

Replay of Very Early Encoding Representations during Recollection

Anna Jafarpour,^{1,2} Lluís Fuentemilla,^{3,4} Aidan J. Horner,^{1,5} Will Penny,⁶ and Emrah Düzel^{1,2}

¹Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom; ²Institute of Cognitive Neurology and Dementia Research, Medizinische Fakultät, 39120 Magdeburg Otto-von-Guericke, Germany; ³Cognition and Brain Plasticity Unit, Institute of Biomedicine Research of Bellvitge, Hospital Duran i Reynals, 08908 Hospitalet de Llobregat, Barcelona, Spain; ⁴Department of Basic Psychology, University of Barcelona, Secretaria 08035 Barcelona, Spain; and ⁵Institute of Neurology and ⁶Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, United Kingdom

Long-term memories are linked to cortical representations of perceived events, but it is unclear which types of representations can later be recollected. Using magnetoencephalography-based decoding, we examined which brain activity patterns elicited during encoding are later replayed during recollection in the human brain. The results show that the recollection of images depicting faces and scenes is associated with a replay of neural representations that are formed at very early (180 ms) stages of encoding. This replay occurs quite rapidly, ~500 ms after the onset of a cue that prompts recollection and correlates with source memory accuracy. Therefore, long-term memories are rapidly replayed during recollection and involve representations that were formed at very early stages of encoding. These findings indicate that very early representational information can be preserved in the memory engram and can be faithfully and rapidly reinstated during recollection. These novel insights into the nature of the memory engram provide constraints for mechanistic models of long-term memory function.

Introduction

Recollection is associated with reexperiencing details of events, such as the scenery in which it took place or the faces of individuals who were present (Tulving, 1985). There is now converging evidence that brain activity patterns that participated in representing aspects of these event characteristics during encoding can be later reinstated or “replayed” at retrieval (for review, see Düzel et al., 2010). An intriguing puzzle in memory research is that cortical representations of event contents, such as faces, emerge very rapidly, within 200 ms (McCarthy et al., 1999; Puce et al., 1999; Fisch et al., 2009; Rössion and Caharel, 2011), whereas encoding processes in brain regions that are critical for recollection (i.e., the hippocampus and surrounding medial temporal areas) are initiated at 200 ms (Fell and Axmacher, 2011) and require several hundred milliseconds to unfold, as evidenced in invasive recordings of neural oscillations (Lega et al., 2012) and slow potentials (Fernández et al., 1999; Axmacher et al., 2010). These discrepancies in timing raise the question of whether rapidly emerging cortical event representations formed at early

stages of encoding are conserved in long-term memory and thus can be later replayed.

To capture the precise temporal evolution of neural representations during memory encoding and retrieval in the human brain, we used magnetoencephalography (MEG)-based multivariate pattern classifiers (MVPCs) as outlined in Jafarpour et al. (2013b). MEG combines the advantage of high temporal resolution with sampling neural activity from almost the entire cortical mantle. This distributed sampling accounts for the possibility that the cortical representation of a perceived event is distributed (Marr, 1971; Haxby et al., 2001; Hoffman and McNaughton, 2002; Fries et al., 2003) and may be spatially recoded in the course of encoding. It should be noted that methodological approaches based on univariate statistics are also suitable to detect neural reactivation, as has recently been demonstrated with sophisticated experimental designs using perceptual features of stimuli (e.g., flickering frequency and visual lateralization; Waldhauser et al., 2012; Wimber et al., 2012). For a systematic comparison between univariate and multivariate decoding and advantages of multivariate analyses, see Jafarpour et al. (2013a).

Healthy young adults were instructed to encode images of scenes and faces that were paired with words (Fig. 1B). Later, the words were used to probe image recollection (Fig. 1C). We trained MVPCs to decode oscillatory (8–45 Hz) brain activity responses to images of faces and scenes during encoding, when only the images were on the screen. MVPC analysis was performed every 66 ms, permitting us to capture the temporal evolution of neural representations. We then used classifiers that successfully classified the oscillatory activities into faces and scenes to detect the timing of replay of the same neural activity

Received May 3, 2013; revised Oct. 14, 2013; accepted Oct. 29, 2013.

