Research Report

Behavioural and neurophysiological signatures in the retrieval of individual memories of recent and remote real-life routine episodic events

Berta Nicolás [a,b,c], *Xiongbo Wu* [a,b,c], *Josué García-Arch* [a,b,c],
Mariella Dimiccoli [d], *Joanna Sierpowska* [e,l], *Cristina Saiz-Masvidal* [f,g],
Carles Soriano-Mas [f,h,i], *Petia Radeva* [j,k] *and Lluís Fuentemilla* [a,b,c,*]

[a] *Cognition and Brain Plasticity Group, Bellvitge Biomedical Research Institute-IDIBELL, Hospitalet de Llobregat, Spain*
[b] *Department of Cognition, Development and Educational Psychology, University of Barcelona, Barcelona, Spain*
[c] *Institute of Neurosciences, University of Barcelona, Spain*
[d] *Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain*
[e] *Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, the Netherlands*
[f] *Bellvitge Biomedical Research Institute-IDIBELL, Barcelona, Spain*
[g] *Department of Clinical Sciences, University of Barcelona, Spain*
[h] *CIBER Salud Mental (CIBERSAM), Spain*
[i] *Department of Psychobiology and Methodology in Health Sciences, Universitat Autònoma de Barcelona, Spain*
[j] *Department of Mathematics and Computer Science, University of Barcelona, Spain*
[k] *Computer Vision Center, Spain*
[l] *Radboud University Medical Center, Donders Institute for Brain Cognition and Behaviour, Department of Medical Psychology, Nijmegen, the Netherlands*

ARTICLE INFO

ABSTRACT

Autobiographical memory (AM) has been largely investigated as the ability to recollect specific events that belong to an individual's past. However, how we retrieve real-life routine episodes and how the retrieval of these episodes changes with the passage of time remain unclear. Here, we asked participants to use a wearable camera that automatically captured pictures to record instances during a week of their routine life and implemented a deep neural network-based algorithm to identify picture sequences that represented episodic events. We then asked each participant to return to the lab to retrieve AMs for single episodes cued by the selected pictures 1 week, 2 weeks and 6—14 months after encoding while scalp electroencephalographic (EEG) activity was recorded. We found that participants were more accurate in recognizing pictured scenes depicting their own past than pictured scenes encoded in the lab, and that memory recollection of personally experienced events rapidly decreased with the passing of time. We also found that the retrieval of real-life picture cues elicited a strong and positive 'ERP old/new effect' over frontal regions and that the magnitude of this ERP effect was similar throughout memory

tests over time. However, we observed that recognition memory induced a frontal theta power decrease and that this effect was mostly seen when memories were tested after 1 and 2 weeks but not after 6—14 months from encoding. Altogether, we discuss the implications for neuroscientific accounts of episodic retrieval and the potential benefits of developing individual-based AM exploration strategies at the clinical level.

1. Introduction

Autobiographical memories (AMs) are specific individualized compilations of our personal past daily life episodic experiences. The ability to recollect detailed information about past autobiographical events is a hallmark of episodic memory (Tulving, 2002). However, the vast majority of behavioural and neuroimaging studies of episodic retrieval have used laboratory-encoded stimuli, such as words or pictures, as memory probes. While such stimuli provide researchers with tight experimental control over the perceptual qualities, exposure duration, and retention interval of the events being tested, laboratory stimuli lack the richness of most real-world experiences (Chow et al., 2018; Chow & Rissman, 2017; Diamond & Levine, 2020; Nielson et al., 2015; St. Jacques et al., 2011). Thus, it is not unsurprising that performance on standard laboratory-based memory tasks may be largely unrelated to one's autobiographical retrieval abilities, as demonstrated by individuals with "highly superior autobiographical memory" (LePort et al., 2012, 2017; Patihis et al., 2013) and "severely deficient autobiographical memory" Fuentemilla, Palombo, & Levine, 2018; Palombo et al., 2015).

A hallmark in the advancement of our understating of how episodic memory serves to retrieve real-life autobiographical experiences would be to have methods that allowed the automatic recording of daily life episodes prospectively at the individual level. Through such an approach, researchers would have the opportunity to examine an exhaustive collection of realistic real-life experience material of an individual ahead of sampling control during encoding. Previous research efforts proved effective in cueing AMs sampled at the individual level. Most notably, the use of self-recorded audiotapes or videos documenting selected real-life event experiences has helped characterize the involvement of a core brain network supporting the retrieval of AMs including the medial temporal lobe and the frontal and parietal regions (Levine et al., 2004; Svoboda & Levine, 2009), coordinated via neural oscillatory mechanisms in the range of the theta band (4—8 Hz) (Fuentemilla et al., 2014). However, this approach requires individuals to actively record selected experiences during their daily life routine, and therefore the effectiveness of the retrieval cues may still be partially explained by additional processes engaged during encoding such as selection, organization, and rehearsal of the recorded material.

The recent incorporation of portable technology, such as wearable cameras, to study the cognitive and neural basis of AM retrieval appeared to be a promising venue for addressing the previous concern. This technology allows the automatic capture (e.g., every 30 sec) of face-front sequence of pictures of daily life activity without the need for the participant to be actively engaged in the recording process. Researchers have already shown that the presentation of pictures acquired with a portable camera engaged the core AM retrieval network (Cabeza et al., 2004; Rissman et al., 2016) and elicited a strong sense of first-person retrieval in participants, even when they were confronted with others' pictures depicting the same content (St. Jacques et al., 2011). The use of pictures collected from a wearable camera has also been shown to be a valuable approach to enhance AM retrieval in healthy young and older adults (Chow & Rissman, 2017; Xu et al., 2020) and in patients with memory impairments (Alle et al., 2017), with Alzheimer's disease (Woodberry et al., 2014) and with limbic encephalitis (Berry et al., 2007, 2009). While the specific neural mechanisms that make picture cues taken in the real world by participants such a powerful retrieval cue remains elusive, functional neuroimaging studies highlighted that this type of cue may engage neural correlates of processes that are difficult to study using laboratory stimuli, including complex constructive processes, recollective qualities of emotion and vividness, and remote memory retrieval (Cabeza & St Jacques, 2007).

In the current study, we sought to use this technology to investigate the neurophysiological underpinnings supporting the retrieval of individual AMs over the passage of time. To this end, we asked healthy participants to retrieve their own AMs cued by pictures taken automatically (i.e., every 30 sec) by a wearable camera carried during one week of daily life routine, after 1—2 weeks, and 6—14 months from the encoding period. To ensure pictures presented during the test cued most of the episodic events that unfolded during the encoding week, we implemented a convolutional network-based algorithm (Dimiccoli et al., 2015) on the entire recorded picture set that automatically grouped together temporally adjacent images sharing contextual and semantic attributes, akin to how we conceive what underlies an event episode from a perception and memory perspective (Zacks & Swallow, 2007; see also Jeunehomme & D'Argembeau, 2018 and 2019). In doing so, the large picture set collected reflecting an entire day's life activity (e.g., ~400 pictures) is grouped into a workable number of picture subsets (e.g., ~20) depicting sequences of temporally adjacent episodic events (e.g., breakfast at home, commuting to work, buying oranges in the corner shop, eating a sandwich at the park). We reasoned that by picking a representative picture from each of the subsets, it would then be possible to investigate whether an individual is capable of retrieving information about a single past episodic event. Additionally, the same participants were asked to enrol in a separate study that required them to encode and retrieve, one week after the fact,

pictures depicting indoor and outdoor scenes. This task, akin to standard lab-based experimental scenarios commonly used in memory research, was thought to help delineate differences between retrieval processes for when participants' memory was cued by real-life autobiographical versus lab-based event experiences.

Complementing the behavioural data, we also aimed to analyse well-known neural response activity widely studied in the context of recognition memory in humans, the Event-Related Potentials (ERPs) and neural oscillations. ERPs have been employed to study the neural correlates of successful retrieval since the early 1990s. These studies have consistently identified a retrieval-related effect that takes the form of more positive-going ERPs for correctly classified old recognition test items relative to new items. In general, ERPs are more positive 300 msec after presentation of correctly recognized studied items compared to correctly reject non-studied items. This difference has been termed the 'ERP old/new effect'. Several groups have found spatially and temporally distinct ERP modulations within the old/new effect that appear to be specifically associated with the retrieval of items with contextual information (i.e., recollection) or with a lack of such contextual details (i.e., familiarity). An early (~300—500 msec) component, distributed over frontal electrodes, is correlated with familiarity while a later (~500—800 msec) component, distributed over parietal electrodes, appears modulated by recollection (Curran, 2000; Curran et al., 2001; Duzel et al., 1997; Wilding & Rugg, 1996). An additional later (~600—1600 msec) and positive sustained ERP modulation, in most prominent frontal scalp sites, has been described in recognition memory studies. This modulation has been observed in a number of studies and is thought to be related to post-retrieval processing (Allan & Rugg, 1997; Curran et al., 2001; Donaldson & Rugg, 1999; Ranganath & Paller, 2000; Wolk, 2006). In the context of the current study, we hypothesized that ERP signatures could provide information about the participants' ability to retrieve contextual information associated with individual real-life episodic events and how the retrieval of these memory events changed over time from encoding.

Indeed, the notion that memory representations change over time has concerned psychologists and neuroscientists for decades. A widely accepted view is that initial encoding of experiences renders temporary, labile memory and that it may become transformed into a more stable, long-lasting form via systems consolidation mechanisms. Systems consolidation refers to the gradual reorganization of the brain systems that support memory (Dudai & Morris, 2013; Squire & Alvarez, 1995). By this process, the hippocampus gradually becomes less important for storage and retrieval, and a more permanent memory develops in distributed regions of the neocortex. This view has received support from studies showing that memory recollection is supported by the coordinated activity of hippocampus and neocortical regions via neural oscillations at the theta band (~3—8 Hz) (Fuentemilla et al., 2014; Nyhus & Curran, 2010; Herweg et al., 2016). While this perspective aligns well with the findings that successful retrieval would be accompanied by an increase in theta power during retrieval (Burgess & Gruzelier, 1997; Gruber et al., 2008; Guderian & Düzel, 2005; Klimesch et al.,

2001; Osipova et al., 2006), recent studies suggested that this picture remains controversial (see, Herweg, Solomon, Kahana, & Kahana, 2020 for a recent discussion on this topic) as other studies have shown the opposite effect, that theta power decreased during memory retrieval (Hanslmayr et al., 2010; Khader & Rösler, 2010; Staudigl et al., 2010). Thus, while the specificity of the theta rhythm in retrieving past memories is a prevalent one in the literature, many questions still remain unclear. In the current study, we sought to investigate the role of theta rhythm during retrieval by examining neural oscillatory power changes upon successful retrieval of individual AM from real life.

