Episodic retrieval involves early and sustained effects of reactivating information from encoding

Jeffrey D. Johnson *, Mason H. Price, Emily K. Leiker

Department of Psychological Sciences, University of Missouri, USA

ABSTRACT

Several fMRI studies have shown a correspondence between the brain regions activated during encoding and retrieval, consistent with the view that memory retrieval involves hippocampally-mediated reinstatement of cortical activity. With the limited temporal resolution of fMRI, the precise timing of such reactivation is unclear, calling into question the functional significance of these effects. Whereas reactivation influencing retrieval should emerge with neural correlates of retrieval success, that signifying post-retrieval monitoring would trail retrieval. The present study employed EEG to provide a temporal landmark of retrieval success from which we could investigate the sub-trial time course of reactivation. Pattern-classification analyses revealed that early-onsetting reactivation differentiated the outcome of recognition-memory judgments and was associated with individual differences in behavioral accuracy, while reactivation was also evident in a sustained form later in the trial. The EEG findings suggest that, whereas prior fMRI findings could be interpreted as reflecting the contribution of reinstatement to retrieval success, they could also indicate the maintenance of episodic information in service of post-retrieval evaluation.

© 2014 Elsevier Inc. All rights reserved.

Introduction

Successful memory retrieval is thought to depend on the degree to which processes engaged during encoding of the memory are re-engaged upon presentation of a retrieval cue (Bower, 1972; Damasio, 1989). This principle is reflected by a common experience in daily life, when difficulty recalling a memory is sometimes resolved by retracing the activities leading up to the initial experience. In experimental psychology, studies providing initial evidence for this idea demonstrated improvements in memory performance as a result of having subjects re-engage encoding-related cognitive processes during a memory test (Tulving and Thomson, 1973; Morris et al., 1977). Further support for the idea comes from highly-influential neurobiological models that have guided episodic memory research for the past 20 years (Alvarez and Squire, 1994; McClelland et al., 1995; Rolls, 2000). A fundamental feature of these models, with respect to the conscious retrieval ("recollection") of episodic information, is the involvement of the hippocampus. During encoding, cortical activity patterns elicited by an event are represented rapidly and sparsely by the hippocampus (Marr, 1971; Teyler and DiScenna, 1986; Shastri, 2002). Upon retrieval cue presentation, the hippocampal representation is reactivated, thereby reinstating the cortical pattern that was present during encoding and allowing for the recollection of additional episodic information

(i.e. not inherent in a partial cue; Hasselmo and Wyble, 1997; Norman and O'Reilly, 2003).

Functional magnetic resonance imaging (fMRI) studies have been instrumental in providing empirical evidence of neural reinstatement during episodic retrieval (for reviews, see Rugg et al., 2008; Danker and Anderson, 2010). These studies test the simple prediction that the pattern of brain activity at encoding should match that at retrieval. In an early study, Wheeler et al. (2000) had subjects encode words paired with either a picture or an auditory stimulus. An important feature of their design, carried forth to subsequent studies and the present one, was that the cues presented during a later memory test were perceptually uninformative about the previous encoding condition in which they appeared. Thus, any brain activity specific to a previous encoding condition was inferred to have resulted from retrieving the associated picture or auditory information. Wheeler et al. demonstrated that cues from each condition during retrieval reactivated a subset of the regions active at encoding (also see Kahn et al., 2004; Johnson and Rugg, 2007). More recently, multivariate pattern analysis (MVPA; see Haynes and Rees, 2006; Norman et al., 2006; Tong and Pratte, 2012) of fMRI data has been used to investigate reinstatement by training a pattern classifier to distinguish brain activity associated with different conditions at encoding and then independently evaluating that classifier on data from a memory test (for review, see Rissman and Wagner, 2012). The classifier's ability to identify an item's encoding history depends on the encoding-retrieval similarity of neural patterns, thus providing an index of reinstatement. In one study, Johnson et al. (2009) used MVPA to assess reinstatement across two subjectively-

* Corresponding author at: Department of Psychological Sciences, 210 McAlester Hall, University of Missouri, Columbia, MO 65211, USA.
   E-mail address: johnsonjeffre@missouri.edu (J.D. Johnson).

different retrieval experiences — an acontextual sense of familiarity and the recollection of specific details from encoding (Yonelinas, 2002). The magnitude of pattern reactivation increased in a graded manner across these experiences, suggesting that they are supported by changes in a common neural process. Together with other results based on this analysis strategy (e.g., Polyn et al., 2005; McDuff et al., 2009; Kuhl et al., 2011, 2012; Staresina et al., 2012), these findings demonstrate the utility of tracking reactivation across different memory judgments to further understand the role(s) of reinstatement in retrieval.

Whereas the fMRI studies described above have convincingly demonstrated the involvement of encoding-related reactivation during retrieval, the fact that they rely on the relatively slow hemodynamic response is a limiting factor in assessing the timing of reinstatement effects. One notable exception is the study by Polyn et al. (2005), which temporally isolated reactivation occurring before a retrieved item was reported; however, given that those effects likely reflect preparatory (or categorical) rather than item-specific processing, the sub-trial timing of neural events related to retrieval success are yet to be determined. In keeping with the models described earlier, hippocampally-mediated reinstatement is thought to allow episodic information to become available for conscious retrieval (Norman and O'Reilly, 2003; Norman, 2010). Implicit in this assertion is the assumption that reactivation effects should precede, or at least coincide with, the neural correlates of retrieval success. The findings of episodic retrieval studies employing neural measures with high temporal resolution, such as electroencephalography (EEG), are relevant to this issue. These studies have consistently identified a retrieval success correlate over the left posterior scalp that onsets within about 500 ms following cue presentation (Friedman and Johnson, 2000; Rugg and Curran, 2007). This EEG correlate has additionally been associated with subjective aspects of retrieval, as evidenced by its enhancement when subjects consciously recollect details or are highly confident about retrieval (Wilding and Rugg, 1996; Düzel et al., 1997; Curran, 2004). The effect thus serves as a landmark from which other retrieval-supporting processes, having distinct scalp topographies and time courses, can be dissociated. According to this framework, neural processes or events that contribute to retrieval success must occur quickly enough upon cue presentation so as to inform other neural events and, ultimately, the behavioral response. Alternatively, the timing of reactivation might follow retrieval success effects, such as is the case with neural correlates of post-retrieval processing (e.g., Wilding and Rugg, 1996; Hayama et al., 2008). In this latter case, reinstatement would be considered to reflect the maintenance of retrieved information in service of evaluative processes that vary with retrieval demands (Koriat and Goldsmith, 1996).

Recent studies have begun to investigate the sub-trial timing of retrieval-related reactivation with EEG (Wimber et al., 2012) and magnetoencephalography (MEG; Jafarpour et al., 2014) in order to overcome the relatively poor temporal resolution of fMRI. In one study, Jafarpour et al. (2014) employed MEG data and MVPA to investigate the reinstatement of neural activity that occurred early during a given encoding trial. Subjects studied words paired with scene or face stimuli, and then undertook a memory test in which they distinguished judgments associated with recollection from those based on high-confidence familiarity (with a variant of the "remember/know" task; Tulving, 1985). An early MEG component discriminating between the scenes and faces was identified at 180 ms after encoding stimulus onset. This component was then shown to be reactivated by around 500 ms after the onset of word retrieval cues that were designated with remember judgments. These effects are thus in the appropriate time window for contributing to retrieval success. However, Jafarpour et al. (2014) note that they had insufficient numbers of trials to test for reactivation associated with other retrieval judgments. As a result, it is unclear whether reactivation varied across judgments, as would be expected if it played a role in retrieval outcome (Kahn et al., 2004; Johnson and Rugg, 2007; Johnson et al., 2009). A similar critique can

be made regarding the results of the study by Wimber et al. (2012). Wimber et al. (2012) had subjects encode words in the context of flickering visual stimuli (6 and 10 Hz) and later make confidence judgments about the words. Whereas some aspects of the frequency information in the EEG data (orientation for 6 Hz and variance for 10 Hz) indicated greater reactivation during high-confidence judgments compared to misses, a secondary behavioral experiment showed that subjects were at chance in identifying the flicker associated with test items. These findings therefore raise the possibility that the reactivation signifies some implicit availability (i.e. priming) of encoding content rather than being functionally related to conscious retrieval. In the current study, we address this issue by employing encoding conditions that have previously been shown to elicit accurate source-memory judgments (McDuff et al., 2009).

The enhanced temporal resolution of reactivation effects described above brings with it another issue about timing that must also be addressed. Although identifying the sub-trial time course of reactivation during retrieval is important (as in Wimber et al., 2012; Jafarpour et al., 2014), the detection of such effects could also be limited by variability (mismatch) in the timing of activated patterns going from encoding to retrieval. For example, a pattern activated late during an item's encoding episode could be activated relatively early during the corresponding retrieval of that item; likewise, the order in which patterns are activated may also differ between the two phases. Systematic examples of this timing mismatch can be found in animal studies in which the reactivation of hippocampal neural firing sometimes occurs more rapidly, and even in the reverse order, compared to that during learning (Lee and Wilson, 2002; Foster and Wilson, 2006). Jafarpour et al. (2014) accounted for such variability, in part, by focusing on a neural pattern at a specific time during encoding and then testing for the reactivation of that pattern throughout the retrieval trial. However, it is reasonable to assume that representations and processes engaged at different time points of an encoding trial could also potentially be reactivated during retrieval of that item. Indeed, encoding tasks are typically structured to allow for this sort of elaborative processing by providing subjects with ample time (multiple seconds) on each item. There is a need therefore to not only track information across successive retrieval time points, but also maximize the accumulation of encoding-related information at each of those points. We accomplish this in the current study with a procedure recently used by Fuentemilla et al. (2010). In that study, pattern classifiers were trained on MEG data from an encoding period (3 s) but then tested on data from a longer period of working-memory maintenance (5 s). The key feature of this procedure is in tracking the *number* of reactivations during maintenance, as opposed to reactivation strength (cf. Johnson et al., 2009), allowing for timing disparities between the training and testing data (for similar application to longer, resting periods, see Staresina et al., 2013). Here, we extend this procedure to the sub-trial time courses of different processes involved in episodic retrieval, as identified using EEG.