Author contributions: L.F. and E.D. designed research; L.F. performed research; A.J. and W.P. contributed unpublished reagents/analytic tools; A.J., L.F., A.H., W.P., and E.D. analyzed data; A.J., L.F., A.H., W.P., and E.D. wrote the paper.

This work was supported by the Spanish Government (Grant PSI2010–15024) and the Ramón y Cajal program. We thank David Bradbury for support during data collection and to thank Nico Bunzeck with help on data collection and analysis and Wellcome Trust Centre for Neuroimaging at UCL for providing facilities.

The authors declare no competing financial interests.

Correspondence should be addressed to Anna Jafarpour, Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK. E-mail: annajafarpour@gmail.com or a.jafarpour@ucl.ac.uk.

DOI:10.1523/JNEUROSCI.1865-13.2014

Copyright © 2014 the authors 0270-6474/14/330242-07\$15.00/0

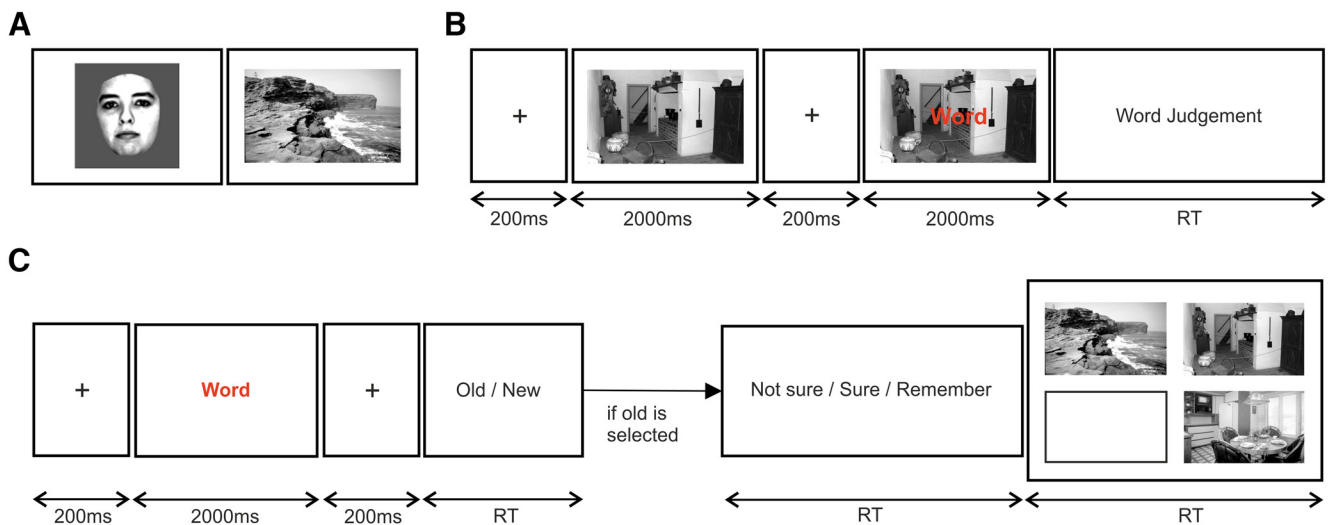


Figure 1. Schema of experimental paradigm. **A**, Samples of stimuli used in this experiment: a face and a scene. **B**, Schema of experiment pipeline during encoding. We trained and tested the classifier for decoding image category (faces or scenes) based on MEG oscillatory activity (8–45 Hz) at different time bins when the images were shown alone (without the word). **C**, Schema of experimental pipeline during retrieval (when the word recognition Old/New response is “Old”). The replay of associated image after onset of the cue (the paired word) was decoded.

pattern at retrieval, when the word associated to the image was shown as a memory cue.

Materials and Methods

Participants. Eleven right-handed healthy adults with normal or corrected-to-normal vision participated in this experiment (6 females; mean age 23 ± 2 years). All participants gave written informed consent to participate. The study was approved by the University College London Research Ethics Committee for Human-Based Research. All participants were financially compensated for their participation.