If our experimental approach is suitable to examine AM retrieval, we expect that participant's ability to recollect detailed and rich contextual information from specific encoded experienced episodes to decrease with the passage of time. To the extent that the hypothesized behavioural effects were related to neural mechanisms associated with memory retrieval, we expect they should be accompanied by changes at the ERPs and neural oscillations at the theta band level.

2. Material and methods

2.1. Participants

Sixteen healthy participants (8 females) participated in experiments 1 and 2 (see Rissman et al., 2016 and Bonnici et al., 2012 for similar sample sizes). The range in age was 22-37 years old (mean = 27.68, SD = 4.22). All participants provided informed written consent for the protocol approved by the Ethics Committee of the University of Barcelona. Participants received financial compensation for their participation. Two of the participants could not complete the follow-up test in experiment 1 (see details below) and these participants were excluded from all analyses. We report how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. However, no part of the study procedures or analyses was pre-registered prior to the research being conducted.

2.2. Experiment 1: retrieval of real-life memories

2.2.1. Design overview
Participants were asked to come to the training session a few days before starting the study. They were informed about how the camera worked and they read, understood, and signed the informed consent. We took special care in providing details on privacy issues so that all participants were fully aware of them before the study began. In the current study, participants wore the wearable camera for a period of 5—7 consecutive days (mean = 6.64 days, SD = .63), including both week and weekend days. The data collection period was established from morning to evening (between 12 and 14 h per day). Once participants finished data collection, they were requested to return the materials to be processed. Participants confirmed not having checked the
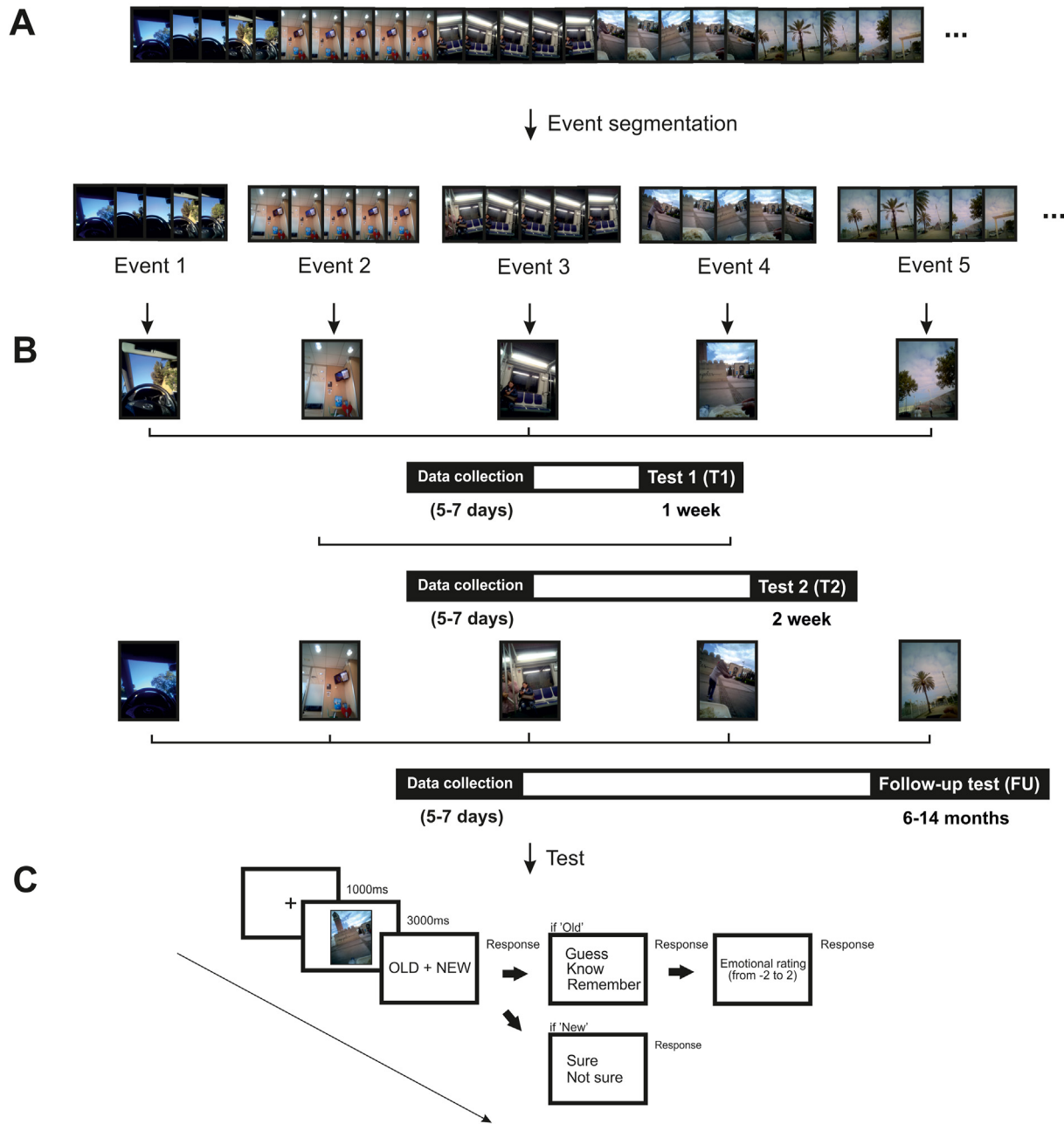
Fig. 1 – Experimental design. (A) Example of a stream of pictures obtained from one participant after one week of data collection. The implementation of the SR-Clustering algorithm allows the automatic organization of picture sequences into a set of meaningful events by the identification of similar context and semantic features. (B) Experimental design. One representative photo from each event sequence of pictures was selected and distributed to each of the three memory tests that followed the picture collection week. A recognition memory test was implemented one week (T1), two weeks (T2) and from 6 to 14 months (Follow-up test (FU)) after the data collection. Events extracted from picture sequences were numbered consecutively and pictures related to even numbered events were used as memory cues in T1 test and odd numbered event pictures in T2 test. Different pictures from same events tested in T1 and T2 were selected as cues in FU test. (C) Recognition memory task. Pictures were presented on the screen for 3000 msec. Afterward, an 'Old/New' question appeared on the screen. If the pictures were seen as 'Old', a "Remember/Know/Guess" appeared on the screen. Next, participants were asked to rate their emotional response towards the event indicating whether it was positive or negative on a scale ranging from 2 (maximum positive) to −2 (maximum negative), with 0 being neutral emotion. If the picture was seen as 'New', a 'Sure/Not sure' judgment appeared on the screen.

pictures during the encoding week. Participants returned to the lab to be tested one week later (T1), two weeks later (T2), and from 6 to 14 months after the last day of data collection (Follow-up, hereafter FU; mean = 10.87 months, SD = 2.02 months). See Fig. 1 for a summary of the experimental design.

2.2.2. Wearable camera

We used the wearable Narrative clip 2 camera® (http://getnarrative.com/) with a camera sensor of 8 MP and a resolution of 3264 × 2448(4:3). The camera was programmed to automatically take images every thirty seconds and produced pictures with an egocentric viewpoint. Participants were instructed to wear the device on a lanyard around the neck. Narrative clip 2 incorporated a downloading app that allowed participants to download the pictures directly to a hard drive. Participants were instructed to not watch the pictures until the experiment finalized. None of the participants reported having done so at the end of the study.

2.2.3. Picture selection

We implemented a deep neural network-based algorithm, SR-Clustering, to automatically organize the stream of each participant's pictures into a set of temporally evolving meaningful events (Dimiccoli et al., 2015). The algorithm segments picture sequences into discrete events (e.g., having breakfast in a kitchen, commuting to work, being in a meeting) based on its ability to identify similar contextual and semantic features from the picture stream.

The implementation of the SR-Clustering algorithm provided boundaries for a variable number of discrete events for each participant per day, during which 8 to 20 pictures were taken. Each participant's events were then manually inspected and those which displayed non-meaningful episodes (e.g., all pictures were blurred, or when the camera was pointing to the roof or was blocked by clothes) were discarded from the study. Picture events that included faces from peer interactions were excluded from the study to avoid the use of personally relevant memory cues. Three independent experimenters rated and selected the set of event pictures for each participant on the basis of these criteria, and only those events that were consistently selected by the three raters were included in the final set of picture events in the study. Note that the consistency across experimenters was set to ensure that the events captured by the algorithm were meaningful and did not involve implementing a subjective inclusion/exclusion selection criterion to which events should be included later in the memory test by the experimenters. The variability observed in the number of events between participants reflected the diversity of each participant's daily life activities (e.g., a person working indoors for 8 h results in fewer events compared with people working outdoors).

Once the images were organized into discrete events, we selected a representative picture from each event; thereby ensuring most of the past episodic experience was brought into the test. We then numbered the sequence of event pictures and assigned even-numbered pictures to be used as memory cues for test T1 and odd-numbered pictures to be included in the T2 test (Table 1). FU included picture cues used in T1 and T2 in the

Table 1 — Total number of photos presented for each participant and retrieval test (Experiment 1).

Participant	Total number of photos in Test 1	Total number of photos in Test 2	Total number of photos in Follow-up test
1	118	118	192
2	168	158	110
3	200	212	118
4	198	192	80
5	194	196	166
6	246	246	190
7	198	188	124
8	256	210	168
9	204	240	144
10	290	156	138
11	252	188	158
12	238	228	180
13	160	180	142
14	222	238	120
Mean (SD)	210 (45)	196 (36)	145 (33)

same proportion. Pictures cues presented to one participant depicting her own past (Old) were also presented to another participant as New images. For each participant, the number of New and Old images was roughly similar. In cases where the number of New images exceeded the number of Old images, we randomly selected a subset of New images. This ensured that differences between Old and New pictures presented to each participant were only based on the image's direct link to ones' past while preserving the rest of the characteristics intact during the test (e.g., angle of view, picture image features, description of routine daily life activities). Old and New images were presented in random order during the experiment. See Fig. 1B for a summary of the experimental design.