The current study investigated the involvement of reactivation effects that are considered early and late, with respect to neural correlates of retrieval success, during episodic retrieval. Subjects completed an encoding phase in which they were presented with a series of words in the context of three tasks, designed to elicit distinct neural activity patterns. The *Artist* task involved thinking about how an artist would draw the item denoted by the word, the *Function* task instructed subjects to generate different functions for the item, and the *Cost* task directed subjects to think about the item's relative cost (cf. Johnson et al., 2009; McDuff et al., 2009). The use of different tasks for the encoding manipulation also allowed us to address whether early reactivation was restricted to stimulus-related (sensory) information from encoding (e.g., the visual flicker in Wimber et al., 2012; the faces and scenes in Jafarpour et al., 2014) or could be extended to the domain of task-related cognitive operations that are largely generated by subjects (e.g., given the Function cue, a number of possible functions

may come to mind). During a later retrieval phase, subjects made confidence-based recognition memory judgments. EEG data from encoding and retrieval were used in combination with MVPA to assess whether neural patterns distinguishing the tasks were reactivated during retrieval. Importantly, the MVPA approach employed here involved training several pattern classifiers separately on data from a series of 100-ms time bins at encoding. The performance of those classifiers was then aggregated across a series of time bins (also 100-ms each) during retrieval. The analyses thus provided the temporal resolution necessary to investigate the sub-trial time course of reactivation, and allowed these effects to be placed in the context of other, well-established EEG correlates of retrieval success. As described earlier, evidence of reactivation occurring shortly after retrieval cue onset (within ~500–800 ms) would be consistent with the standard proposal of the modeling literature that reinstatement contributes to retrieval success. Moreover, reactivation of encoding-related patterns beyond this early period would provide novel evidence in support of the idea that reinstatement is involved in maintaining the products of retrieval for further, post-retrieval evaluation.

Materials and methods

Subjects

Twenty-one University of Missouri (MU) students participated for either course credit or monetary compensation. All subjects were right-handed, native-English speakers, with no history of neurological disease. Informed consent was obtained in accordance with the MU Institutional Review Board. Data from three subjects were excluded from all analyses: one for excessive artifact in the EEG, and the other two due to poor behavioral performance ($d' \approx 0$). The final sample of 18 subjects (8 males) were 18 to 29 years of age (M = 21).

Stimuli

The stimulus pool consisted of 306 words drawn from the MRC database (Coltheart, 1981; Wilson, 1988; http://websites.psychology. uwa.edu.au/school/MRCDatabase/uwa_mrc.htm). The words were between 4 and 9 letters long (M = 5.5, SD = 1.3), had written frequencies between one and 50 per million (M = 16.9, SD = 13.2; Kucera and Francis, 1967), and had scores of at least 500 on scales of familiarity (M = 580.9, SD = 34.6), concreteness (M = 539.2, SD = 27.5), and imagability (M = 581.7, SD = 31.3). For each subject, 162 words were randomly selected to be presented during three encoding-phase blocks (54 words each). Within each block, equal numbers of words were randomly assigned to the three encoding tasks. These words were presented again during the retrieval phase, along with the 144 remaining words from the pool. The retrieval phase was also divided into three blocks, each with 54 old and 48 new words. The old words for each retrieval block were selected to equate the numbers from each encoding block and task. Forty-two additional words were used for instructions and practice. Stimulus presentation was controlled by the Cogent 2000 toolbox (http://www.vislab.ucl.ac.uk) in MATLAB (The MathWorks, Natick, MA). All words were shown in white upper-case 36-point Arial font on the black background of a 24-inch wide-screen LCD monitor (cropped to 1024 × 768 resolution) that was viewed at a distance of about 1 m.

Behavioral procedures

The experimental session lasted approximately 2 h. Instructions and practice on the encoding phase were administered first. Subjects were then fitted with an electrode cap and seated at a computer monitor where they completed three encoding blocks followed by three retrieval blocks. Blocks were separated by breaks of around three minutes each. Instructions and practice on the retrieval phase were delayed

until immediately prior to its start to prevent any influence on encoding processing. Following the final retrieval block, the electrode cap was removed, and subjects were debriefed.

Subjects first completed three encoding tasks for a series of words (one task per word). The *Artist* task required subjects to rate the difficulty (1 = "easy" to 4 = "hard") an artist would have drawing the item denoted by the word. The *Function* task required subjects to respond with the number of functions they could generate (from "1" to "4") for the item. The *Cost* task directed subjects to rate the relative cost of the object (1 = "low" to 4 = "high"; cf. Johnson et al., 2009; McDuff et al., 2009). Responses to each word were made by pressing keyboard keys with the right index through little fingers. To help subjects engage in distinctive processing for the three tasks, words were grouped into mini-blocks in which a single task was completed for three consecutive words. Each mini-block began with a 3-s display indicating the task and possible response options, above and below where the words were centrally presented. This display remained on the screen throughout the mini-block, but subjects were encouraged to refer to it only when necessary, in order to minimize eye movements. Each word was displayed for 5 s, with an asterisk appearing above the word for the final 2 s to indicate that subjects should make their response. The next word (or mini-block instruction) followed immediately. Mini-blocks were pseudo-randomly ordered to prevent consecutive completion of the same task.

The retrieval phase consisted of a recognition memory test on a series of intermixed old and new words. For each word, subjects were instructed to rate their confidence that the word was old or new on a six-point scale (1 = "sure new", 2 = "probably new", 3 = "maybe new", 4 = "maybe old", 5 = "probably old", and 6 = "sure old"). The responses were mapped respectively to keyboard keys pressed with the ring through index fingers of the left hand and the index through ring fingers of the right hand. The response options were displayed continuously at the bottom of the screen, but subjects were instructed to refer to them only when necessary. Test words were displayed for 2 s and replaced by a plus sign for 1 s until the next word. To obtain the initial reactions to the test words, subjects were encouraged to respond within the 2-s that the word was displayed. Responses occurring beyond the 2-s time window, along with omitted and multiple responses, accounted for about 1% of all trials and were not analyzed.

EEG acquisition and preparation

EEG was recorded continuously during both the encoding and retrieval phases. The EEG data were acquired with a BrainAmp Standard system (Brain Vision LLC; Durham, NC; http://www.brainvision.com) from 59 Ag/AgCl sintered ring electrodes embedded in an elastic cap (Easycap, Herrsching, Germany; http://www.easycap.de). The electrode locations were based on the International 10–20 system (Easycap montage no. 11) and included the following sites: Fpz/1/2, AFz/3/4/7/8, Fz, F1/2/3/4/5/6/7/8, FC1/2/3/4/5/6, FT7/8, Cz, C1/2/3/4/5/6, T7/8, CPz, CP1/2/3/4/5/6, TP7/8, Pz, P1/2/3/4/5/6/7/8, POz/3/4/7/8, and O1/2. The data were recorded with reference to an electrode at FCz, and a ground electrode was embedded in the cap at FT10. Two additional electrodes were adhered to the mastoids, and vertical and horizontal EOG data were recorded with electrodes adhered below the left eye and on the outer canthi. Prior to the start of the experiment, electrodes were adjusted until impedances were below 5 kΩ. The data were recorded at a sampling rate of 1000 Hz and an amplifier bandwidth of 0.01–100 Hz.

Offline preparation of the EEG data was implemented with the EEGLAB toolbox (Delorme and Makeig, 2004; http://sccn.ucsd.edu/eeglab/) in MATLAB. The continuous data were high-pass filtered (.05 Hz), down-sampled (200 Hz), re-referenced to linked mastoids, epoched (−200 to 2000 ms relative to item onset), and baseline-corrected according to the pre-stimulus period. Independent component analysis (ICA) was used to identify data appearing to correspond to

artifacts (e.g., eye movements, blinks, and muscle activity), which were manually rejected on the basis of scalp topography and power spectra (see Jung et al., 2000). Epochs were also rejected if the signal exceeded ± 100 mV. The data were then low-pass filtered (40 Hz).

EEG analysis

Multivariate analyses

Multivariate analyses of the EEG data were implemented with the Princeton Multi-Voxel Pattern Analysis toolbox (The Princeton Neuroscience Institute, Princeton, NJ; https://code.google.com/p/princeton-mvpa-toolbox/) and custom code in MATLAB. These analyses first involved using the EEG data from the encoding phase to train pattern classifiers to distinguish between the three tasks (Artist, Function, and Cost). The trained classifiers were then independently evaluated on data from the retrieval phase to provide measures of how well the prior encoding task associated with a test item could be determined. Classifier performance was taken as an indicator that encoding patterns were reactivated. The classification analyses were conducted separately for each subject.