Experimental design. The experiment contained six runs consisting of two separate phases: the study (encoding) and the test (retrieval) phase. An arithmetic distraction task separated the two phases. In the study phase of each run, participants were required to memorize a set of 20 trial unique images associated to 20 trial unique words. All images were gray scaled and normalized to a mean gray value of 127 and SD of 75, of dimensions 300×300 pixels, and shown upon a gray background (gray value of 127) subtending ~ 6 degrees of horizontal and vertical visual angle. In each run, images were randomly selected from faces (5 female and 5 male) or scenes (5 indoor and 5 outdoor; Fig. 1A). The paired words denoted either living (50%) or nonliving (50%) objects with a Kucera–Francis frequency of 20–24. Image–word associations were not semantically related and were shown only once during encoding and randomized across participants.

Participants were instructed to learn the association between the image and the word. For each association, scene or face images were presented for 2000 ms preceded and followed by a 1500 ms fixation period. Immediately thereafter, the same image appeared with the associated word in red for 3000 ms on top of it, which was followed by a living/nonliving judgment about the word (responding with the index or middle finger of their right hand). After a random intertrial interval of 1500, 2000, or 2500 ms, the next image and image–word association was presented. An arithmetic task of 5 min separated the Study and Test phases to eliminate active rehearsal of the last image–word pairs studied in each run (Fig. 1B).

In the test phase, a word (in red) was presented for 2000 ms. Afterward, 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. In each run, 20 Old and 20 New words were presented in a randomized order. Thereafter, a confidence judgment task (2000 ms) followed. Here, new judgments were followed by “Sure/Not sure” and old judgments were followed by “Remember/Sure/Not sure.” Participants were instructed to make confidence judgments following

old judgments with respect to their ability to recollect the image associated to that word at encoding. They responded “not sure” when they did not have any memory for the associated image, “sure” when they believed that they could recognize what image was associated to the word, and “remember” when they had the associated image vividly in mind.

Each trial ended with a source memory (image-selection) test during which three images and an empty square were presented in the four corners of the screen. The three images always included the face/scene originally paired with the word and two familiar images (i.e., presented in the Study phase with different words) from the same category as the paired image. Participants were required to select, within a 3000 ms time limit, which of these images was paired with the word or had the opportunity to select the empty square if they could not identify the match. A random intertrial interval of 1500, 2000, or 2500 ms preceded the beginning of the next trial. After the test phase, participants had a short rest period before the next run (Fig. 1C).

MEG recordings. MEG data were recorded with a 274 channel CTF Omega whole-head gradiometer system (VSM MedTech) with a 600 Hz sampling rate. Head position inside the system was tracked via head localizer coils attached to the nasion and 1 cm anterior to the left and right preauricular points. Participants were seated upright and the stimuli were back projected onto a screen 1 m in front of them.

MEG preprocessing and data preparation. Data were preprocessed using MATLAB 2009 and SPM8 (www.fil.ion.ucl.ac.uk/spm/). The main noise (49–51 Hz) was filtered out of the data and then MEG single-trial epochs of -1000 to 2500 ms relative to the onset of the images on the screen in the study phase, when the images were shown for the first time, were extracted and baseline corrected (subtraction by the average amplitude of the epoch). Next, the signals from individual trials were transferred to the time-frequency (TF) domain using 5 cycle Morlet wavelets. For the multivariate analysis, 38 wavelets were used for this transformation (from 8 to 45 Hz in steps of 1 Hz) and the power of the TF signal was calculated. The 8 to 45 Hz frequency range covered a broad range of frequencies without compromising temporal resolution too much by including lower frequencies. In addition, a model based on information theory suggests that power decreases in the alpha/beta frequency range reflect information coding in long-term memory (Hanslmayr et al., 2012), and the 30–40 Hz frequency range is suggested to include physiognomic information about faces (Gao et al., 2013). The TF-transformed data were then down-sampled to 300 Hz and normalized by z-scoring the power value at each time, frequency, and channel across trials.

For three subjects, one run of six experimental runs was discarded. Two of the subjects did not follow the experimental instructions during the first run. For one subject, there was a problem with data acquisition in

the last run. Therefore, for these three subjects, five of six runs were analyzed.