By design, none of the participants were friends with each other, and we never encountered an instance where two concurrently enrolled participants came into direct contact with one another while wearing their cameras.

2.3. Recognition memory task

In the test, pictures were presented on the screen for 3000 msec. Afterwards, when an "Old/New" question appeared on the screen, participants were required to judge whether the picture reflected an event from the participant's own daily life (Old) or was experimentally novel, signalling with the right index and middle fingers, respectively. Next, participants were asked to judge whether they were "Sure/Not sure" when indicated that an image was "New" and "Remember/Know/Guess" when images were seen as "Old". Participants were instructed that "Guess" referred to when they had no contextual memory reference for what was depicted in the image, but they recognized the content as being from their own life (e.g., viewing one's living room). "Know" was the signal for when the visual content in the picture was highly familiar but the subject could not determine what unfolded in it, perhaps because the event in the test was part of a routine (e.g., playing football on Thursdays), while "Remember" was the signal for when the picture elicited a vivid memory of that specific event and it could be located in time. Participants' ability to order each event depicted

in the pictures along the encoding week was tested afterward more concretely, when they were asked to indicate whether the pictures depicted an event that took place at the "beginning, middle or end" of the encoding week as the image appeared on the screen. Finally, to explore the data further, participants were asked to rate the degree to which each of the pictures elicited an emotional response and to indicate whether it was positive or negative on a scale that ranged from 2 (maximum positive) to −2 (maximum negative), with 0 being neutral emotionally. See Fig. 1C for a summary of the recognition memory task. Temporal ratings are not shown in the design overview.

2.4. Experiment 2: retrieval of lab-based memories

The experimental design was similar to experiment 1, but differed in that images depicting real-life experiences were replaced by neutral images of indoor and outdoor scenes extracted from previous experiments (e.g., Bunzeck & Düzel, 2006; Fuentemilla et al., 2010). The experimental design involved an encoding phase and a test phase administered after each participant finished the retrieval session at the FU session in experiment 1.

In the encoding phase, participants were instructed to indicate whether scene pictures were indoor or outdoor images. There were 80 scenes (40 indoor and 40 outdoor, presented in random order). Each scene was presented for 2000 msec preceded by a 1500 msec fixation period and followed by the text "indoor/outdoor" that prompted participants' response (responding with the index or middle finger of their right hand). A period of 10 min of rest separated the study from the test phase.

In the test phase, a scene picture was presented on the screen for 3000 msec. Afterwards, when an "Old/New" question appeared on the screen, participants were required to judge whether the word was presented in the previous study phase (Old) or was experimentally novel (New) with the right index and middle finger, respectively. The test phase included 160 scene images in total (80 Old and 80 New, randomly presented). Thereafter, as in experiment 1, a confidence judgment task followed. Here, new judgments were followed by "Sure/Not sure" and old judgments were followed by "Remember/Know/Guess". Participants were instructed to make confidence judgments following old judgments with respect to their ability to vividly retrieve the contextually associated information related to the image during encoding. They were instructed to respond "Guess" when they were unsure about their previous Old judgment, "Know" when they recognized the scene image but could not retrieve any contextual feature linked to it, and "Remember" when the scene image brought a vivid recollection of the specific context that surrounded the encoding of that particular image during encoding.

2.5. Behavioural data analysis

A repeated measures Analysis of Variance (ANOVA) for hits (i.e., Correct Old responses) and Correct Rejections (i.e., Correct New responses), including time of the test (T1, T2 and FU) as a within-subject factor, was implemented to assess for statistical differences. Paired t-test comparisons

were used as a post-hoc test. Significance threshold was set at $p < .05$.

Given the dependency between category responses in confidence judgment (Guess, Know, Remember) (Experiments 1 and 2) and emotionality (Experiment 1) measures, we performed a Bayesian ordinal regression fitted using Monte Carlo Markov Chains algorithm via the brms (Bayesian Regression Models using Stan) package in R (Bürkner, 2017) to assess for changes in the likelihood of participants' responses as a function of time of the test. In order to deal with the dependency between observations from the same participants and to accommodate the repeated measures study design, we used a multilevel approach within the Bayesian ordinal regression analysis, where time was considered a constant effect and participant was considered a varying effect [commonly known as fixed and random effects but see recommendations about this terminology in Gelman and Hill (2006)]. In addition, we could not assume that our predictor (time) would have the same effect on all response categories (e.g., neutral emotional ratings could increase over time while positive emotional ratings decrease). This was explicitly modelled by allowing for category specific effects (CSE), which imply estimating as many regression coefficients per category specific predictor as possible thresholds (C − 1 = 2 in our case). However, given that it has been suggested that fitting CSE in cumulative models is problematic (Bürkner & Vuorre, 2019), we used an adjacent category model instead.

For both models, the Multilevel ordinal Bayesian regression was computed with 4 chains [2000 iterations per chain (warmup = 1000) and weakly informative priors] (Bürkner, 2017). Both models were compared to their respective non-multilevel version (without including participants as varying effect) and with their respective null models (intercepts only) with an approximate leave-one-out cross-validation, in which interpretability is similar than the one used to interpret the Akaike's information criterion (Bürkner & Vuorre, 2019).

2.6. EEG recordings and preprocessing

EEG was recorded at a 500 Hz sampling rate (High-pass filter at .016 Hz, notch filter at 50 Hz) from the scalp using a BrainAmp amplifier tin electrodes mounted in an electrocap (Electro-Cap International) located at 29 standard positions (Fp1/2, Fz, F7/8, F3/4, FCz, FC1/2, FC5/6, Cz, C3/4, T3/4, Cp1/2, Cp5/6, Pz, P3/4, T5/6, PO1/2, Oz) and at the left and right mastoids. An electrode placed at the lateral outer canthus of the right eye served as an online reference. EEG was re-referenced offline to the linked mastoids. Vertical eye movements were monitored with an electrode at the infraorbital ridge of the right eye (EOG channel). Electrode impedances were kept below 3 kΩ. EEG was band-pass filtered offline at .1 - 40 Hz. Independent Component Analysis (Delorme & Makeig, 2004) was applied to the continuous EEG data to remove blinks and eye movement artefacts. EEG data from two participants were lost due to technical problems and were not able to be included in the rest of the EEG analysis.

2.7. Event-related potentials (ERPs) analysis

The continuous sample EEG data were then epoched into 3100 msec segments (0−3000 msec relative to trial onset), and

the pre-stimulus interval (−100 to 0 msec) was used as the baseline for baseline removal procedure. Trials exceeding ±100 μV in EEG and/or EOG channels within −100 to 3000 msec time window from stimulus onset were rejected offline and not used in ERPs and time-frequency analysis (see details below). For each participant, we obtained trial epochs that were correctly classified by that participant as either Old or New. We were unable to analyse ERP data for "Remember/Know/Guess" responses separately because of the low number of trials in some of the experimental conditions (please see results below).

2.8. Time-frequency analysis

The power of neural oscillatory activity was calculated by means of the continuous complex Morlet wavelet. It is a biologically plausible wavelet modulated by a Gaussian function which depends on the number of cycles the sinusoidal wave segment comprises. In the current study, the cycles of the Morlet wavelets used for convolution ranged from 4 to 10, increasing logarithmically as frequency increased. We adopted this modified wavelet approach to optimize the trade-off between the temporal resolution at lower frequency band and the frequency resolution at the higher frequency band. For all conditions in experiment 1 and 2, time-frequency analysis was carried out on for each participant and at single trial basis for hits in the memory tests, with epochs of 3500 msec time-locked to the presentation of photo starting at 500 msec before its onset. The convolution with Morlet wavelet was conducted for each frequency value from 1 Hz to 40 Hz, with 50 steps increasing logarithmically. Power values for each frequency were averaged across trials for each channel and then baseline-corrected by decibel conversion.

2.9. Cluster-based statistics of the ERP and time-frequency data

To account for ERP differences elicited by Old and New pictures, a cluster-based permutation test was used (Maris & Oostenveld, 2007) to identify clusters of significant points in the resulting spatiotemporal 2D matrix (time and electrodes) in a data-driven manner and addressing the multiple-comparison problem by employing a nonparametric statistical method based on cluster-level randomization testing to control for the family-wise error rate. Statistics were computed for each time point, and the spatiotemporal points whose statistical values were larger than a threshold ($p < .05$, two-tail) were selected and clustered into connected sets on the basis of x, y adjacency in the 2D matrix. The observed cluster-level statistics were calculated by taking the sum of the statistical values within a cluster. Then condition labels were permuted 1000 times to simulate the null hypothesis, and the maximum cluster statistic was chosen to construct a distribution of the cluster-level statistics under the null hypothesis. The nonparametric statistical test was obtained by calculating the proportion of randomized test statistics that exceeded the observed cluster-level statistics.

To assess for differences between Old and New conditions at the time-frequency level a similar statistical approach was adopted. However, clusters ($p < .05$, two-tail) were determined by connected sets of data samples that were contiguous on the basis of temporal, frequency, or spatial adjacency in the 3D matrix. Cluster statistics and null distribution were created following the same approach as for the ERP statistical approach.

3. Results

3.1. Experiment 1

3.1.1. Behavioural results
Participants were highly accurate in correctly distinguishing pictures that depicted their own past (Old pictures) from those that belonged to others' past (New pictures) (Fig. 2A). False Alarms [T1 test: Mean (M) = .05, Standard Deviation (SD) = .05; T2 test: M = .03, SD = .03; FU test: M = .03, SD = .02] and Omissions (T1 test: M = .08, SD = .06; T2 test: M = .09, SD = .06; FU test: M = .11, SD = .09) were very rare in all tests. However, a repeated measures ANOVA, including the three-memory test (T1, T2, and FU) as a within-subject factor in the analysis, revealed that hit rate differed significantly across them [$F(2,26) = 4.94$, $p = .01$] (Fig. 2A). A series of paired t-test comparisons showed that hit rate decreased as a function of time from encoding. Thus, significant differences were found when T1 and FU hit rate were compared [$t(13) = −2.68$, $p = .02$] but not for T1 and T2 [$t(13) = 1.37$, $p = .19$], nor T2 and FU [$t(13) = 1.93$, $p = .07$]. These differences cannot be accounted for by a general decrease in performance over time as participants' ability to identify New images (i.e., correct rejections) was similar in the three tests [$F(1.29, 16.85) = 2.14$, $p = .16$]. Correct rejections: T1 test: M = .94, SD = .05; T2 test: M = .97, SD = .03; FU test: M = .97, SD = .02.