Prior to classification, the data were divided into 100-ms time bins and the voltage amplitudes were averaged at each electrode and bin. Data from a particular bin at encoding were used to train a classifier, and that classifier was then tested on the data from each bin at retrieval, providing a time course of classification performance. The training–testing procedure was repeated for each bin at encoding, with the final results combined for reporting. Because the classifications were conducted independently on the data from each encoding bin, it is important to note that the classifiers for multiple time bins might be sensitive to similar patterns (i.e. reflecting sustained or revisited representations). Classifiers consisted of two-layer (no hidden layer) feed-forward neural networks, with the input layer consisting of the patterns of 59 voltages (one per electrode), and the output layer corresponding to the three tasks. For each classification there were equal numbers of input patterns (trials) for the two tasks. The classification parameters were similar to those we previously employed for fMRI data (Johnson et al., 2009; also see Polyn et al., 2005; McDuff et al., 2009), such that input–output weights were initialized to random values and adjusted with a scaled conjugate gradient descent version of backpropagation (Bishop, 1995; Duda et al., 2001). Classifier output was determined by a log-sigmoid transfer function. Training was stopped after 200 epochs or when the mean square error between the actual and computed outputs fell below .001. To reduce the prediction error associated with initializing the weights randomly, each classification was repeated 20 times, and the results were averaged over these repetitions.

The three tasks were classified in a pair-wise manner: Artist vs. Function, Artist vs. Cost, and Function vs. Cost (also see Kuhl et al., 2012). Classifier output consisted of a value between 0 and 1 for each task. For each classification, the output for the correct task was thresholded at .95 (regardless of output for the incorrect task), signifying what we refer to as a "reactivation" (following Fuentemilla et al., 2010). For a given time bin at retrieval, the reactivations were then summed over the classifiers from the 20 time bins (100-ms each) at encoding, resulting in a total of 20 possible reactivations per time bin and 400 reactivations over the 0–2000 ms recording epoch. To statistically test the reactivations, we constructed null distributions for each subject by conducting 500 classifications based on shuffled task labels for the training patterns. The permutation analyses were carried out using the same parameters described for the foregoing analyses (repeating each classifier 20 times and averaging, and using the same stopping criteria), and shuffling was done in a pseudo-random manner that maintained the constraint that each mini-block only corresponded to a single task. The resulting comparisons with these shuffled classifications took the form of paired T-tests. While the classifications and the corresponding permutation analyses were carried out separately for each

pair-wise combination of tasks, the results for a given trial (e.g., an Artist trial) were averaged over the corresponding combinations (in this example, Artist vs. Function and Artist vs. Cost) for reporting. Finally, to mitigate the inflated family-wise error rate (FWER) due to making multiple comparisons across the 100-ms time bins, we employed a cluster-wise correction procedure (Maris and Oostenveld, 2007; also see Osipova et al., 2006). This procedure involved contrasting each instance of the permuted data according to the primary contrast of interest ("sure old" vs. all other judgments; see below), and tracking the maximum number of consecutively-significant bins and the associated sum of T-values. These maximal values were then sorted to determine that a T-value sum of 6.40 within three consecutive time bins corresponded to a cluster-corrected threshold of P < .05.

A secondary goal of the EEG analysis was to identify neural activity that distinguished the different retrieval judgments, irrespective of the reactivation of encoding patterns. Due to the number of trials eliciting each confidence response, we compared "sure old" judgments (hereafter referred to as *SureOld*) to a category of responses that collapsed over the remaining judgments corresponding to low confidence and new (hereafter, *Unsure/New*). This classification was restricted to retrieval data by employing a leave-one-out cross-validation procedure in which data from two blocks of the retrieval phase were used for training, while data from the remaining (left-out) block served to test the classifier. The procedure was then repeated until each retrieval block served as testing data. These analyses used the same stopping criteria and repetition methods (to account for random initialization of weights) as those described above for the reactivation analyses. Due to SureOld trials being more frequent than the Unsure/New trials, we randomly selected a subset of the former to equate the number of input patterns for each classification. Additionally, because these analyses sought to identify retrieval correlates previously shown with univariate analyses to be consistent in timing across subjects, we trained and tested the classifiers within each 100-ms bin rather than fully crossing them. Doing so allowed us to make use of a simpler measure of classifier accuracy. Finally, to control the FWER, we conducted permutations in the same manner as described above for the reactivation analyses (Maris and Oostenveld, 2007). Similar to the results of the foregoing analyses, a cluster-wise corrected threshold of P < .05 corresponded to a T-value sum of 5.76 over three consecutive time bins.

To identify the electrodes that were most influential to the above classifications, we created "importance" maps by multiplying the trained network weight for a given electrode by its average voltage amplitude (normalized) during the training phase. Because we employed pair-wise (two-way) classifiers, evidence in favor of one condition is equivalent to evidence not in favor of the other task. In other words, when the weight and voltage are both positive for one condition, they are both negative for the other condition. Thus, we created importance maps for each of the two conditions involved in a classifier by using only the positive weights and positive voltages. Electrodes for which the voltage and weight were oppositely-signed were rare and are not reported (also see Johnson et al., 2009; McDuff et al., 2009; Newman and Norman, 2010). This procedure was carried out for each subject, and then the results were averaged across subjects. The resulting importance maps for (a) the classifications of task-related reactivation and (b) the classification of retrieval judgments (SureOld vs. Unsure/New) are reported in the Supplemental Material. For brevity, we provide the maps for 200-ms time bins; the 100-ms maps provided little in the way of novel insight, but are available on request.

Univariate analyses

Univariate EEG analyses of retrieval success employed two approaches. First, we directed the analyses to a time period and set of electrodes that have been previously shown to identify a prominent subjective-retrieval effect (e.g., Wilding and Rugg, 1996; Düzel et al., 1997; Curran, 2004). This analysis was based on the amplitude data during the 500–800 ms time period from a set of nine electrodes across

the scalp (Fz/3/4, Cz/3/4, and Pz/3/4). The data were submitted to an ANOVA that included factors of judgment (SureOld, Unsure/New), anterior/posterior electrode location (frontal, central, parietal), and electrode laterality (left, midline, right). The second approach involved assessing the temporal onset of differences between SureOld and Unsure/New judgments. This analysis was also based on averaged amplitude data from the nine electrodes used in the previous analysis. The key difference was that the data were contrasted separately at each electrode and 100-ms bin. Paired T-tests were used for these contrasts, with a critical T-value of 2.11 (P < .05, two-tailed).

Results

Behavioral performance

The response proportions made during the retrieval phase are summarized in Fig. 1A. Recognition performance was quite high, and items from the three encoding task conditions (Artist, Function, and Cost) appeared to give rise to similar levels of performance. To confirm this, a measure of accuracy, d′, was computed for each of the conditions, based on the overall hit and false alarm rates (regardless of confidence level). These data are provided in Fig. 1B. A one-way ANOVA indicated no significant differences in this accuracy measure across tasks ($F_{2,34} = 1.09$, P = .35). A similar pattern was evident when estimating the areas under binormal receiver-operating characteristic (ROC) curves constructed from the confidence data (M = .90, .91, and .90 for the Artist, Function, and Cost tasks, respectively; F < 1). Further analyses were focused on test items correctly judged as "sure old" (SureOld). Accuracy measures based only on the hits and false alarms that were designated by SureOld responses (d′$_{SureOld}$) are summarized for each encoding task in Fig. 1B. As with overall accuracy, a one-way ANOVA

of these data gave rise to no significant differences according to encoding task (F < 1). Because there were insufficient numbers of old items eliciting each of the lower-confidence and new judgments, we collapsed these items into a single category (Unsure/New). These judgments served as a comparison group in two-way ANOVAs, with factors of task and judgment, on the response proportions and associated response times (RTs). Both ANOVAs revealed main effects of judgment, indicating that SureOld judgments were more frequent ($F_{1,17} = 24.57$, P < .001; see Fig. 1A) and faster than Unsure/New judgments ($F_{1,17} = 65.66$, P < .001; see Fig. 1C). Follow-up ANOVAs restricted to the SureOld data further confirmed that there were no significant differences according to task among the proportions or RTs (both Fs < 1).

EEG results

Reactivation across all old items

The EEG analyses first followed an approach that has been employed numerous times with fMRI data to investigate reactivation (Polyn et al., 2005; Johnson et al., 2009; McDuff et al., 2009; Kuhl et al., 2011, 2012; Staresina et al., 2012). This approach involved training pattern classifiers to distinguish between experimental conditions at encoding – tasks in the current study – on the basis of the corresponding EEG data. The trained classifiers were then evaluated with EEG data from the retrieval phase. Given that the task-related information was not physically present at retrieval, classification performance depends on the degree to which encoding patterns are reactivated by the retrieval cues.