Pattern classifier analysis. A support vector machine (SVM) with a linear kernel (Vapnik, 2000) was used to classify the preprocessed MEG signals of face versus scene samples. The SVM algorithm is implemented in the MATLAB bioinformatics toolbox. A pattern classifier was trained on MEG TF responses elicited when images of scenes and faces were shown at encoding (i.e., when the face/scene was first displayed on the screen, without the associated word). There were 60 samples of faces and scenes for seven of the analyzed subjects and 50 samples of faces and scenes for three subjects. One subject was excluded due to behavioral memory performance. We used an equal number of samples from each category (random selection) for training and tested an equal number of the remaining samples from each category (by random selection). The classification accuracy reported here is the performance of the classifier averaged over categories (faces and scenes), subjects, and cross-validation folds (see below). The classifiers were trained separately for each participant and time bin. We used 13 time bins, each of duration 66 ms, and centered at $-19, 46, 113, 180, 246, 313, 380, 446, 513, 580, 646, 713,$ and 780 ms relative to stimulus onset. Each of the classifiers used spectral power in 274 MEG channels and at 21 time points within each time bin. For each time bin, there were therefore 218,652 possible features ($274 \text{ channels} \times 21 \text{ time points} \times 38 \text{ frequencies}$).

For each of the 13 pattern classifiers (i.e., time bin) 10-fold cross-validation was adopted for validating the accuracy of the trained model. Accordingly, 10 classification iterations were run and 10% of samples from each category were left out at each iteration for testing the accuracy of the classifier. Before training, in each cross validation iteration, a feature-selection step was conducted by performing a univariate statistical analysis across the training set (excluding the validation set) on spectral power at each frequency, time point, and channel that constituted the features for the classifier. The testing dataset was never included in the feature selection step. Those features that were found to be significantly different between categories by a two-tailed paired Student's *t* test ($p < 0.05$) were selected. This data-led process served to reduce the dimension of the pattern classification problem by 95%. In each cross-validation iteration, the model was used to predict the category of the left-out trials (i.e., test trials). The classification performance was calculated as the average across the cross-validation iterations.

Classification performance at encoding was further investigated as follows. First, we tested whether the classification accuracy during encoding relied on the event-related field (ERF) component (M170; Liu et al., 2002; Gao et al., 2013). For each subject, we averaged the (low-pass-filtered 45 Hz) signal to obtain ERFs in each category and subtracted the average category-specific ERFs from the signal in each trial. The resulting signal was preprocessed as mentioned above (exactly the same as for the original signal) and cross-validation was again the same process as for the original data. Second, the time bins in which pattern classifiers performed significantly above chance (in the main analysis) after multiple-comparisons correction were selected and then classifiers were trained on all trials from that encoding time bin. The trained classifiers were then used to classify all time bins during encoding. This analysis was performed to assess whether the spatiotemporal frequency patterns that consistently contributed to classification at a specific time bin (e.g., 180 ms) were repeated at other time bins during encoding.

We next analyzed the retrieval data in a similar fashion. We first selected time bins from encoding that showed significant classification performance in the initial cross-validation analysis. We trained classifiers for each time bin using all the trials for those encoding bins and tested on each time bin at retrieval, in which memory for the images was cued with the associated word. Testing was performed at 13 separate time bins: $-19, 46, 113, 180, 246, 313, 380, 446, 513, 580, 646, 713,$ and 780 ms from onset of the memory cue (the same time bins used in the encoding phase). The classification accuracy was calculated in relation to the category of the paired image (i.e., the image that the participant should have successfully retrieved). We studied retrieval in two steps. First, we looked at replay in all the trials when the words were recognized correctly as "Old" (recognition hits). In the second step, we analyzed the trials in

which the image associated with the word was selected correctly (source memory hits; recollection).

Between-subject ("second-level") analysis of classification accuracy was implemented using SPM8 for MEG data. To test the accuracy of the classifiers against chance (i.e., 50%) we used a one-sample *t* test with a correction for multiple comparisons (familywise error; FWE) using random field theory (RFT) implemented in SPM8 (Kilner et al., 2005; Litvak et al., 2011). As is standard in neuroimaging, we made inferences using a cluster-level threshold. RFT procedure adjusts the *p*-value statistics as a function of number of time points (classification repetition here). Such adjustment is similar to Bonferroni correction; however, Bonferroni correction is suitable for datasets that are independent at each repetition (or data point), whereas here, time-frequency data are naturally not independent of adjacent time points and RFT is more suitable for multiple-comparisons correction (Kilner et al., 2005). To avoid numerical problems (e.g., infinite *z*-scores) in the input data for second-level analysis in SPM8, we changed any 100% and 0% classification accuracies to 99.9% and 0.01%, respectively (*z*-scores of which are 3 and -3 , respectively).