Participants' confidence judgements for Hits for each of the memory tests are displayed in Fig. 2B. The β estimates for the model predicting differences in participants' confidence judgements as a function of time of the memory test [T1 (reference), T2, FU] included the mean, the standard error (SE) and the 95% credible intervals (CrI) of the posterior distribution of each parameter of interest. Category-specific effects are reported for changes over time on participants' responses for "Know" versus "Guess" and for "Remember" versus "Know". The results showed that the likelihood of "Know" versus "Guess" responses increased [β = .44, SE = .11, CrI (.22, .66)] but that "Remember" versus "Know" responses decreased in T2 compared to T1 [β = −.32, SE = .07, CrI (-.45, −.18)]. Similarly, the likelihood that participants judged their responses as "Know" versus "Guess" increased in FU test [β = .27, SE = .11, CrI (.05, .48)] and so did decrease the likelihood of "Remember" versus "Know" when compared to T1 [β = −.67, SE = .07, CrI (-.82, −.53)]. To assess the likelihood that participants confidence judgment responses changed between T2 and FU, we performed the same analysis but replacing the reference category to T2. This analysis showed that the likelihood that participants responded to "Remember" versus "Know" was lower in FU than in T2 [β = −.40, SE = .05, CrI (-.50, −.31)] while the likelihood that participants responded
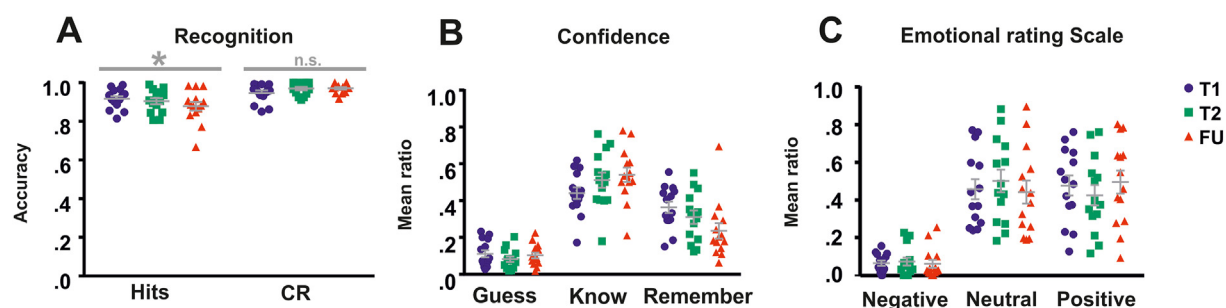
Fig. 2 − Behavioural data in healthy young participants from Experiment 1. (A) Averaged participants' accuracy and correct rejections in selecting their own events (old) compared to others' events (new) over three time periods (T1: one week, T2: two weeks, FU: 6−14 months). (B) Confidence judgments over three time periods (T1, T2, and FU). (C) Emotional ratings over three time periods (T1, T2, and FU). * indicates $p < .05$. Error bars represent SEM.

"Know" over "Guess" did not change [$\beta = .12$, SE $= .08$, CrI (-.09, .28)]. Altogether, the results indicate that participants' tendency to rate their confidence judgements as "Know" increased over time.

The indicators selected to assess model fit and performance showed that there were no problems in model convergence (all effective sample sizes > 1000; all Rhat = 1; all Pareto k estimates < .05; 0 of 4000 iterations ended with a divergence). We additionally compared this model with a null model (intercepts model) as well as with an alternative model without the inclusion of varying effects for participants ID. The results showed that the inclusion of time as an exogenous variable considerably improved the model fit from the null model (diff. Looic $= -212.2$; SEdiff $= 20.9$) and that the inclusion of the varying effect of participants ID resulted in a better model fit in comparison with the model without varying effects (diff. Looic $= -193.6$; SEdiff $= 20.2$). The inclusion of participants ID as a varying effect was also supported by the results on SD (intercept) which showed appreciable variability between participants [95% CrI (.31, .77)].

We next examined if participants' emotional ratings for Hits differed as a function of memory test (T1, T2 and FU). We observed that participants rarely indicated maximum negative (i.e., −2 in the scale) or positive (i.e., +2) ratings in the tests (negative: M $= 1.14\%$, SD $= 1.94\%$; positive: M $= 14.35\%$, SD $= 13.61\%$, averaged over the three test). Therefore, we grouped participants ratings −2 and −1 as negative and +1 and + 2 as positive, leaving 0 as indicating neutral emotion (Fig. 2C). The results showed that the likelihood that participants rated differently neutral and negative responses did not change in T2 compared to T1 [$\beta = -.13$, SE $= .12$, CrI (-.36, .10)] but that it decreased the likelihood they rated positive frequently than neutral their choices in T2 in comparison to T1 [$\beta = -.24$, SE $= .12$, CrI (-.37, -.12)]. When we analyzed the likelihood that participants' emotional ratings changed as a function of memory test, we found that the likelihood of judging the images as neutral versus negative [$\beta = .19$, SE $= .13$, CrI (-.06, .44)] increased while positive versus neutral [$\beta = -.06$, SE $= .06$, CrI (-.19, .06)] did not differ in T2 when compared to T1. When we ran the same analysis but replacing T2 instead of T1 as a reference, we found that participants' likelihood to indicate positive versus neutral [$\beta = -.05$, SE $= .05$, CrI (-.13,

.04)] and positive versus negative [$\beta = .05$, SE $= .09$, CrI (-.13, .23)] did not change in the FU test when compared to T2. In sum, these results indicate that emotional ratings upon the retrieval of AMs did not change between memory tests.

The indicators selected to assess model fit and performance showed that there were no problems in model convergence (all effective sample sizes > 1000; all Rhat = 1; all Pareto k estimates < .05; 0 of 4000 iterations ended with a divergence). We additionally compared this model with a null model (intercepts only) and an alternative model without the inclusion of varying effects for participants ID. The results showed that the inclusion of time as an exogenous variable considerably improved the model fit from the null model (diff. Looic $= -317.9$ SEdiff $= 24.7$) and that the inclusion of the varying effect of participants ID resulted in a better model fit in comparison with the model without varying effects (diff. Looic $= -332.7$ SEdiff $= 25.4$). The inclusion of participants' ID as a varying effect was also supported by the results on SD (intercept) which showed appreciable variability between participants [95% CrI (.42, 1.00)].

Finally, participants' response accuracy to temporal order memory was random (M $= .49$, SD $= .22$), thereby indicating this test was not suitable to capture the participants' ability to retrieve temporal representations from the tests.

3.1.2. ERPs results

Our analytical strategy was to first assess whether Old and New images elicited different patterns of brain activity in the participants. To address this issue, we averaged, at the participant level, the ERPs elicited by Old and New image conditions across tests and ran a cluster-based permutation test between these two conditions. This analysis revealed a significant cluster showing that Old images elicited higher ERP positive amplitude from 400 msec at stimulus onset, which lasted over the rest of the temporal window in the analysis and comprised frontal and central electrodes at the scalp (Fig. 3A). To assess how the identified Old/New ERP effect varied as a function of memory test, we selected the averaged ERP activity from the cluster to a repeated-measures ANOVA including experimental condition (Old, New) and memory test (T1, T2, and FU) as a within-subject factor. As expected, this analysis revealed a main effect of condition [$F(1,11) = 60.24$,
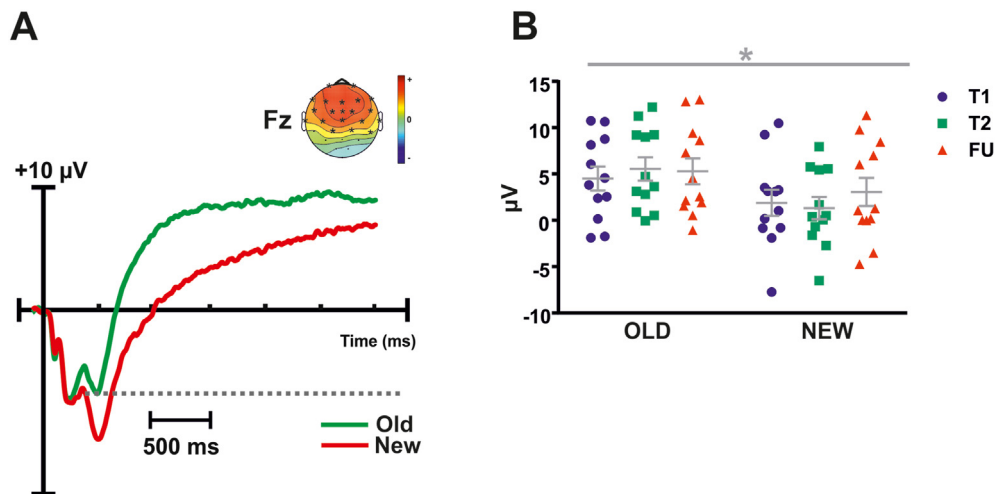
Fig. 3 − ERP old/new effect in Experiment 1. (A) Across participants grand-average Event-Related Potentials (ERPs) for Experiment 1 for Old and New conditions at Fz. A cluster-based permutation analysis between the two conditions revealed that Old pictures elicited greater ERPs amplitude than did fronto-central scalp regions. Dashed line indicates the temporal window of significance ($p < .05$, corrected). (B) Cluster-averaged individual ERP data for each of the three recognition memory tests (T1, T2, and FU) and conditions. * indicates $p < .05$ (corrected). Error bars represent SEM.

