by the cerebral cortex. Current Opinion in Neurology 17, 17-25.

Plöchl, M., Ossandón, J.P., König, P., 2012. Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. Frontiers in human neuroscience 6, 278.

Rajkai, C., Lakatos, P., Chen, C.M., Pincze, Z., Karmos, G., Schroeder, C.E., 2008. Transient cortical excitation at the onset of visual fixation. Cerebral Cortex 18, 200-209.

Ries, A.J., Touryan, J., Ahrens, B., Connolly, P., 2016. The Impact of Task Demands on Fixation-Related Brain Potentials during Guided Search. PLoS One 11, e0157260.

Ryan, J.D., Shen, K., Liu, Z.X., 2020. The intersection between the oculomotor and hippocampal memory systems: empirical developments and clinical implications. Annals of the New York Academy of Sciences 1464, 115-141.

Sederberg, P.B., Gauthier, L.V., Terushkin, V., Miller, J.F., Barnathan, J.A., Kahana, M.J., 2006. Oscillatory correlates of the primacy effect in episodic memory. Neuroimage 32, 1422-1431.

Shen, K., Bezgin, G., Selvam, R., McIntosh, A.R., Ryan, J.D., 2016. An Anatomical Interface between Memory and Oculomotor Systems. Journal of Cognitive Neuroscience 28, 1772-1783.

Staudigl, T., Hanslmayr, S., 2013. Theta oscillations at encoding mediate the context-dependent nature of human episodic memory. Curr Biol 23, 1101-1106.

Staudigl, T., Hanslmayr, S., Bauml, K.H., 2010. Theta oscillations reflect the dynamics of interference in episodic memory retrieval. J Neurosci 30, 11356-11362.

Tatler, B.W., Gilchrist, I.D., Land, M.F., 2005. Visual memory for objects in natural scenes: from fixations to object files. Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology 58, 931-960.

Team, R.C., 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Thickbroom, G.W., Knezevic, W., Carroll, W.M., Mastaglia, F.L., 1991. Saccade onset and offset lambda waves: relation to pattern movement visually evoked potentials. Brain Research 551, 150-156.

Tulving, E., 1983. Ecphoric Processes in Episodic Memory. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 302, 361-371.

Turk-Browne, N.B., 2019. The hippocampus as a visual area organized by space and time: A spatiotemporal similarity hypothesis. Vision Research 165, 123-130.

Tyson-Carr, J., Soto, V., Kokmotou, K., Roberts, H., Fallon, N., Byrne, A., Giesbrecht, T., Stancak, A., 2020. Neural underpinnings of value-guided choice during auction tasks: An eye-fixation related potentials study. Neuroimage 204, 116213.

Voss, J.L., Bridge, D.J., Cohen, N.J., Walker, J.A., 2017. A Closer Look at the Hippocampus and Memory. Trends in cognitive sciences 21, 577-588.

Voss, J.L., Warren, D.E., Gonsalves, B.D., Federmeier, K.D., Tranel, D., Cohen, N.J., 2011. Spontaneous revisitation during visual exploration as a link among strategic behavior, learning, and the hippocampus. Proc Natl Acad Sci U S A 108, E402-409.

Waldhauser, G.T., Johansson, M., Hanslmayr, S., 2012. Alpha/beta oscillations indicate inhibition of interfering visual memories. J Neurosci 32, 1953-1961.

Wynn, J.S., Shen, K., Ryan, J.D., 2019. Eye Movements Actively Reinstate Spatiotemporal Mnemonic Content. Vision (Basel) 3.

Yagi, A., 1979. Saccade size and lambda complex in man. Physiological Psychology 7, 370-376.