Cluster-level FWE-corrected *p*-values were used to examine the classification accuracy during encoding and retrieval of recognition hit trials. Follow-up decoding analyses were conducted using source memory hit trials and, for these analyses, only time windows showing replay in recognition hit trials were considered. For these targeted analyses, we used the conservative Bonferroni-corrected α level for *t* tests.

Time-frequency analysis. For a post hoc classical univariate TF analysis at retrieval, similar to the preprocessing steps used for pattern classification, 5 cycle Morlet wavelets from the 3–45 Hz frequency range in steps of 1 Hz were used. The power was then transformed to logarithmic scale and baseline corrected by the average power in a -150 to 0 ms time window relative to onset of the word cue. For the second-level analysis, we used paired *t* tests in SPM. SPM employs an FWE-corrected statistical threshold (set at $p < 0.05$) for extracting the significance of statistical results (Litvak et al., 2011).

In the second-level analysis, we assessed spectral power differences (ranging from 3 to 45 Hz) at the time window during which MVPA indicated memory replay: 400–550 ms. The averaged power over the 400–550 ms time window was calculated for each frequency and channel and then compared between hits and correct rejections (CRs). A similar analysis was done for the 250–400 ms and 550–700 ms time windows, which are the adjacent time windows to 400–550 ms and have the same time length. In the final step of the second-level analysis, we compared the power (in the time windows of interest) between source hits and recognition misses using the same time windows.

Results

Behavioral results

Behaviorally, participants recognized the words at test (corrected hit rate: $M = 87.96\%$ and $SD = 5.05\%$) equally well regardless of the category of the paired image (hit rate for words associated with faces: $M = 87.21\%$ and $SD = 6.85\%$ and hit rate for words associated with scenes: $M = 88.84\%$ and $SD = 5.72\%$, paired-sample *t* test: $t_{(9)} = -0.599, p = 0.563$). However, their source memory for scenes (hit rate: $M = 80.11\%$ and $SD = 11.83\%$) was better than for faces (hit rate: $M = 67.17\%$ and $SD = 16.82\%$; $t_{(9)} = 2.91, p = 0.017$). Repeated-measures ANOVA, conducted on the source memory test as a function of source memory confidence and image category, revealed no main effect of image category ($F_{(1,9)} = 1.21, p = 0.276$); but there was a significant effect of confidence level ($F_{(2,18)} = 31.46, p < 0.001$) and a confidence \times image category effect ($F_{(2,18)} = 6.96, p = 0.002$). Post hoc paired-sample *t* tests indicated that subjects had more confidence ("Remember") in selecting the correct scenes than the correct faces (for faces: $M = 33.19\%$, $SD = 22.76\%$, and for scenes: $M = 57.40\%$, $SD = 19.44\%$; $t_{(9)} = 5.97, p < 0.001$) and they said "Sure" more frequently for correctly selecting faces (for faces: $M = 25.85\%$, $SD = 14.32\%$, and for scenes: $M = 15.23\%$, $SD = 9.29\%$; $t_{(9)} = 4.77, p = 0.001$), but accuracy did not significantly