$p < .01$] and a significant condition x memory test interaction [$F(2,22) = 6.45, p < .01$] (Fig. 3B) but not a main effect of memory test [$F(2,22) = .79, p = .47$]. However, separate repeated-measures ANOVAs for ERP values to Old and New images revealed that none of them showed a statistically significant effect as a function of memory test [Old: $F(2,22) = .66, p = .53$; New: $F(2,22) = 2.58, p = .09$].

3.1.3. Time-frequency results
Following the ERP analytical strategy, we first implemented a cluster-based permutation test to identify, in a data-driven manner, the existence of a main Old/New difference pattern of neural oscillatory response along the temporal x spatial x spectral dimension. Thus, spectral power measures elicited at the onset of Old and New correct responses were averaged
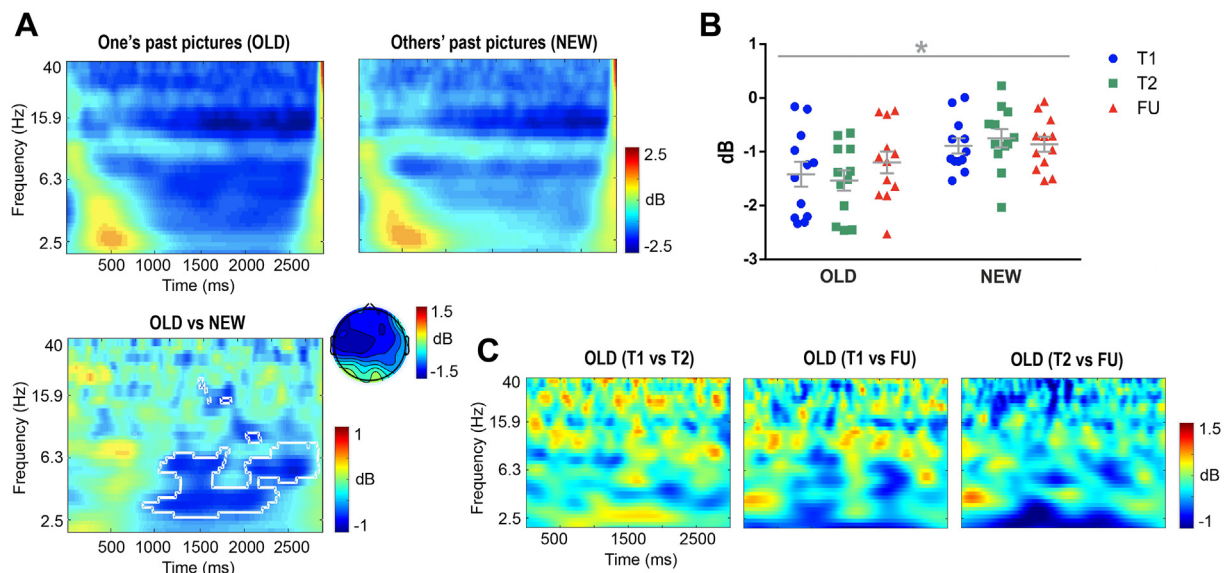
Fig. 4 − Theta oscillations in Experiment 1. (A) Group-averaged changes in spectral power over the three periods (T1: 1 week, T2: 2 weeks, and FU: 6−14 months after encoding) at a representative electrode (i.e., Fz) elicited by pictures related to a participant's own personal events (Old) compared to others' events (New). Difference between Old and New conditions is also displayed. The thick white line depicts the timing of the significant cluster between conditions ($p < .05$, cluster-based permutation test). Decreased theta power was observed in old responses over fronto-central regions of the scalp. (B) Average theta power over three periods of time (T1, T2, and FU) for each condition separately. Significant ($p < .05$) interaction effect is indicated with a grey star. Error bars represent SEM. (C) Group-averaged differences in spectral power changes elicited by Old stimuli between the three tests distributed over time at a representative electrode (i.e., Fz).

across the three tests (T1, T2, and FU) and were then compared. This analysis revealed the existence of a significant cluster initiating at around 1000 msec from stimulus onset, comprising low-frequency activity in the theta range (~3—8 Hz) (Fig. 4A). More specifically, these results showed that Old responses were accompanied by a decrease in theta power and that this effect was over frontal and central regions of the scalp. A repeated-measures ANOVA, including as within-subject factors memory test (T1, T2 and FU) and image type (Old and New), showed a main effect of image type [F(1,11) = 56.85, $p < .01$], no significant main effect of memory test [F(2,22) = .33, $p = .73$], but the existence of a significant image type x memory test interaction [F(2,22) = 5.15, $p = .01$] (Fig. 4B). These results suggest that theta Old/New effect differed between memory tests. However, paired t-test analysis showed that Old versus New theta differences were significant at T1 [t(11) = 4.21, $p < .01$], at T2 [t(11) = 7.21, $p < .01$], and at FU [t(11) = 3.69, $p < .01$], thereby hindering the possibility of establishing the source of this interaction clearly in our data.

To gain further insight into the decrease in theta changes over time, we compared spectral power response elicited by Old stimuli between the memory tests. This analysis revealed that Old stimuli elicited a significant decrease of theta power in T1 vs FU ($p = .03$, corrected) in a cluster that included frequencies in the lower theta range (i.e., 1—3.1 Hz) (Fig. 4C). Spectral power changes elicited by Old stimuli were not found to be statistically different when comparing T1 and T2 and T2 and FU ($p > .05$, corrected at cluster level).

3.2. Experiment 2

3.2.1. Behavioural results

Participants' hits and correct rejection (CR) rates are displayed in Fig. 5A. Overall, participants' performance was much poorer in this experiment. To assess this statistically, we compared participants' Hits and CRs in experiment 1 and in experiment 2. This analysis used participants' behavioural data obtained in condition T1 from experiment 1, as this shared a one-week time frame between encoding and retrieval in the two experiments. This analysis confirmed significant differences in the proportion of Hits [F(1,13) = 47.09 $p < .01$] and in CRs [F(1,13) = 31.12, $p < .01$] between experiments, thereby indicating, as expected, that participants were much more

accurate in recognizing pictures related to their own past real-life experience than those encoded "artificially" in a lab-context.

To assess for differences in confidence judgment ratings ("Remember", "Know", "Guess") for Hits in Experiment 2 we carried out a multinomial Bayesian regression without predictors (intercepts only). These intercepts reflect the difference in the probability of choice of each category (in a multinomial logit scale) with respect to the reference category and constitute a direct test that respects the dependence between categories, given that the sum of their respective probabilities is 1. The reference category was set to "Guess" for the contrasts "Guess" versus "Know" and "Guess" versus "Remember" and it was then replaced to "Know" for the remaining contrast ("Know versus Remember"). The results showed that the likelihood that participants indicated "Know" [$\beta = -.47$, SE = .18 CrI (-.84, −.09)] and "Remember" [$\beta = -1.30$, SE = .29, CrI (-1.89, −.73)] over "Guess" was lower (Fig. 5B). In addition, the likelihood that participants indicated "Remember" [$\beta = -.85$, SE = .28, CrI (-1.39, −.31)] over "Know" was also lower. These results indicate that participants' sense of recollection for laboratory-based material was not as vivid as when they judged their own stimuli.

3.2.2. ERPs results

EEG data were analysed as in experiment 1, by contrasting ERP patterns of activity elicited by Old and New images through a data-driven cluster-based permutation test. This analysis revealed a significant cluster showing that Old images elicited higher ERP positive amplitude at 400 msec from stimulus onset. However, compared to the ERP differences observed in experiment 1, this cluster of EEG activity was much shorter and less well distributed over the scalp (Fig. 6A). Nevertheless, a repeated-measures ANOVA including response type (Old and New) and experiment (T1 from experiment 1 and experiment 2) as within-subject factor revealed that the Old/New ERP effects were statistically similar in the two experiments [main effect of experiment: F(1,9) = 1.96, $p = .19$; experiment x response type interaction: F(1,9) = 3.29, $p = .10$].

3.2.3. Time-frequency results

The implementation of a cluster-based permutation test assessing differences between Old and New responses to
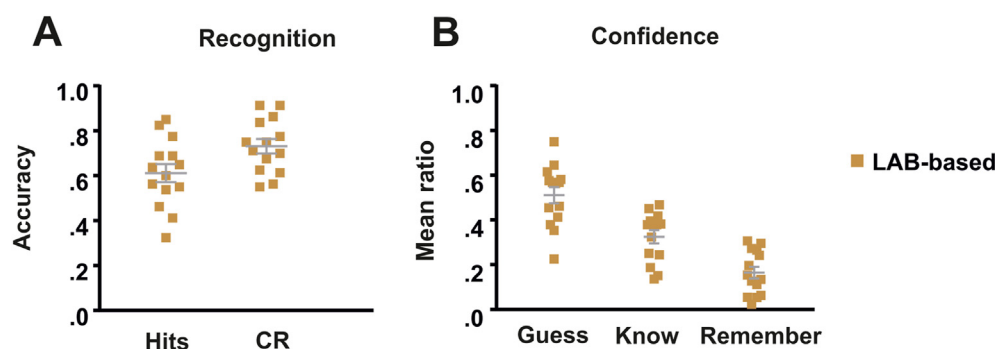
Fig. 5 — Behavioural data in healthy young participants from Experiment 2. (A) Averaged participants' accuracy (Hits) and correct rejections (CR) in selecting pictures previously encoded (old) compared to pictures never seen (new). (B) Confidence judgments based on laboratory stimuli. * indicates $p < .05$. Error bars represent SEM.
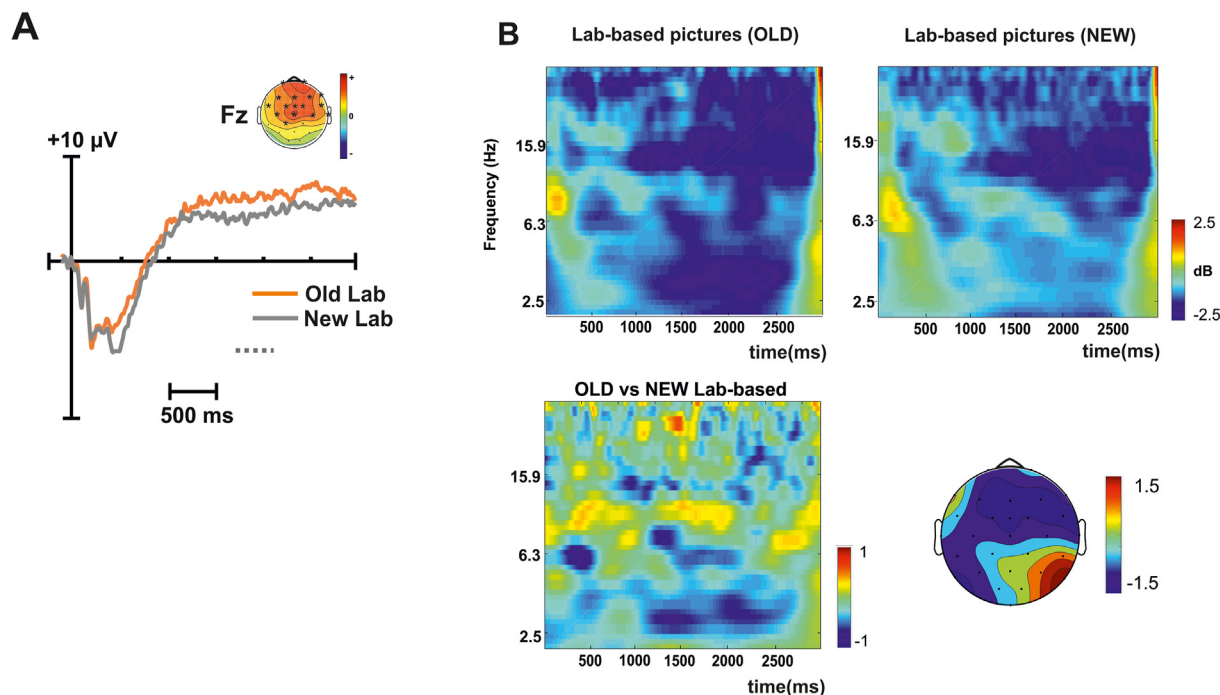
Fig. 6 – ERP and time-frequency data in Experiment 2. (A) Across-participants grand-average event-related potentials (ERPs) for Experiment 2 for Old and New conditions at Fz. A cluster-based permutation analysis between the two conditions revealed that Old pictures elicited greater ERP amplitudes than did fronto-central scalp regions. Dashed line indicates the temporal window of significance (*p* < .05, corrected). (B) Group-averaged changes in spectral power (averaged over all scalp sensors) elicited by pictures related to own personal events (old) compared to others' events (new).