Prior to classification, the EEG data (voltage amplitudes) from the encoding phase were segregated into 100-ms time bins across the 2000-ms post-stimulus interval. Classifiers were trained to distinguish pairs of the three encoding tasks separately for each subject and for
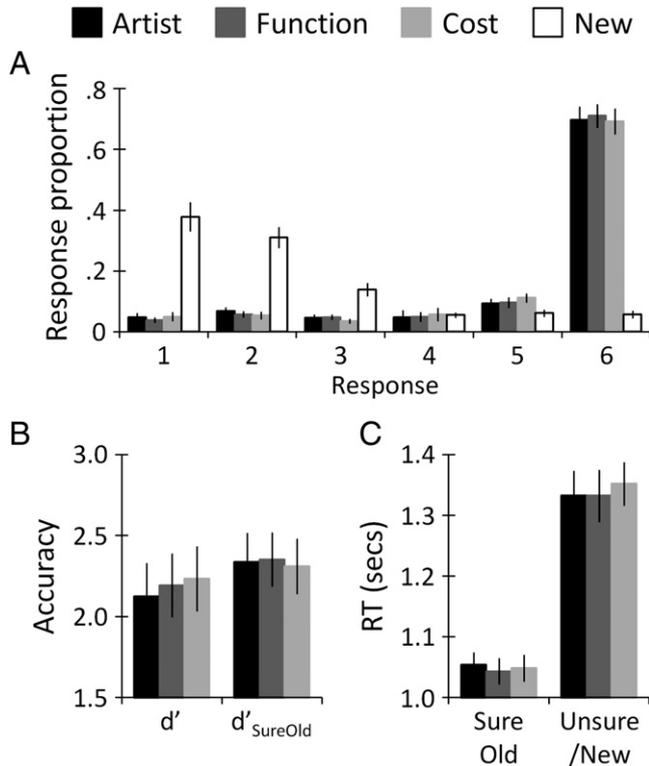

Fig. 1. Behavioral performance means (±SEM) during the retrieval phase. (A) Proportions of each confidence response (from 1 = "sure new" to 6 = "sure old") according to test item condition. (B) Accuracy according to overall hit and false alarm rates (d′) and hit and false alarm rates based on high-confidence old responses (d′$_{SureOld}$). (C) Response times (RTs) associated with old test items. SureOld refers to "sure old" responses, whereas Unsure/New constitutes all other responses.
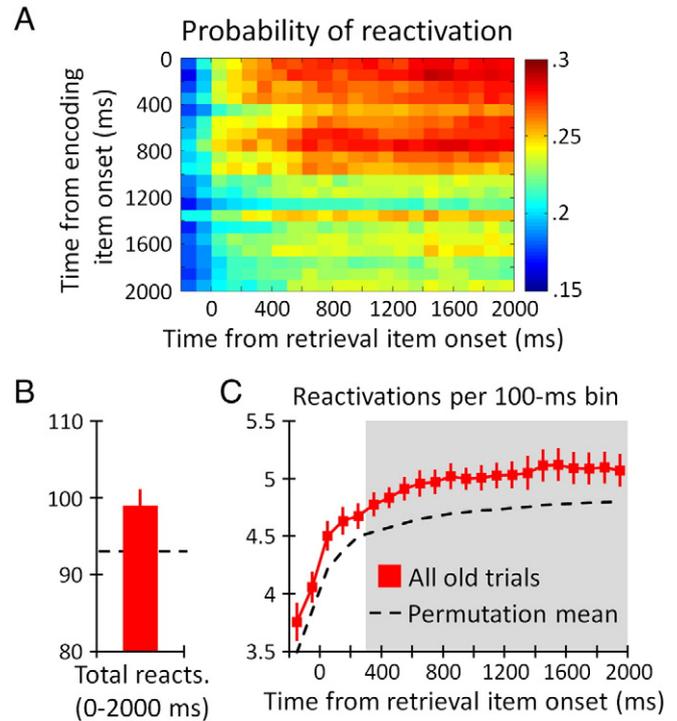

Fig. 2. Reactivation results for all old test items. (A) Across-subject mean probabilities of reactivation according to 100-ms time bins at encoding and at retrieval. (B) Mean (±SEM) total reactivations across the retrieval recording epoch. Dashed line indicates the mean total of reactivations determined by the permutation analysis, and the P-value corresponds to the permutation-based significance test. (C) Mean (±SEM) reactivations according to 100-ms bins at retrieval. Dashed line corresponds to the mean time course for the permutation analysis. Gray shaded area (300–2000 ms) indicates the bins showing significant effects.

each bin. Classifier testing was then based on data from an entire series of 100-ms time bins at retrieval, ranging from 200 ms pre-stimulus to 2000 ms post-stimulus onset. Fig. 2A displays the probabilities of reactivations, averaged over all subjects, according to the time of the pattern at encoding as well as the time of the reactivation at retrieval. Note that these data are presented here purely for descriptive purposes, but we revisit them later in more detail. As shown in the figure, the reactivated patterns came primarily from the early portion (before 1000 ms) of the encoding trials. This finding is likely due to the organization of encoding tasks into mini-blocks (see Behavioral procedures), which allowed subjects to prepare and maintain the appropriate task-related processing for the upcoming encoding items (as opposed to waiting for the item to appear before engaging in the task). Importantly, the prevalence of reactivations during a given retrieval trial appeared to increase from the pre-stimulus period to around 500-ms post-stimulus onset and were then sustained throughout the recording epoch. Finally, consistent with our hypothesis outlined in the Introduction, there appeared to be little evidence that the activation of encoding and retrieval patterns followed similar time courses, as would be indicated by the negative diagonal being highlighted in Fig. 2A.

Fig. 2B shows the total number of reactivations (M = 98.97, SD = 9.03) for a given old item during retrieval, collapsed over all of the time bins at encoding and irrespective of the retrieval response made. To assess this number statistically, we carried out 500 classifications for each subject in which the task labels for the training patterns were pseudo-randomly shuffled (see EEG analysis). For all old trials, the total number of reactivations across the recording epoch was significantly greater than the paired subject values derived from these permutations (T_{17} = 2.89, P < .01, one-tailed; M = 93.21 reactivations per trial for the permuted data). To investigate the sub-trial time course of reactivations, we segregated the classifier results according to 100-ms bins at retrieval. Fig. 2C displays these data for all old test items. As shown, the reactivations increased early on in the trial and reached an asymptote by around 600 ms post-stimulus onset. Furthermore, this level of reactivation was sustained through the end of the recording epoch and, as is apparent by Fig. 2A, is comprised of encoding patterns that themselves appear to be reactivated in a sustained manner, as opposed to being transient throughout the retrieval period. The data for each retrieval time bin were statistically tested with the aforementioned permutation procedure (here, segregating the permuted results by bin), revealing that the earliest time of significance was at 300–400 ms after item onset (T_{17} = 2.16, P < .025, one-tailed). (Note that, although the reactivations prior to this time period were not significantly different from the permutation results, there was a marked increase in reactivation from item onset. This could be due merely to visual activity resulting from item presentation which, given the blurred topographical nature of scalp EEG data, happened to be similar to task-related patterns.) The results for each subsequent bin also reached significance (range of Ps = .002 to .018). The extended time course of this effect also ensured its significance according to the cluster-wise threshold (see EEG analysis).

Reactivation according to retrieval outcome

We next investigated the occurrence of reactivation with respect to the different judgments made during the retrieval phase. Based on previous fMRI findings (e.g., Johnson and Rugg, 2007; Johnson et al., 2009; McDuff et al., 2009; Kuhl et al., 2011; Staresina et al., 2012), we hypothesized that more reactivations would be associated with SureOld relative to Unsure/New judgments. Analogous to the fMRI results, we collapsed the reactivation measures over the entire recording epoch at retrieval. Fig. 3A shows these data for a given old item according to its judgment designation. As with the findings for all old items, SureOld judgments were associated with significant reactivation (T_{17} = 2.99, P < .005, one-tailed; based on the paired permutation results for all old trials). By contrast, the total reactivations during trials in which
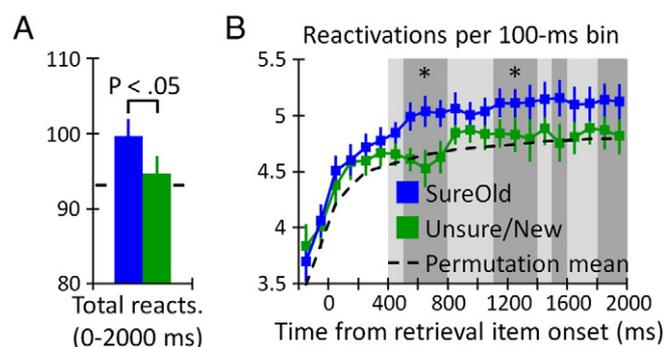

Fig. 3. Reactivation results according to retrieval judgment. (A) Mean (±SEM) total reactivations across the retrieval recording epoch according to SureOld (blue) and Unsure/New (green) judgments. Dashed line indicates the mean total of reactivations of the permutation analysis (for all old trials). P-value corresponds to the SureOld vs. Unsure/New comparison. (B) Mean (±SEM) reactivations according to 100-ms time bins at retrieval, segregated by judgment category. The number of reactivations differed significantly between judgments in the bins denoted by the dark gray shading. The asterisks indicate two time periods (three 100-ms intervals each) in which the SureOld > Unsure/New effects survived the cluster-wise significance level. The number of reactivations for SureOld judgments differed from chance in both the dark and light shaded areas.

old items elicited Unsure/New judgments did not reach significance (P = .25). Comparing these values to their corresponding permutation results within each judgment type gave rise to a similar pattern of significance (SureOld: T_{17} = 3.02, P < .005; Unsure/New: P = .28). The mean numbers of reactivations for the permutations were also similar for SureOld and Unsure/New: M = 93.14 and M = 93.40, respectively. To examine these data for different judgments in more detail, the time courses of reactivation, displayed in Fig. 3B, were also assessed. SureOld judgments were accompanied by significant reactivation starting during 400–500 ms (T_{17} = 2.42, P < .025, one-tailed), and the effects were sustained for all subsequent time bins (range of Ps = .001 to .014; also cluster-wise significant, given the extended time course). By contrast, test items designated with Unsure/New judgments were not associated with significant reactivation at any of the time points (all Ps > .05, uncorrected). (As before, we also compared the results to permutations segregated according to judgment, which resulted in similar findings. These data are available on request.)