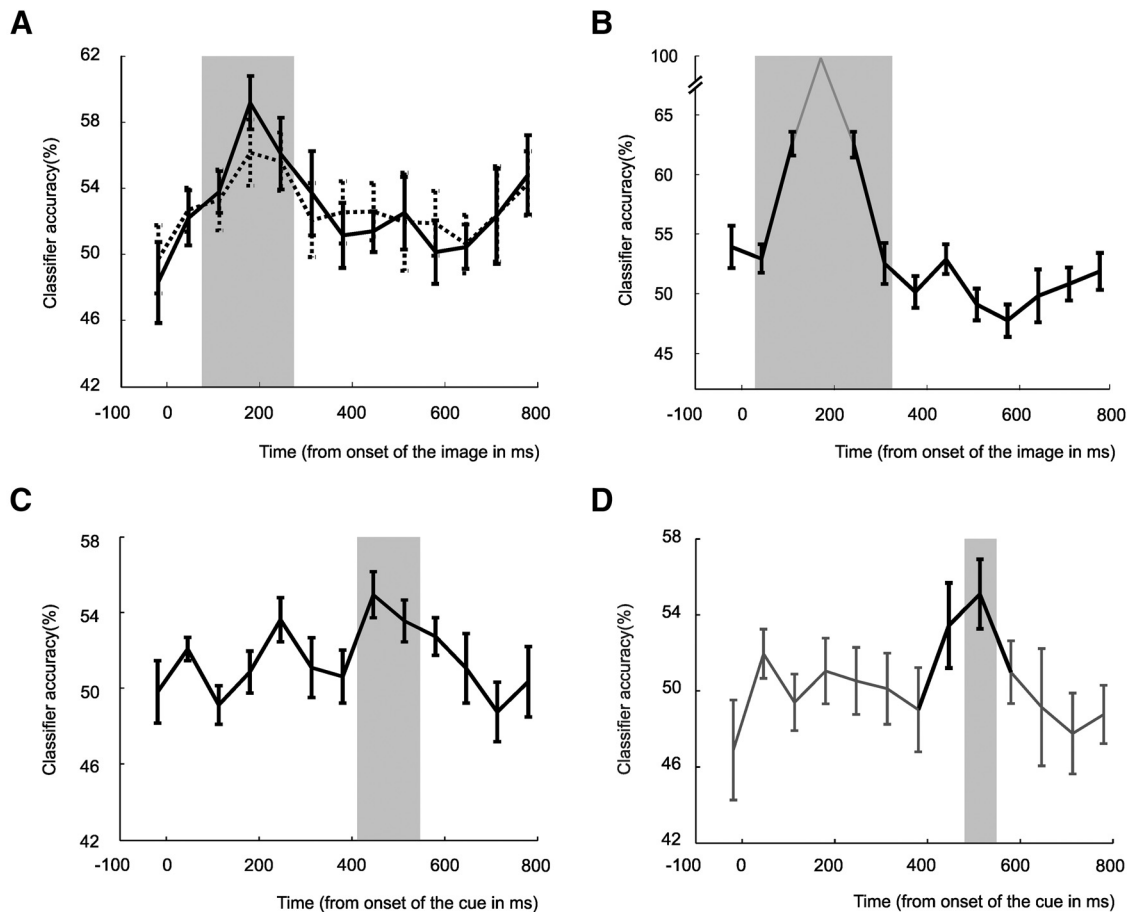


Figure 2. Face- and scene-specific representations during encoding and retrieval. **A**, Cross-validated accuracy of separate (for each time bin) pattern classifiers decoding faces and scenes during encoding (solid line). A rapidly (at 113 to 246 ms and peak at 180 ms) emergent pattern classifier decoded faces and scenes. Dotted lines show the cross-validated accuracy for decoding the signal from which the average category-specific ERF was subtracted (at 118 and 246 ms uncorrected $p = 0.009$ and cluster-level FWE-corrected $p = 0.061$). **B**, The performance of 180 ms classifier in decoding other stages of encoding (significant only at the immediately adjacent time bins). The 0 ms in **A** and **B** correspond to the onset of the images (face or scene) during encoding. **C**, The same 180 ms classifier from the encoding period showed significant replay of associated image information at 446 to 513 ms from onset of the cue, “Old” words, during correct word recognition. **D**, In trials in which the associated image is also correctly identified (recollected), replay is detected at 513 ms after onset of the cue. The 0 ms in **C** and **D** correspond to the word onset during retrieval. The time bins with significant classifications, multiple-comparisons corrected $p < 0.05$, are highlighted in gray (see Materials and Methods and Results sections for details). Error bars indicate SEM. In **B** and **D**, the classification accuracy only from time points depicted with black lines were considered in second-level analyses.

differ in selecting the correct image when they were “Not sure” (for faces: $M = 8.12\%$, $SD = 5.90\%$, and for scenes: $M = 7.48\%$, $SD = 5.00\%$; $t_{(9)} = 0.48$, $p = 0.642$).