scene images encoded in the laboratory yielded non-significant results along the temporal x frequency x spatial (scalp sensors) dimensions Fig. 6B. Thus, contrary to the theta Old/New effects found in experiment 1, differences were not observable in experiment 2. Nevertheless, to assess the extent to which the Old/New theta effects found in experiment 1 differed from the same theta modulations in experiment 2, we ran a repeated-measures ANOVA including response type (Old and New) and experiment (T1 from experiment 1 and experiment 2) as within-subjects factors. Given the null effects found in experiment 2, theta power in this experiment was extracted by averaging data points that were within the temporal x frequency x spatial cluster data points identified in experiment 1. This analysis confirmed the main effect of image type [$F(1,9) = 9.57$, $p = .01$] but uncovered no significant main effect of the memory test [$F(1,9) = 1.96$, $p = .19$] nor type x memory test interaction [$F(1,9) = 3.29$, $p = .10$], which hindered the possibility of concluding that the Old/New ERP effect was greater in one of the experiments. However, note that the clusters of activity included in this analysis were much larger in ERPs from experiment 1 than in ERPs from experiment 2.

4.　Discussion

Real-life experience is characterized by continuous inputs carved by memory systems into an organized, yet intricate, memory network of episodic event representations. How the memory systems retrieve past episodic experience from daily life routine and how the retrieval of these episodes changes with the passage of time are a current hallmark in the Auto-biographical Memory (AM) research field that has important implications at the clinical level (Urbanowitsch et al., 2013). In the current study, we examined an experimental approach that sought to cue episodic retrieval by using pictures depicting instances of individual past real-life experience. The use of neural network computational methods to chunk picture sequences into segments of meaningful event episodes within the daily life routine activity help reduce the number of possible cues to be used at test without compromising the possibility of exploring memory retrieval for the entire past experience. We found that at test, participants were much more accurate in recognizing pictures depicting scenes from their real life than picture scenes encoded in the lab, but that memory recollection of personally experienced events elicited by real life pictures rapidly decreased over time. We also observed that retrieval for real-life events elicited a high proportion of familiarity (i.e., "know" responses), even one week after encoding, and that this tendency increased when memory was tested 2 weeks (T2) and more prominently after 6–14 months (FU) after encoding. Our findings revealed that the retrieval of real-life picture cues elicited a strong and positive 'ERP old/new effect' over frontal regions and that the magnitude of this ERP effect was similar throughout memory tests over time. However, we found recognition memory induced a frontal theta power decrease and that this effect was mostly seen when memories were tested after 1 and 2

weeks but not after 6−14 months from encoding. These findings provide evidence of distinct neurophysiological signatures supporting the retrieval of real-life routine episodes and their dynamics with retrieval over the passage of time. At the same time, they show that memories for routine events, i.e., memories for episodes that lacked the memory benefit of novelty, salience or context specificity do also decline over the course of time.

In experiment 1, we examined the ability of healthy adults to retrieve AMs of real-life event episodes encoded over a period of a week. To gain insight into how memories may be retrieved from experiences that best accounted for daily life activity, we explicitly instructed participants that the encoding period should take place during a week in which no salient or rare events were expected to take place (e.g., birthday parties, important job meetings, exceptional visits from old friends). We reasoned that this approach would provide information about the retrieval for experienced episodes that dominated most of our everyday life routines, which are characterized by having a high degree of overlap (e.g., having dinner at home, working at the same office) and unfold continuously over time. Thus, while previous investigation showed that the ability to recollect detailed and vivid memories from past experiences is modulated by their saliency and emotional content during encoding (Holland & Kensinger, 2010), it is less clear how AMs are retrieved for real-life experiences that are highly repetitive and less exceptional, though they are those that better represent an individual's routine. The results from the current study revealed that participants were accurate in recognizing picture cues that depicted scenes from these past routine experiences but that they were not strong enough to induce high degrees of vivid and detailed recollection of specific episodes, even when memories were tested after one week from encoding. Thus, while participants' memory accuracy rates were much higher for picture scenes depicting real-life experience than for picture scenes encoded at the lab after one week from encoding, the associated high subjective ratings of familiarity by participants suggests that our ability to recollect details of our routine past episodes declines quite rapidly over time. Nevertheless, the increased ratings of familiarity upon picture cues when memories were tested 2 weeks and most prominently after 6−14 months after encoding suggests that our design was suitable to capture a progressive memory decline associated with the passage of time. This increase in the feeling of familiarity might be explained by the expected decline in memory strength over time that could cause the loss of contextual features, attributed to an inherent time-dependent memory transformation due to consolidation processes (Squire & Alvarez, 1995; Dudai & Morris, 2000). In fact, it has been suggested that memories for individual event episodes that take place in the same context may be more susceptible to being integrated into a semantic memory network. Over time, having multiple related memory traces may facilitate the extraction of factual information into a semantic representation of the gist of the episode (Winocur et al., 2010; Winocur & Moscovitch, 2011). This information would be integrated into a semantic knowledge and become independent of the specific episode. In the context of the current experiment,

pictures depicting contextual features derived from repeated routine episodes (e.g., a picture of the sofa from the living room), may elicit the retrieval of semantic representations linked to multiple individual episodes, thereby facilitating process of familiarity rather than recollection at retrieval. Additionally, over time, specific episodic memories related to routine activity have the potential to be interfered with by multiple similar experiences that took place in same contexts. The repeated sequential encoding of events that shared similar contextual information could interfere with previously encoded memories, rendering them less prone to be recollected (Forcato et al, 2007, 2009; Hupbach et al., 2007). While the current experimental findings cannot disambiguate between these explanations, we also believe the use of similar experimental strategies may also offer the possibility of addressing them in future research.

The strong positive-going frontal "ERP old/new effect" elicited by picture cues that were correctly recognized by the participants indicate the retrieval of AMs from picture cues engaged similar neural responses widely described previously in lab-based memory research. Specifically, we found that this frontal effect started 400 msec at picture onset and remained sustained throughout the entire time window of analysis (i.e., 2500 msec). The fact that the ERP effects were not found at parietal scalp regions during retrieval lends support to the notion that retrieval was mostly accounted for by neural mechanisms supporting the retrieval of information with a lack of contextual details (i.e., familiarity) (Curran, 2000; Curran et al., 2001; Duzel et al., 1997; Rugg et al., 1998; Wilding & Rugg, 1996). A surprising finding of the current study is that such frontal ERP old/new effect was not modulated in magnitude throughout the different memory tests, an issue that does not fit with the behavioural data showing a decreasing tendency of the participants to retrieve contextual episodic information by a picture cue. A plausible explanation is that the sustained frontal ERP old/new effects was reflecting the need to engage post-retrieval monitoring (Allan & Rugg, 1997; Curran et al., 2001; Donaldson & Rugg, 1999; Ranganath & Paller, 2000; Wilding & Rugg, 1997a,b; Wolk et al., 2006), that may not necessarily be associated with retrieval success (Ranganath & Paller, 1999). It has been suggested that monitoring is engaged whenever the outcome of a retrieval attempt is ambiguous (Rugg and Helnson, 2002). This could occur, for example in our study, because the retrieved information is poorer, causing uncertainty about the item's study status (as in the case of an unrecollected item eliciting a weak familiarity signal). An alternative possibility would be that the positive-going frontal "ERP old/new effect" reflected neural and cognitive processes associated with "scene reconstruction". Some researchers have postulated that autobiographical memories, particularly remote ones, are essentially a reconstruction of original episodes out of schematic knowledge and inference (Barry & Maguire, 2019). This view argues that, even in the absence of an original trace, the hippocampus accomplishes this by assembling consolidated neocortical elements into spatially coherent scenes that form the basis of unfolding memory events. This reconstruction is likely facilitated by input from the medial prefrontal cortex. Thus, our findings that the positive-going frontal "ERP old/new effect" did not decrease in retrieval over the passage of time could

also reflect the medial prefrontal cortex engagement in mental reconstruction of memories that are faded, impoverished or schematic in nature.