More relevant to our primary hypothesis was whether there were differences in reinstatement according to retrieval judgment and, importantly, how these differences developed over the sub-trial time course. To address these issues, we directly compared the reactivations across judgments. At the entire-trial level (see Fig. 3A), the total number of reactivations for a given SureOld judgment (M = 99.69, SD = 9.85) was significantly greater than that for an Unsure/New judgment (M = 94.67, SD = 9.97; T_{17} = 2.00, P < .05, one-tailed). The time courses of these differences, shown in Fig. 3B, were tested with a two-way ANOVA with factors of judgment (SureOld, Unsure/New) and time (20 consecutive 100-ms bins). This analysis gave rise to a significant main effect of time ($F_{9,153}$ = 2.56, P < .01) and a significant judgment × time interaction ($F_{9,153}$ = 3.72, P < .001). Follow-up T-tests (one-tailed) for each time bin revealed significant differences from 500–800 ms (for the three respective bins: T_{17} = 2.47, 2.55, and 2.10; Ps = .012, .011, and .025) and at six subsequent bins (1100–1400, 1500–1600, and 1800–2000 ms; range of significant P-values: .027 to .049). Based on the cluster-corrected significance threshold, the two effects during 500–800 and 1100–1400 ms were significant.

As was shown for all old trials in Fig. 2A, the reactivated patterns during retrieval primarily came from the first half (0–1000 ms) of the encoding epochs. This raises a question, which can be addressed with the temporal resolution provided by EEG, about whether the reactivation effects for different retrieval judgments are driven by the same encoding patterns. The encoding time × retrieval time plots of the probabilities of reactivations (averaged over subjects) are provided
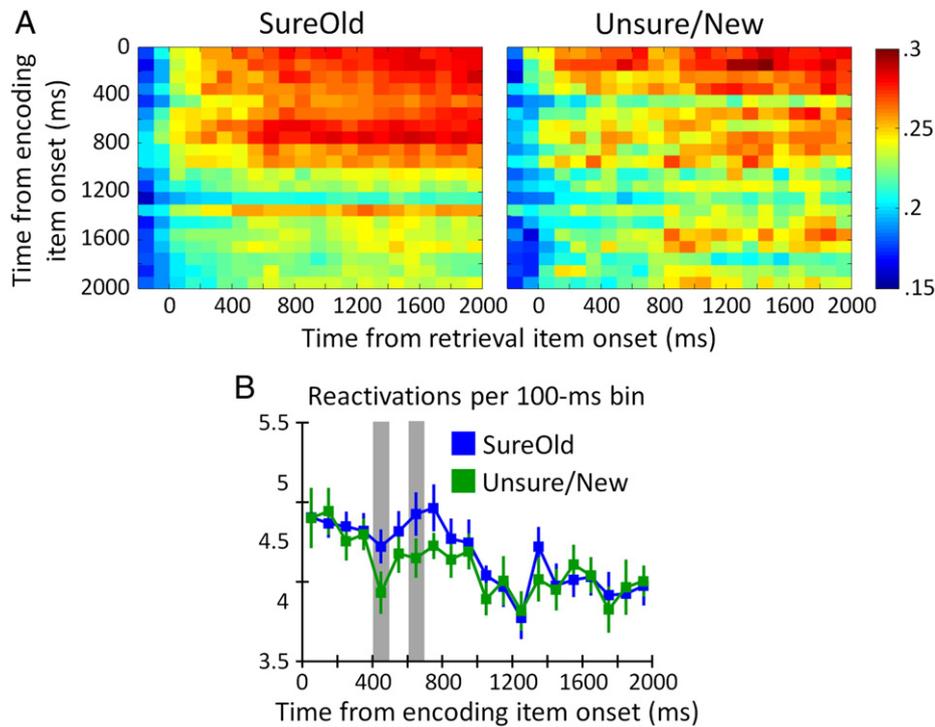
Fig. 4. Reactivation results according to encoding time. (A) Mean probabilities of reactivation according to 100-ms time bins at encoding and at retrieval, segregated according to SureOld and Unsure/New judgments. (B) Mean (±SEM) reactivations according to 100-ms bins at encoding, segregated by judgment category. Reactivations differed significantly between judgments in the bins denoted by the gray shading.

in Fig. 4A according to judgment. As shown, the decrease in reactivation for Unsure/New responses appears to be due largely to differences in the time period of about 400–800 ms post-stimulus onset. Fig. 4B displays these differences more clearly in the form of the number of reactivations with respect to encoding time (collapsing over retrieval time). To assess these differences, we submitted the data to an ANOVA that included factors of judgment and time (20 bins of 100-ms each at encoding). The ANOVA gave rise to significance of both main effects (judgment: $F_{1,17} = 50.11$, $P < .001$; time: $F_{9,153} = 1.97$, $P < .05$) and the interaction ($F_{9,153} = 2.65$, $P < .01$). Follow-up T-tests comparing the judgment effect at each bin indicated significance at 400–500 ms ($T_{17} = 2.81$, $P < .025$) and 600–700 ms ($T_{17} = 2.78$, $P < .025$), whereby the number of reactivations diminished going from SureOld to Unsure/New. Because each of these effects was short-lived, they did not pass the cluster-corrected significance threshold. However, as is apparent in Fig. 4 and was noted earlier for Fig. 2A, more reactivations came from the first half of the encoding epoch compared to the second half. An additional ANOVA that collapsed the 20 bins into two halves (0–1000, 1000–2000 ms) confirmed this with a significant effect of epoch half ($F_{1,17} = 14.135$, $P < .005$).

Reactivation relative to other retrieval correlates

To address the role of the aforementioned reactivation effects in retrieval processing, we next sought to put these effects in the context of a known EEG correlate of episodic retrieval. As noted in the Introduction, numerous studies have identified a correlate of retrieval success, maximal over left parietal electrodes and at around 500–800 ms post-stimulus onset, which has been shown to be further enhanced for judgments involving high levels of confidence as well as recollection (see Wilding and Rugg, 1996; Düzel et al., 1997; Curran, 2004). The foregoing analyses show that our reactivation effects coincide with the timing of this retrieval correlate. To confirm the presence of a retrieval correlate analogous to that described above for the current data set, we conducted a series of multivariate and univariate analyses. Because we were interested in identifying the earliest point at which the retrieval correlate was significant, we first conducted multivariate

(classification) analyses to capitalize on the potential increase in sensitivity over standard univariate analyses (see Jimura and Poldrack, 2012). Classifiers were trained to distinguish the EEG data at retrieval according to SureOld vs. Unsure/New judgments. Note that because the timing of the retrieval success correlate should be consistent across trials, we held constant the boundaries of the training and testing time bins rather than crossing all bins (as was the case for the foregoing analyses). The resulting time course of classifier accuracy is shown in Fig. 5A. Beginning at 500–600 ms, analyses confirmed that accuracy was significantly above chance (50%, given two judgment categories; $T_{17} = 2.33$, $P < .025$, one-tailed), and the effects continued through to the end of the recording epoch (range of Ps = 8×10^{-6} to .018; the extended time period of this effect ensured that it passed the cluster-wise significance threshold). (Note that the means of the permuted data also remained at 50% across all time bins.) The mean level of accuracy across this 500–2000 ms time window was 55% (SD = 3.9). Although this accuracy level appears modest in light of the consistency with which the left parietal effect has been demonstrated across numerous studies, the present results are likely limited by the inclusion of data from all electrodes; in other words, the focal nature of the parietal effect may lend itself better to univariate analysis. Additionally, because the SureOld and Unsure/New categories were consistently associated with different hands for responding, these results might be driven in part by motor preparation and/or execution. (The results of univariate analyses could be subject to the same confound, although counterbalancing the response hands across subjects would presumably alleviate any such differences. Controlling the influence of this confound on the accuracy measure from multivariate analysis would require a more elaborate approach such as switching response hands on a block-by-block basis, within subjects.) Importantly though, the classification results do not appear to onset any earlier than the results of typical univariate EEG analyses (also see below), allowing us to maintain our conclusions about the temporal correspondence of reactivation and retrieval success.

The retrieval success correlate was further tested using univariate analyses. In particular, we were again interested in the left parietal effect
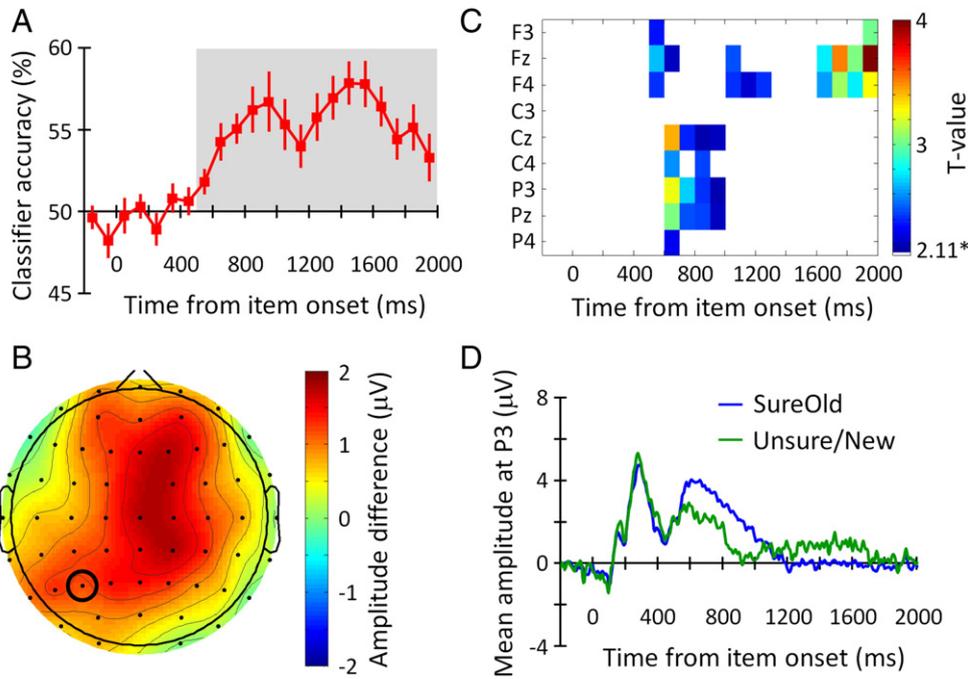
Fig. 5. EEG correlates of retrieval success. (A) Mean (± SEM) classifier accuracy of SureOld vs. Unsure/New judgments, according to 100-ms time bins at retrieval. Gray shading indicates when accuracy was significantly greater than chance (50%). (B) Topographical plot of the voltage amplitude differences between the SureOld and Unsure/New judgments during the 500–800 ms time period. Red indicates where SureOld was more positive-going. The P3 electrode is circled, and the corresponding data are plotted in panel D. (C) Results of comparing the voltage amplitudes for SureOld vs. Unsure/New judgments across time bins at retrieval. Y-axis refers to nine electrodes from representative scalp locations. *Denotes critical T-value. (D) Event-related potentials from the P3 electrode (see panel B) according to the SureOld and Unsure/New judgments at retrieval.