MEG-based decoding

MVPCs were used at different time bins during encoding to decode the emergence of category specific neural activity elicited by picture onset. At encoding, cross-validation analysis (Fig. 2A, solid line) revealed that significant above chance classification peaked at 180 ms after onset of the image (averaged classification accuracy = 59.20% at 180 ms; peak-level $t_{(9)} = 5.37$, cluster-level FWE-corrected $p = 0.001$, including 113, 180 [peak] and 246 ms). We next investigated whether the significant classification at 180 ms was driven by the event-related M170. Once we subtracted the mean category-specific ERFs from each individual trial, the subsequent classification analysis still revealed a significant above chance classification at 180 and 246 ms after image onset (average classification accuracy = 56.08% at 180 ms; peak-level $t_{(9)} = 2.90$, cluster-level uncorrected $p = 0.009$, including 180 and 246 [peak] ms; Figure 2A, dotted line). This result suggests that the classification at 180 ms is not primarily driven by any category-specific ERF response. We note, however, that this

classification analysis was marginally significant FWE corrected ($p = 0.061$), perhaps suggesting that the ERF component did contribute, albeit minimally, to classifier performance in our main encoding analysis.

We then tested whether the category-specific oscillatory patterns, which emerged at early time windows and peaked at 180 ms, were replayed at any other time point within the first 800 ms of the encoding period. This was done by training the classifier on the oscillatory pattern at the 180 ms time bin and testing during other encoding time points. The 180 ms pattern was detected only at an early time window during encoding (peak-level $t_{(9)} = 11.74$, cluster-level FWE-corrected $p < 0.001$ including 46, 113 [peak], 246, 313 ms; Figure 2B). Correct classification rapidly dropped before and after the early time cluster. This suggests that face- and scene-related neural representations present at early time bins did not reemerge at later time bins during the encoding period that we analyzed. In sum, we saw a category-specific oscillatory pattern at 180 ms that was not replayed at later time points and was not primarily driven by a category-specific event-related response.

Next, we sought to investigate whether neural patterns identified at 180 ms during encoding were replayed during retrieval,

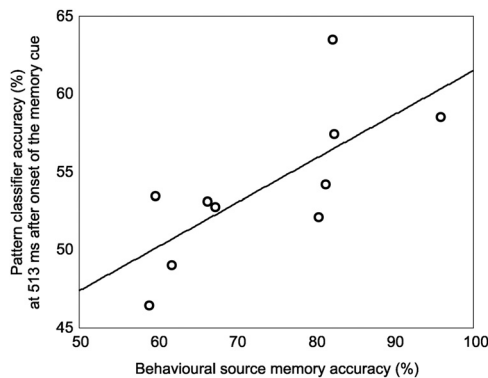


Figure 3. Classification accuracy correlates with source memory accuracy. Classification accuracy in decoding faces and scenes at 513 ms after onset of the memory cue (an “Old” word) correlated positively with behavioral accuracy in source memory ($r = 0.73$ and $p = 0.017$). In this analysis, all trials were considered in which the word was correctly recognized as “Old” (recognition hits). Each circle represents a participant.

when the memory was cued by the associated word. For this analysis, we used all the trials in which participants correctly recognized the word cue (i.e., Hits; averaged number of trials across subjects = 112 and SD = 9). The decoding revealed significant classification of oscillatory patterns elicited by the onset of words to the images associated with that word at 446 to 513 ms (peak-level $t_{(9)} = 3.06$, cluster-level FWE-corrected $p = 0.022$ including 446 [peak] and 513 ms; Figure 2C). We therefore saw the same category-specific 180 ms oscillatory pattern seen at encoding during retrieval at ~ 450 ms after word onset.

Finally, we used these two time windows (446 and 513 ms) for hypothesis-driven testing of classification accuracy only in those trials in which subjects correctly selected the associated image. This was done to identify whether the replay at 446 and/or 513 ms is associated with recollection (averaged number of trials across subjects = 66 and STD = 19). Congruent with this notion, we found significant classification at 513 ms (peak-level $t_{(9)} = 2.64$, Bonferroni corrected $p = 0.026$ for testing two time windows; Fig. 2D) for recollected trials. Furthermore, when all correct word recognition trials (regardless of source memory responses) were considered, the classification accuracy at 513 ms ($r = 0.73$ and $p = 0.017$) was predictive of source accuracy (Fig. 3); however, this was not the case at 446 ms ($r = -0.07$, $p = 0.833$). Therefore, only the time point at which the classification performance was predictive of source memory performance showed significant classification (replay) when recollected trials were used selectively (Fig. 2D). This relationship between replay and source memory performance suggests a link between category-specific replay and the ability to recollect the contextual details of a previous event.