Perhaps less expectedly, we found that the cue picture recognition elicited a decrease in theta power over frontal scalp regions. Increase of spectral power at the theta band has been well documented during retrieval success in humans (e.g., Burgess & Gruzelier, 1997; Guderian & Düzel, 2005; Gruber et al., 2008; Klimesch et al., 2001; Osipova et al., 2006). On the other hand, theta oscillations have been shown to be negatively related to retrieval success in interference and inhibition studies (Hanslmayr et al., 2010; Khader & Rösler, 2010; Staudigl et al., 2010). Interestingly, a recent investigation provided evidence of a way to reconcile these views by showing the distinct slow (~3 Hz) and fast (~7 Hz) scalp EEG theta power dynamics coexisted during cue word recognition from a studied word list (Pastötter & Bäuml, 2014). They showed that the positive effect in slow theta power was predominant in the first half of the cue presentation interval (i.e., .5–1 sec after word onset) and was localized over fronto-central and right-parietal electrodes. In contrast, the negative effect in fast theta power was predominant in the second half of the cue presentation (i.e., 1–2 sec after word onset) interval and was localized over mid-frontal and left-temporal electrode sites. These results suggested that theta oscillations recorded from scalp EEG may have distinct functional roles in episodic memory retrieval, with slow theta oscillations being related to processes of recollection and conscious awareness, and fast theta oscillations being linked to processes of interference and interference resolution. In our study, EEG theta power decrease was mostly observed at 1–2 sec from picture cue onset and distributed over left fronto-temporal regions at the scalp. In line with the aforementioned study, we reasoned that such theta decrease may reflect processes of interference resolution derived from participants' retrieval effort to relate a picture cues to specific memory-encoded episodes in the past. This interpretation lends support to the notion that the use of picture scenes depicting context information shared by multiple episodic events experienced during a routine week of activity (e.g., living room, desk at the office) as memory cues may also engender the need to deal with the concomitant retrieval of target (linked to a specific episode) and non-target information (linked to previous episodes that share similar context) during retrieval. Indeed, individual differences in the ability to retrieve personal episodic memories in response to cue words (e.g., "home") have been associated with inhibitory control and cognitive flexibility (Guler & Mackovichova, 2019). The decrease of theta power effects upon picture recognition in the FU test in our study may be explained by the lack of need to deal with concomitant target and non-target competitive retrieval of specific episodic memories, as at that time, after 6–14 months from encoding, most of the episodic details associated to the encoded events may have been forgotten.

Wearable cameras have been utilized by memory researchers to capture photographs of real-world experiences that can later be presented in the fMRI scanner to probe various aspects of AM. These studies have helped confirm that core cortical (i.e., anterior and posterior midline, and lateral frontal and posterior cortices) and medial temporal lobe system supporting episodic memory retrieval in laboratory-based studies (e.g., Kim, 2012; St.-Jacques, Kragel, & Rubin, 2011; for reviews see; Andrews-Hanna, et al., 2015; Robin & Moscovitch, 2017) can generalize to the retrieval of real-world memories encoded in naturalistic settings (Chow and Rissman, 2017). At the same time, recent fMRI studies using wearable camera technology provided valuable information that could not be easily ascertained through other methods. For example, the use of this approach combined with multivariate pattern analysis revealed that the human hippocampus represents the spatial and temporal location of memories for real-life events at scales of up to 30 km and a month of time (Nielson et al., 2015). Scalp EEG data is blind to the neural sources and therefore our results are not suitable to provide information about the contribution of AM-related neural structures during the retrieval of everyday real-life episodes. However, the fine-grained temporal resolution of the EEG complements fMRI findings by allowing us to investigate the neural dynamics sustained during very rapid cognitive operations, which is critical to test the prediction that neural mechanisms should appear intertwined and occur very rapidly during retrieval. In addition, different neural mechanisms can be simultaneously measured with EEG activity, thereby enabling the possibility of 2 neural signals being measured independently of each other. Concretely, we investigated the possibility that retrieval of memories for real-life event events were supported by evoked neural responses, registered as event-related potentials (ERPs) and changes in ongoing neural oscillatory activity, specifically in the theta range (3–9 Hz). Indeed, the combined study of ERPs and neural oscillations has provided critical insights into the timing and the neural dynamics operating during encoding, maintenance and retrieval of memory episodes (Duzel et al., 1997; Rugg et al., 1998; Paller & Wagner, 2002; Düzel, Penny, & Burgess, 2010, and they have been shown to be sensitive to medial temporal lobe lesions (including the hippocampus) (Düzel, Vargha-Khadem, Heinze, & Mishkin, 2001). Furthermore, the opportunity to investigate these two neural mechanisms with scalp EEG recordings represents a valuable contribution to clinical environments. Thus, the simple placing of a few EEG sensors, even at the ambulatory level, would allow a fine-grained exploration in neurological patients of basic neural mechanisms that are very relevant in the retrieval of memories for their individual past.

In conclusion, the results of the current study provide evidence of a methodological approach to exploring the retrieval of individual AMs for everyday real-life activity. Our findings help contribute to the emerging body of literature emphasizing the advantages of using naturalistic material rather than artificial lab-based stimuli to explore the cognitive and neural underpinnings of episodic memory retrieval. Current findings may be relevant beyond basic research as they may help improve understanding of AM functioning in the clinic at the individual level.

Transparency and data availability

The conditions of our ethics approval do not permit public archiving of anonymized study data. Readers seeking access

to the data should contact the lead author (BN). Access will be granted to named individuals in accordance with ethical procedures governing the reuse of sensitive data. Specifically, requestors must meet complete a formal data sharing agreement. The Matlab® and R set of scripts of our experiment are available at https://osf.io/j6gmx/

Credit author statement

Acknowledgements

REFERENCES

Allan, K., & Rugg, M. D. (1997). An event-related potential study of explicit memory on tests of cued recall and recognition. *Neuropsychologia, 35*, 387–397.

Allé, M. C., Manning, L., Potheegadoo, J., Coutelle, R., Danion, J. M., & Berna, F. (2017). Wearable cameras are useful tools to investigate and remediate autobiographical memory impairment: A systematic PRISMA review. *Neuropsychological Review, 27*, 81–99.

Andrews-Hanna, J. R., Saxe, R., & Yarkoni, T. (2015). Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, restingstate connectivity, and fMRI meta-analyses. *Neuroimage, 91*, 324–335.

Barry, D. N., & Maguire, E. A. (2019). Remote memory and the Hippocampus: A constructive critique. *Trends in Cognitive Sciences, 2019*(23), 128–142.

Berry, E., Hampshire, A., Watson, Kapur N., Owen, A., Browne, G., Rowe, A., Hodges, S., Wood, K., & Smyth, G. (2009). The neural basis of effective memory therapy in a patient with limbic

encephalitis. *Journal of Neurology, Neurosurgery and Psychiatry, 80*, 1202–1205.

Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B., & Wood, K. (2007). The use of SenseCam as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis - a preliminary report. *Neuropsychological Rehabilitation, 17*, 582–601.

Bonnici, H. M., Chadwick, M. J., Lutti, A., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2012). Detecting representations of recent and remote autobiographical memories in vmPFC and hippocampus. *Journal of Neuroscience, 32*(47), 16982–16991.

Bunzeck, N., & Düzel, E. (2006). Absolute coding of stimulus novelty in the human substantia Nigra/VTA. *Neuron, 51*(3), 369–379.

Burgess, A. P., & Gruzelier, J. H. (1997). Short duration synchronization of human theta rhythm during recognition memory. *Neuroreport, 8*, 1039–1042.

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software, 80*(1), 1–28.

Bürkner, P. C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science, 2*(1), 77–101.

Cabeza, R., Prince, S. E., Daselaar, S. M., Greenberg, D. L., Budde, M., Dolcos, F., LaBar, K. S., & Rubin, D. C. (2004). Brain activity during episodic retrieval of autobiographical and laboratory events: An fMRI study using a novel photo paradigm. *Journal of Cognitive Neuroscience, 16*(9), 1583–1594.

Cabeza, R., & St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends in Cognitive Sciences, 11*, 219–227, 2007.

Chow, T. E., & Rissman, J. (2017). Neurocognitive mechanisms of real-world autobiographical memory retrieval: Insights from studies using wearable camera technology. *Annals of the New York Academy of Sciences, 1396*, 202–221.

Chow, T. E., Westphal, A. J., & Rissman, J. (2018). Multi-voxel pattern classification differentiates personally experienced event memories from secondhand event knowledge. *Neuroimage, 176*, 110–123.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28*, 923–938.

Curran, T., Schacter, D. L., Johnson, M. K., & Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience, 13*, 201–216.

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*(1), 9–21.

Diamond, N., & Levine, B. (2020). Linking details to temporal context reinstatement in recall of real-world experiences. *Psychological Science, 31*(115), 1557–1572.

Dimiccoli, M., Bolaños, M., Talavera, E., Aghaei, M., Nikolov, S. G., & Radeva, P. (2015). SR-Clustering: Semantic regularized clustering for egocentric photo streams segmentation. *Computer Vision and Image Understanding, 155*, 55–69.

Donaldson, D. I., & Rugg, M. D. (1999). Event-related potential studies of associative recognition and recall: Electrophysiological evidence for context dependent retrieval processes. *Cognitive Brain Research, 8*, 1–16.

Dudai, Y., & Morris, R. (2000). To consolidate or not to consolidate: what are the questions? https://doi.org/10.1093/acprof:oso/9780198524823.003.0011

Dudai, Y., & Morris, R. G. M. (2013). Memorable trends. *Neuron, 80*, 742–750.

Duzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H., & Tulving, E. (1997). Event-related potential correlates of two states of conscious awareness in memory. *Proceedings of the National*

Academy of Sciences of the United States of America, 94,
5973–5978.

Düzel, E., Vargha-Khadem, Heinze, & Mishkin. (2001 Jul 3). Brain
activity evidence for recognition without recollection after
early hippocampal damage. *Proceedings of the National Academy
of Sciences of the United States of America, 98*(14), 8101–8106.