that onsets by about 500 ms following stimulus presentation. For this analysis, the average voltage amplitudes from a set of nine electrodes during the time period of 500–800 ms were submitted to an ANOVA that included factors of judgment (SureOld, Unsure/New), anterior/posterior electrode location (frontal, central, parietal), and electrode laterality (left, midline, right). The ANOVA gave rise to a main effect of judgment ($F_{1,17} = 5.61$, P < .05), a judgment × laterality interaction ($F_{2,34} = 6.30$, P < .01), and a three-way interaction ($F_{4,68} = 3.12$, P < .025). (Effects not involving the judgment factor are not reported here.) As shown in Fig. 5B, the judgment-related differences (SureOld > Unsure/New) during this period were rather distributed across the scalp. To further assess the localization and the onset latency of these differences, we conducted a series of T-tests contrasting the SureOld and Unsure/New data at each of the nine electrodes and in each 100-ms bin. The significant results of these analyses are

summarized in Fig. 5C. As is apparent from the figure, the earliest judgment-related differences were at frontal electrodes beginning at 500–600 ms. Shortly thereafter, effects over central and parietal electrodes reached significance, lasting from about 600 to 1000 ms. Finally, effects were evident over midline and right frontal electrodes later during the recording epoch (1600–2000 ms). For the data specifically from the hypothesized posterior region of the scalp (see Friedman and Johnson, 2000; Rugg and Curran, 2007), the maximal T-value was at a left parietal site (P3; see Fig. 5C). The voltage waveforms from this site are plotted in Fig. 5D.

The above findings that the earliest difference in reactivation roughly coincided with the onset of the left parietal retrieval correlate is prerequisite (although not necessarily sufficient) for claiming that reactivated information plays a direct role in retrieval success. In a final analysis, we extended these findings by investigating the
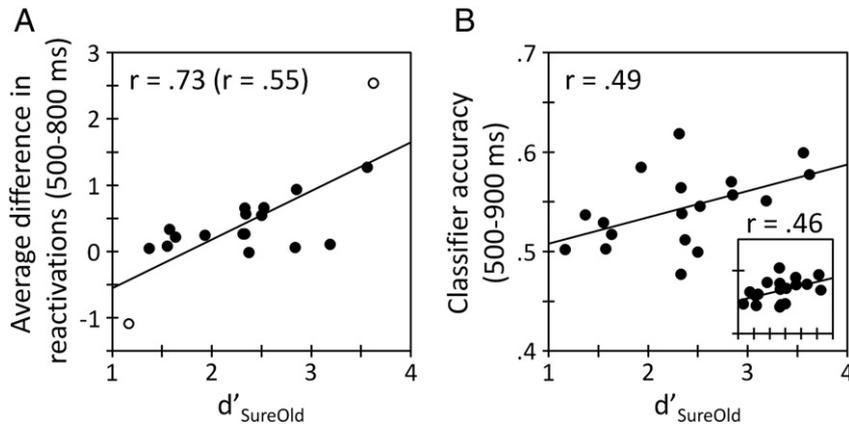

Fig. 6. Correlations between neural effects and behavior. (A) Plot of behavioral accuracy ($d'_{SureOld}$) vs. the reactivation differences from 500–800 ms at retrieval. The first r-value denotes the across-subject correlation including all subjects; the second r-value is based on the data excluding the apparent outliers (open circle data points). Both values were significant (P < .001 and P < .05, respectively). (B) Plot of behavioral accuracy vs. the classifier accuracy distinguishing SureOld and Unsure/New judgments during the 500–900 ms retrieval time period. The r-value at upper left denotes the across-subject correlation (P < .05). The inset and corresponding r-value correspond to data from 500–800 ms (P = .053).

relationship between reactivation and a measure of behavioral performance. This involved correlating the reactivation difference during the 500–800 ms interval (see Fig. 3B) with the level of memory accuracy based on SureOld responses ($d'_{SureOld}$) across subjects. The data for this analysis are provided in Fig. 6A. As shown, an increase in the difference in reactivations was associated with increased behavioral discrimination (Pearson r = .73, P < .001). Since two subjects had extreme values of the reactivation difference – one low and one high (−1.09 and 2.54, respectively) – we repeated the analysis while excluding those subjects. The correlation remained significant: r = .55, P < .05. A similar analysis testing for correlation between the classifier accuracy for the SureOld vs. Unsure/New distinction and behavioral performance also reached significance, but only when using data from a slightly longer interval of 500–900 ms (r = .49, P < .05; for 500–800 ms, r = .46, P = .053; see Fig. 6B). (Analyses directed at correlating behavioral performance with the reactivation differences occurring subsequent to the retrieval success correlate gave rise to no significant results, regardless of whether all data points [800–2000 ms] or only those showing significant differences [1100–1400, 1500–1600, and 1800–2000 ms] were usEd.)

Discussion

Several recent fMRI studies have demonstrated that neural activity elicited as episodes are encoded into memory is reactivated when those memories are retrieved (for reviews, see Rugg et al., 2008; Danker and Anderson, 2010; Rissman and Wagner, 2012), consistent with the notion that successful retrieval depends on hippocampally-mediated reinstatement (completion) of patterns of cortical activity. The temporal resolution of fMRI data, however, has been a limiting factor in the ability to test specific hypotheses regarding the functional role of reinstatement in episodic memory retrieval. In the present study, EEG measures of encoding-related reactivation during retrieval were employed to provide information about the precise timing of the effects (on the order of 100 ms), thereby allowing us to situate the reactivation effects in the context of other correlates – both EEG and behavioral – of retrieval success. The main finding was that the onset of significant reactivation was concurrent with the timing of an established EEG correlate of retrieval success that begins by about 500 ms following the retrieval cue presentation (for reviews, Friedman and Johnson, 2000; Rugg and Curran, 2007). Importantly, we report here that the prevalence of these reactivations differs according to the upcoming behavioral response (cf. Wimber et al., 2012; Jafarpour et al., 2014), suggesting that reinstatement might contribute to or signify the outcomes of retrieval attempts. Finally, the early-onsetting effects were accompanied by later differences in reactivation, occurring well beyond the time period of the retrieval success EEG correlate, potentially reflecting the involvement of reinstatement in post-retrieval processing.

The present findings constitute a significant advance in our understanding of the reinstatement during memory retrieval. As described in the Introduction, the existing fMRI evidence regarding this issue can be interpreted in at least two ways (for related discussions, see Johnson and Rugg, 2007; Johnson et al., 2008). One interpretation is that the reactivation of neural patterns occurs shortly after the presentation of a memory cue, as would be expected if it is closely tied to the conscious retrieval of episodic information (i.e. recollection; Hasselmo and Wyble, 1997; Norman and O'Reilly, 2003; Norman, 2010). Consequently, this form of reinstatement could have an immediate effect on the engagement of other retrieval processes, such as those sensitive to the subjective aspects of the retrieval experience. An alternative interpretation of fMRI-based reactivation is that it supports cognitive processes that occur subsequent to the recovery of episodic information but are still necessitated by the demands of the memory task (for similar arguments, see Maratos et al., 2001; Woodruff et al., 2005; Johnson and Rugg, 2007). Whereas the fMRI findings are ambiguous regarding this

functional distinction, the present study advances its resolution by demonstrating that reactivation effects are evident concurrently with and subsequent to the EEG correlate of retrieval success (see Fig. 5). Perhaps of utmost theoretical significance, therefore, is that it would be premature to attribute the body of fMRI evidence solely to the sort of early-onsetting reactivation effects that have been emphasized by recent EEG and MEG studies (Wimber et al., 2012; Jafarpour et al., 2014).

The foregoing interpretation that the reactivation of encoding patterns has direct bearing on other processes supporting retrieval, as well as on the subsequent decision, brings with it the following question: What do additional reactivations add to the phenomenological experience of retrieval? As described earlier, old items correctly judged with the "sure old" response were associated with more reactivations than were old items judged with the remaining responses. In addition, reactivation in the time period of 500–800 ms was correlated across subjects with highly-confident behavioral performance (see Fig. 6A). Taking the nature of our memory test at face value, the simplest answer to the above question is that the reactivations increased subjects' confidence that a test item was previously studied. Implicit in this question and answer is the assumption that neural reactivation gives rise to confidence-related processing and its neural correlates. We warn, however, that the current study provides no evidence for such a causal effect, as the results are only correlational (and largely simultaneous in timing). Our assumption instead has a theoretical underpinnings, based on findings that information (reactivation, here) is accumulated in service of making decisions (Ratcliff, 1978; also see Ratcliff et al., 2009) and that regions such as posterior parietal cortex serve to track and/or maintain the retrieved information (Wagner et al., 2005; Vilberg and Rugg, 2012). An alternative possibility, which we cannot rule out, is that task-related encoding information is retrieved as a consequence (i.e. downstream) of item-related retrieval. (We thank an anonymous reviewer for a helpful discussion on this issue.)

Both of the accounts described above are consistent with findings we reported previously, in which fMRI-based reinstatement decreased in magnitude from items judged with high confidence to those judged with lower confidence levels (including misses; Johnson et al., 2009). We must, however, qualify this across-study comparison, given that the high-confidence judgments in the two studies likely differed in an important way. Namely, whereas the previous study included an additional response option to segregate recollection-based memory judgments from those based on confidence, the current study used only confidence ratings. We know from several studies that high-confidence old responses in memory tasks where subjects are just allowed confidence response options (and not a recollection option) are associated not only with a strong sense of familiarity, but also with behavioral measures of recollection (for review, see Yonelinas, 2002). Furthermore, neuroimaging studies have repeatedly demonstrated that the neural correlates of highly-confident judgments – again, when confidence is the only response dimension – are similar to the neural correlates of recollection (e.g., compare Daselaar et al., 2006; Kim and Cabeza, 2009, to Yonelinas et al., 2005). There is no reason to expect that our high-confidence old response is not also comprised of a mixture of trials for which there is a strong sense of familiarity and trials accompanied by recollection. Thus, if these judgments are driven at least in part by recollective processes, they are perhaps described just as well as being consistent with the findings of several fMRI studies showing stronger reinstatement for recollection than for judgments lacking recollection (Kahn et al., 2004; Johnson and Rugg, 2007; Johnson et al., 2009; Staresina et al., 2012). This distinction is thus beyond the scope of the current study. Note, however, that this account does not hinge on whether recognition judgments are thought to rely on two distinct processes (i.e. recollection and familiarity; Yonelinas, 2002) as opposed to a single process (Donaldson, 1996; Dunn, 2004; Wixted and Mickes, 2010). Indeed, the utility of assessing

reinstatement, let alone any orthogonal neural measure, is that it provides a pathway to better understand the processes underlying different behavioral judgments or experimental conditions.

In the same manner that we highlight the reactivation differences from 500–800 ms post-stimulus onset, the effects evident later in the recording epoch also deserve further discussion. As noted earlier, these reactivation effects occurred in the time period following the retrieval success correlate and were apparent even well beyond the mean response times for recognition judgments (see Fig. 1C). A likely role of reactivation in this instance, therefore, seems to be in supporting post-retrieval processes. By this account, reinstatement could serve to maintain or manipulate the products of retrieval for further evaluation, as has been shown to be involved when subjects must make confidence ratings or source memory judgments beyond simple recognition decisions (Koriat and Goldsmith, 1996; Rugg et al., 2012). Several EEG studies have identified a sustained neural correlate of post-retrieval processing that is maximal over right frontal scalp, consistent with the findings reported for the current study (see Fig. 5C; also Wilding and Rugg, 1996; Hayama et al., 2008). Here, we demonstrate that this right frontal positivity for high-confidence old judgments is also accompanied by more frequent reactivation of encoding-related patterns for these same items, relative to those associated with less confidence or missed. Note that not all of the time points during this period exhibited significant reactivation differences according to the retrieval judgment. This inconsistency might be due to the fact that reinstatement is not engaged on all trials or that it is somewhat short-lived, whereby the effects occur earlier for some trials and later on others. A similar lack of uniformity has been demonstrated with the right frontal EEG correlate. Whereas it has been associated with retrieval success (e.g., Wilding and Rugg, 1996; Woodruff et al., 2006; and the current results), it is also occasionally more positive-going for incorrect judgments and recollection failure (Düzel et al., 1997; Trott et al., 1997), suggesting that those items for which memory is weaker elicit prolonged evaluation. It remains to be determined whether these later-onsetting retrieval correlates are less critical to the outcome of retrieval (as suggested by the absence of correlation between these neural effects and behavioral performance), and thereby less consistent than earlier correlates. Nevertheless, these post-retrieval reactivation effects highlight the importance of employing EEG to further investigate reinstatement, as their sustained nature might make them more likely to be detected with fMRI methods than the short-lived effects co-occurring with retrieval success.

The post-retrieval effects reported here also appear inconsistent with the findings of other studies employing EEG/MEG data to investigate reactivation. In particular, the effects reported by Wimber et al. (2012) lasted about 200–300 ms, returning to baseline by 500 ms after stimulus onset. In the study by Jafarpour et al. (2014), the reactivation effects deviated from chance at around 400 ms and lasted for only about 150 ms. There are two obvious explanations for these discrepant findings. First, the previous studies focused on encoding effects that appeared to reflect the processing of sensory-related information. Notably, the encoding effects in Wimber et al. (2012) were evident while the visual flicker remained on the screen, and the identification of effects during the encoding phase of Jafarpour et al. (2014) was restricted to a single, early time point (around 180 ms) likely corresponding to a low-level form of stimulus identification (Liu et al., 2002; also see Jafarpour et al., 2014, for discussion). By contrast, our encoding phase was designed to encourage elaborative processing about the word-task relationship, which may be more accessible to post-retrieval processes than stimulus attributes are (Wilding, 1999). Second, it is possible that our findings of sustained effects at encoding and retrieval are due largely to the analysis procedure being designed to capture such effects. Whereas Jafarpour et al. (2014) had subjects make multiple judgments on each retrieval trial, EEG data from only the first 800 ms were analyzed (and only 1000 ms in Wimber et al., 2012). Furthermore, by separately testing for the activation of different

patterns across the whole recording epoch at encoding, the present results might reflect holding information about the task in working memory for a period of time (~1 s; see Fig. 2A). Similar findings of sustained encoding-related activation were demonstrated in the context of a working memory task in the Fuentemilla et al. (2010) study from which we adapted our MVPA procedures. The current findings thus may represent an altogether different role of the involvement of reinstatement in retrieval, as opposed to the stimulus-driven reactivation in Wimber et al. (2012) and Jafarpour et al. (2014), in which the prolonged maintenance of representations at encoding influences the later retrieval of those items (Schon et al., 2004; Axmacher et al., 2008; also see Hasselmo and Stern, 2006).

To summarize, the EEG findings reported here constituted both early and sustained effects in which neural patterns from encoding were reactivated during retrieval. The reactivation effects evident in a time period concurrent with the left parietal EEG correlate were positively correlated with behavioral accuracy, further suggesting their reflection of reinstatement being directly involved in successful retrieval. Later reactivation effects, by comparison, are consistent with the maintenance of retrieved information in service of further, post-retrieval evaluation. Beyond these results, the present study highlights the limitations of using fMRI data alone to determine the functional significance of reactivation effects. Moreover, similar limitations apply to the use of scalp EEG data, in that we cannot make claims about the brain structures underlying these effects and, more importantly, whether reactivation is confined to the cortical level as opposed to mediated by the hippocampus (Hasselmo and Wyble, 1997; Norman and O'Reilly, 2003). As a result, we must consider the present results in the context of fMRI findings that hippocampal activity is functionally correlated with cortical reactivation (Staresina et al., 2013; Ritchey et al., 2013). Together, the combined findings from these different modalities and analytical approaches are key to further understanding the role(s) of reinstatement in the successful episodic memory retrieval.

Acknowledgments

Conflict of interest

The authors declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2014.11.013.

References

Alvarez, P., Squire, L.R., 1994. Memory consolidation and the medial temporal lobe: a simple network model. Proc. Natl. Acad. Sci. U. S. A. 91, 7041–7045.
Axmacher, N., Schmitz, D.P., Weinreich, I., Elger, C.E., Fell, J., 2008. Interaction of working memory and long-term memory in the medial temporal lobe. Cereb. Cortex 18, 2868–2878.
Bishop, C., 1995. Neural Networks for Pattern Recognition. Oxford University Press, New York.
Bower, G.H., 1972. Stimulus-sampling theory of encoding variability. In: Melton, A.W., Martin, E. (Eds.), Coding Processes in Human Memory. V.H. Winston & Sons, Washington, DC, pp. 85–123.
Coltheart, M., 1981. The MRC psycholinguistic database. Q. J. Exp. Psychol. A 33, 497–505.
Curran, T., 2004. Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. Neuropsychologia 42, 1088–1106.
Damasio, A.R., 1989. Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. Cognition 33, 25–62.
Danker, J.F., Anderson, J.R., 2010. The ghosts of brain states past: remembering reactivates the brain regions engaged during encoding. Psychol. Bull. 136, 87–102.

Daselaar, S.M., Fleck, M.S., Cabeza, R., 2006. Triple dissociation in the medial temporal lobes: recollection, familiarity, and novelty. J. Neurophysiol. 96, 1902–1911.

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

Donaldson, W., 1996. The role of decision processes in remembering and knowing. Mem. Cogn. 24, 523–533.

Duda, R., Hart, P., Stork, D., 2001. Pattern Classification, Second Edition. Wiley, New York.

Dunn, J.C., 2004. Remember-know: a matter of confidence. Psychol. Rev. 111, 524–542.

Düzel, E., Yonelinas, A.P., Mangun, G.R., Heinze, H.J., Tulving, E., 1997. Event-related brain potential correlates of two states of conscious awareness in memory. Proc. Natl. Acad. Sci. U. S. A. 94, 5973–5978.

Foster, D.J., Wilson, M.A., 2006. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. Nature 440, 680–683.

Friedman, D., Johnson Jr., R., 2000. Event-related potential (ERP) studies of memory encoding and retrieval: a selective review. Microsc. Res. Tech. 51, 6–28.