Düzel, E., Penny, wd, & Burgess, N. (2010 Apr). Brain oscillations
and memory. *Current Opinion in Neurobiology, 20*(2), 143–149.
https://doi.org/10.1016/j.conb.2010.01.004

Forcato, C., Argibay, P. F., Pedreira, M. E., & Maldonado, H. (2009).
Human reconsolidation does not always occur when a
memory is retrieved: The relevance of the reminder structure.
Neurobiology of Learning and Memory, 91, 50–57.

Forcato, C., Burgos, V. L., Argibay, P. F., Molina, V. A., Pedreira, M. E.,
& Maldonado, H. (2007). Reconsolidation of declarative memory
in humans. *Learning & Memory, 14,* 295–303.

Fuentemilla, L., Barnes, G. R., Düzel, E., & Levine, B. (2014). Theta
oscillations orchestrate medial temporal lobe and neocortex
in remembering autobiographical memories. *Neuroimage, 85,*
730–737.

Fuentemilla, Lluís, Palombo, D. J., & Levine, B. (2018). Gamma
phase-synchrony in autobiographical memory: Evidence from
magnetoencephalography and severely deficient
autobiographical memory. https://doi.org/10.1016/j.neuropsy
chologia.2017.08.020

Fuentemilla, L., Penny, W. D., Cashdollar, N., Bunzeck, N., &
Düzel, E. (2010). Theta-coupled periodic replay in working
memory. *Current Biology, 20*(7), 606–612.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and
multilevel/hierarchical models.* Cambridge university press.

Gruber, T., Tsivilis, D., Giabbiconi, C. M., & Müller, M. M. (2008).
Induced electroencephalogram oscillations during source
memory: Familiarity is reflected in the gamma band, recollection
in the theta band. *Journal of Cognitive Neuroscience, 20,* 1043–1053.

Guderian, S., & Düzel, E. (2005). Induced theta oscillations
mediate large-scale synchrony with mediotemporal areas
during recollection in humans. *Hippocampus, 15,* 901–912.

Guler, O. E., & Mackovichova, S. (2019). The role of executive
function in autobiographical memory retrieval: Does the type
of cue word matter? *Memory, 27*(4), 423–430.

Hanslmayr, S., Staudigl, T., Aslan, A., & Bäuml, K.-H. T. (2010). Theta
oscillations predict the detrimental effects of memory retrieval.
Cognitive, Affective & Behavioral Neuroscience, 10, 329–338.

Herweg, N. A., Apitz, T., Leicht, G., Mulert, C., Fuentemilla, L., &
Bunzeck, N. (2016). Theta-alpha oscillations bind the
hippocampus, prefrontal cortex, and striatum during
recollection: Evidence from simultaneous EEG-fMRI. *The
Journal of Neuroscience: the Official Journal of the Society for
Neuroscience, 36,* 3579–3587.

Herweg, N. A., Solomon, E. A., Kahana, M. J., & Kahana, M. J.
(2020). Theta oscillations in human memory. *Trends in
Cognitive Sciences, 24*(3), 208–227.

Holland, A. C., & Kensinger, E. A. (2010). Emotion and
autobiographical memory. *Physics of Life Reviews, 7*(1), 88–131.

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation
of episodic memories: A subtle reminder triggers integration of
new information. *Learning & Memory, 14,* 47–53.

Jeunehomme, O., & D'Argembeau, A. (2018). Event segmentation
and the temporal compression of experience in episodic
memory. *Psychological Research, 84*(2), 1–10.

Jeunehomme, O., & D'Argembeau, A. (2019). The time to
remember: Temporal compression and duration judgements
in memory for real-life events. *Quarterly Journal of Experimental
Psychology, 72*(4), 930–942.

Khader, P. H., & Rösler, F. (2010). EEG power changes reflect
distinct mechanisms during long-term memory retrieval.
Psychophysiology, 48, 362–369.

Kim, H. (2012). A dual-subsystem model of the brain's default
network: Self-referential processing, memory retrieval
processes, and AM retrieval. *Neuroimage, 61,* 966–977.

Klimesch, W., Doppelmayr, M., Yonelinas, A., Kroll, N. E. A.,
Lazzara, M., Röhm, D., & Gruber, W. (2001). Theta
synchronization during episodic retrieval: Neural correlates of
conscious awareness. *Cognitive Brain Research, 12,* 33–38.

LePort, A. K. R., Mattfeld, A. T., Dickinson-Anson, H., Fallon, J. H.,
Stark, C. E. L., Kruggel, F., & McGaugh, J. L. (2012). Behavioral
and neuroanatomical investigation of Highly Superior
Autobiographical Memory (HSAM). *Neurobiology of Learning and
Memory, 98*(1), 78–92.

LePort, A. K. R., Stark, S. M., McGaugh, J. L., & Stark, C. E. L. (2017).
A cognitive assessment of highly superior autobiographical
memory. *Memory, 25*(2), 276–288.

Levine, B., Turner, G. R., Tisserand, D., Hevenor, S. J., Graham, S. J.,
& McIntosh, A. R. (2004). The functional neuroanatomy of
episodic and semantic autobiographical remembering: A
prospective functional MRI study. *Journal of Cognitive
Neuroscience, 16*(9), 1633–1646.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical
testing of EEG- and MEG-data. *Journal of Neuroscience Methods,
164*(1), 177–190.

Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S., &
Sederberg, P. B. (2015). Human hippocampus represents space
and time during retrieval of real-world memories. *Proceedings
of the National Academy of Sciences of the United States of America,
112*(35), 11078–11083.

Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta
oscillations in episodic memory. *Neuroscience and Biobehavioral
Reviews, 34,* 1023–1035.

Osipova, D., Takashima, A., Oostenveld, R., Fernández, G.,
Maris, E., & Jensen, O. (2006). Theta and gamma oscillations
predict encoding and retrieval of declarative memory. *The
Journal of Neuroscience: the Official Journal of the Society for
Neuroscience, 26,* 7523–7531.

Palombo, D. J., Alain, C., Söderlund, H., Khuu, W., & Levine, B.
(2015). Severely deficient autobiographical memory (SDAM) in
healthy adults: A new mnemonic syndrome. *Neuropsychologia,
72,* 105–118.

Paller, K. A., & Wagner, A. D. (2002). Observing the transformation
of experience into memory. *Trends in Cognitive Sciences, 6*(2),
93–102. https://doi.org/10.1016/S1364-6613(00)01845-3

Pastötter, B., & Bäuml, K.-H. T. (2014). Distinct slow and fast
cortical theta dynamics in episodic memory retrieval.
Neuroimage, 94, 155–161.

Patihis, L., Frenda, S. J., LePort, A. K. R., Petersen, N.,
Nichols, R. M., Stark, C. E. L., & Loftus, E. F. (2013). False
memories in highly superior autobiographical memory
individuals. *Proceedings of the National Academy of Sciences of the
United States of America, 110*(52), 20947–20952.

Ranganath, C., & Paller, K. A. (1999). Frontal brain potentials
during recognition are modulated by requirements to retrieve
perceptual detail. *Neuron, 22*(3), 605–613.

Ranganath, C., & Paller, K. A. (2000). Neural correlates of memory
retrieval and evaluation. *Cognitive Brain Research, 9,* 209–222.

Rissman, J., Chow, T. E., Reggente, N., & Wagner, A. D. (2016).
Decoding fmri signatures of real-world autobiographical
memory retrieval. *Journal of Cognitive Neuroscience, 28*(4),
604–620.

Robin, J., & Moscovitch, M. (2017). Details, gist and schema:
Hippocampal -cortical interactions underlying recent and
remote episodic and spatial memory. *Current Opinion in
Behavioral Sciences, 17,* 114–123.

Rugg, M. D., & Henson, R. N. (2002). Episodic memory retrieval: An
(event-related) functional neuroimaging perspective. In
A. E. Parker, E. L. Wilding, & T. Bussey (Eds.), *The cognitive*

neuroscience of memory encoding and retrieval. Hove, UK: Psychology Press.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature, 392*, 595−598.

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinion in Neurobiology, 5*, 169−177.

Staudigl, T., Hanslmayr, S., & Bäuml, K.-H. T. (2010). Theta oscillations reflect the dynamics of interference in episodic memory retrieval. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience, 30*, 11356−11362.

St Jacques, P. L., Kragel, P. A., & Rubin, D. C. (2011). Dynamic neural networks supporting memory retrieval. *Neuroimage, 57*(2), 608−616.

Svoboda, E., & Levine, B. (2009). The effects of rehearsal on the functional neuroanatomy of episodic autobiographical and semantic remembering: A functional magnetic resonance imaging study. *Journal of Neuroscience, 29*(10).

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*(1), 1−25.

Urbanowitsch, N., Gorenc, L., Herold, C. J., & Schröder, J. (2013). Autobiographical memory: A clinical perspective. *Frontiers in Behavioral Neuroscience, 7*, 194.

Wilding, E. L., & Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain: a Journal of Neurology, 119*, 889−905.

Wilding, E. L., & Rugg, M. D. (1997a). An event-related potential study of memory for words spoken aloud or heard. *Neuropsychologia, 35*, 1185−1195.

Wilding, E. L., & Rugg, M. D. (1997b). Event-related potentials and the recognition memory exclusion task. *Neuropsychologia, 35*, 119−128.

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society, 17*, 766−780.

Winocur, G., Moscovitch, M., & Bontempi, B. (2010). Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia, 48*, 2339−2356.

Wolk, D. A., Schacter, D. L., Lygizosc, M., Senc, M., Holcombd, P. J., Daffnerc, K. R., & Budsonc, A. E. (2006). ERP correlates of recognition memory: Effects of retention interval and false alarms. *Brain Research, 1096*, 148−162.

Woodberry, E., Browne, G., Hodges, S., Watson, P., Kapur, N., & Woodberry, K. (2014). The use of a wearable camera improves autobiographical memory in patients with Alzheimer's Disease. *Memory, 23*, 340−349.

Xu, Q., Zhang, J., Grandjean, J., Tan, C., Subbaraju, V., Li, L., Lee, K. J., Hsieh, P. J., & Lim, J. H. (2020). Neural correlates of retrieval-based enhancement of autobiographical memory in older adults. *Scientific Reports, 10*, 1447.

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science, 16*(2), 80−84.