Fuentemilla, L., Penny, W.D., Cashdollar, N., Bunzeck, N., Düzel, E., 2010. Theta-coupled periodic replay in working memory. Curr. Biol. 20, 606–612.

Hasselmo, M.E., Stern, C.E., 2006. Mechanisms underlying working memory for novel information. Trends Cogn. Sci. 10, 487–493.

Hasselmo, M.E., Wyble, B.P., 1997. Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. Behav. Brain Res. 89, 1–34.

Hayama, H.R., Johnson, J.D., Rugg, M.D., 2008. The relationship between the right frontal old/new ERP effect and post-retrieval monitoring: specific or non-specific? Neuropsychologia 46, 1211–1223.

Haynes, J.D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. 7, 523–534.

Jafarpour, A., Fuentemilla, L., Horner, A., Penny, W., Düzel, E., 2014. Replay of very early encoding representations during recollection. J. Neurosci. 34, 242–248.

Jimura, K., Poldrack, R.A., 2012. Analyses of regional-average activation and multivoxel pattern information tell complementary stories. Neuropsychologia 50, 544–552.

Johnson, J.D., Rugg, M.D., 2007. Recollection and the reinstatement of encoding-related cortical activity. Cereb. Cortex 17, 2507–2515.

Johnson, J.D., Minton, B.R., Rugg, M.D., 2008. Content dependence of the electrophysiological correlates of recollection. NeuroImage 39, 406–416.

Johnson, J.D., McDuff, S.G., Rugg, M.D., Norman, K.A., 2009. Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. Neuron 63, 697–708.

Jung, T.P., Makeig, S., Westerfield, W., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111, 1745–1758.

Kahn, I., Davachi, L., Wagner, A.D., 2004. Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. J. Neurosci. 28, 4172–4180.

Kim, H., Cabeza, R., 2009. Common and specific brain regions in high-versus low-confidence recognition memory. Brain Res. 1282, 103–113.

Koriat, A., Goldsmith, M., 1996. Monitoring and control processes in the strategic regulation of memory accuracy. Psychol. Rev. 103, 490–517.

Kucera, H., Francis, W.N., 1967. Computational Analysis of Present-day American English. Brown University Press, Providence, RI.

Kuhl, B.A., Rissman, J., Chun, M.M., Wagner, A.D., 2011. Fidelity of neural reactivation reveals competition between memories. Proc. Natl. Acad. Sci. U. S. A. 108, 5903–5908.

Kuhl, B.A., Bainbridge, W.A., Chun, M.M., 2012. Neural reactivation reveals mechanisms for updating memory. J. Neurosci. 32, 3453–3461.

Lee, A.K., Wilson, M.A., 2002. Memory of sequential experience in the hippocampus during slow wave sleep. Neuron 36, 1183–1194.

Liu, J., Harris, A., Kanwisher, N., 2002. Stages of processing in face perception: an MEG study. Nat. Neurosci. 5, 910–916.

Maratos, E.J., Dolan, R.J., Morris, J.S., Henson, R.N.A., Rugg, M.D., 2001. Neural activity associated with episodic memory for emotional context. Neuropsychologia 39, 910–920.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190.

Marr, D., 1971. Simple memory: a theory for archicortex. Philos. Trans. R. Soc. Lond. Ser. B 262, 23–81.

McClelland, J.L., McNaughton, B.L., O'Reilly, R.C., 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. 102, 419–457.

McDuff, S.G., Frankel, H.C., Norman, K.A., 2009. Multivoxel pattern analysis reveals increased memory targeting and reduced use of retrieved details during single-agenda source monitoring. J. Neurosci. 29, 508–516.

Morris, C.D., Bransford, J.D., Franks, J.J., 1977. Levels of processing versus transfer appropriate processing. J. Verbal Learn. Verbal Behav. 16, 519–533.

Newman, E.L., Norman, K.A., 2010. Moderate excitation leads to weakening of perceptual representations. Cereb. Cortex 20, 2760–2770.

Norman, K.A., 2010. How hippocampus and cortex contribute to recognition memory: revisiting the complementary learning systems model. Hippocampus 20, 1217–1227.

Norman, K.A., O'Reilly, R.C., 2003. Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. Psychol. Rev. 110, 611–646.

Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. 10, 424–430.

Osipova, D., Takashima, A., Oostenveld, R., Fernandez, G., Maris, E., Jensen, O., 2006. Theta and gamma oscillations predict encoding and retrieval of declarative memory. J. Neurosci. 26, 7523–7531.

Polyn, S.M., Natu, V.S., Cohen, J.D., Norman, K.A., 2005. Category-specific cortical activity precedes retrieval during memory search. Science 310, 1963–1966.

Ratcliff, R., 1978. A theory of memory retrieval. Psychol. Rev. 85, 59–108.

Ratcliff, R., Philiastides, M.G., Sajda, P., 2009. Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. Proc. Natl. Acad. Sci. U. S. A. 106, 6539–6544.

Rissman, J., Wagner, A.D., 2012. Distributed representations in memory: insights from functional brain imaging. Annu. Rev. Psychol. 63, 101–128.

Ritchey, M., Wing, E.A., LaBar, K.S., Cabeza, R., 2013. Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. Cereb. Cortex 23, 2818–2828.

Rolls, E.T., 2000. Hippocampo-cortical and cortico-cortical backprojections. Hippocampus 10, 380–388.

Rugg, M.D., Curran, T., 2007. Event-related potentials and recognition memory. Trends Cogn. Sci. 11, 251–257.

Rugg, M.D., Johnson, J.D., Park, H., Uncapher, M.R., 2008. Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective. Prog. Brain Res. 169, 339–352.

Rugg, M.D., Vilberg, K.L., Mattson, J.T., Yu, S.S., Johnson, J.D., Suzuki, M., 2012. Item memory, context memory and the hippocampus: fMRI evidence. Neuropsychologia 50, 3070–3079.

Schon, K., Hasselmo, M.E., LoPresti, M.L., Tricarico, M.D., Stern, C.E., 2004. Persistence of parahippocampal representation in the absence of stimulus input enhances long-term encoding: a functional magnetic resonance imaging study of subsequent memory after a delayed match-to-sample task. J. Neurosci. 24, 11088–11097.

Shastri, L., 2002. Episodic memory and cortico-hippocampal interactions. Trends Cogn. Sci. 6, 162–168.

Staresina, B.P., Henson, R.N.A., Kriegeskorte, N., Alink, A., 2012. Episodic reinstatement in the medial temporal lobe. J. Neurosci. 32, 18150–18156.

Staresina, B.P., Alink, A., Kriegeskorte, N., Henson, R.N., 2013. Awake reactivation predicts memory in humans. Proc. Natl. Acad. Sci. U. S. A. 110, 21159–21164.

Teyler, T.J., DiScenna, P., 1986. The hippocampal memory indexing theory. Behav. Neurosci. 100, 147–154.

Tong, F., Pratte, M.S., 2012. Decoding patterns of human brain activity. Annu. Rev. Psychol. 63, 483–509.

Trott, C.T., Friedman, D., Ritter, W., Fabiani, M., 1997. Item and source memory: differential age effects revealed by event-related potentials. Neuroreport 8, 3373–3378.

Tulving, E., 1985. Memory and consciousness. Can. Psychol. 26, 1–12.

Tulving, E., Thomson, D.M., 1973. Encoding specificity and retrieval processes in episodic memory. Psychol. Rev. 80, 352–373.

Vilberg, K.L., Rugg, M.D., 2012. The neural correlates of recollection: transient versus sustained fMRI effects. J. Neurosci. 32, 15679–15687.

Wagner, A.D., Shannon, B.J., Kahn, I., Buckner, R.L., 2005. Parietal lobe contributions to episodic memory retrieval. Trends Cogn. Sci. 9, 445–453.

Wheeler, M.E., Petersen, S.E., Buckner, R.L., 2000. Memory's echo: vivid remembering reactivates sensory-specific cortex. Proc. Natl. Acad. Sci. U. S. A. 97, 11125–11129.

Wilding, E.L., 1999. Separating retrieval strategies from retrieval success: an event-related potential study of source memory. Neuropsychologia 37, 441–454.

Wilding, E.L., Rugg, M.D., 1996. An event-related potential study of recognition memory with and without retrieval of source. Brain 119, 889–905.

Wilson, M., 1988. Psycholinguistic database: machine-usable dictionary, version 2.00. Behav. Res. Methods Instrum. Comput. 20, 6–10.

Wimber, M., Maaβ, A., Staudigl, T., Richardson-Klavehn, A., Hanslmayr, S., 2012. Rapid memory reactivation revealed by oscillatory entrainment. Curr. Biol. 22, 1482–1486.

Wixted, J.T., Mickes, L., 2010. A continuous dual-process model of remember/know judgments. Psychol. Rev. 117, 1025–1054.

Woodruff, C.C., Johnson, J.D., Uncapher, M.R., Rugg, M.D., 2005. Content-specificity of the neural correlates of recollection. Neuropsychologia 43, 1022–1032.

Woodruff, C.C., Hayama, H.R., Rugg, M.D., 2006. Electrophysiological dissociation of the neural correlates of recollection and familiarity. Brain Res. 1100, 125–135.

Yonelinas, A.P., 2002. The nature of recollection and familiarity: a review of 30 years of research. J. Mem. Lang. 46, 441–517.

Yonelinas, A.P., Otten, L.J., Shaw, K.N., Rugg, M.D., 2005. Separating the brain regions involved in recollection and familiarity in recognition memory. J. Neurosci. 25, 3002–3008.