TF analysis

Group-level TF analysis revealed that at 400–550 ms (the time window at which the MVPA indicated memory replay), there was a significant ($p < 0.05$, FWE corrected) theta (3–8 Hz) power increase for hit trials compared with CRs that was maximal in left-temporal channels (Fig. 4A). A similar statistically significant ($p < 0.05$, FWE corrected) theta increase was also apparent in the adjacent time windows 250–400 ms and 550–700 ms. These results are congruent with previous studies contrasting recognition-hits and correct-rejections of word stimuli. Group-level TF analysis revealed that at 400–550 ms (time window at which the MVPA indicated memory replay), there was a significant ($p < 0.05$, FWE corrected) theta

power increase for hit trials compared with CRs that was maximal in left-temporal channels (Fig. 4A). A similar statistically significant ($p < 0.05$, FWE corrected) theta increase was also apparent in the adjacent time windows 250–400 ms and 550–700 ms. These results are congruent with previous studies contrasting recognition-hits and correct-rejections of word stimuli (Düzel et al., 2003, 2005). We also found that beta (23–25 Hz) power decreased for hit compared with CR trials over central and occipital channels at 400–550 ms (Fig. 4B) and the following time window, 550–700 ms.

In a follow-up analysis we tested for power differences between source hits (correct picture selection) and recognition miss trials (misses) in the same three time windows. We found a significant ($p < 0.05$, FWE corrected) power decrease within the beta frequency range (13–25 Hz) at 400–550 ms and at 550–700 ms for source hits compared with recognition miss trials. This difference peaked over central channels, which was similar to findings by Osipova et al. (2006).

Discussion

Our findings indicate that the category-specific neural representations of faces and scenes elicited selectively at early (before 200 ms after stimulus onset) stages of encoding are replayed during recollection. This replay of source information in this hippocampal-dependent task (Horner et al., 2012) occurs relatively rapidly, ~ 500 ms after the onset of the word cue (Fig. 2), and was predictive of behavioral accuracy in the source memory test (Fig. 3). Our results extend functional magnetic resonance imaging (fMRI) studies of single event memories showing that cortical activity patterns elicited during encoding reappear during subsequent memory retrieval (Kahn et al., 2004; Polyn et al., 2005; Johnson et al., 2009; Kuhl et al., 2011; Ritchey et al., 2013; Staresina et al., 2012b). In these studies, the low temporal resolution of fMRI did not permit to determine whether the replayed patterns were established early or late during encoding and at which time bin(s) they were replayed during retrieval.

The time information obtained here addresses two major mechanistic possibilities regarding encoding. First, it shows that encoding is possible in the absence of prolonged maintenance of very early representations during encoding. The possibility that maintenance of information can aid encoding into long-term memory has been recently suggested (for review, see Hasselmo and Stern, 2006). Although our results do not rule out such a possibility, they suggest that, if there is encoding-related maintenance, it does not involve replaying very early cortical representations (Fig. 2B; also see Carlson et al., 2013). Second, given the prolonged nature of encoding processes, the neural representations that are encoded and later replayed could be modified versions of the early cortical representations. According to this possibility, memories are reconstructed during the later stages of encoding and, therefore, the early event representations cannot be reinstated during recollection. Our data ruled out this possibility because they show that early event representations can be reinstated during recollection, akin to representational “snapshots.”

In this study, we have investigated a special (albeit frequently studied) case of recollection in which memory content is composed of associations of single events (Fig. 1). This may limit the generalizability of our findings to mechanisms underlying prolonged events, such as those elicited during continuous spatial navigation (Hoffman and McNaughton, 2002; Fries et al., 2003) or movies (Gelbard-Sagiv et al., 2008). Therefore, although we have positive evidence that very early representations can survive in long-term memory, we cannot exclude the possibility that late representations are also replayed. Furthermore, although our

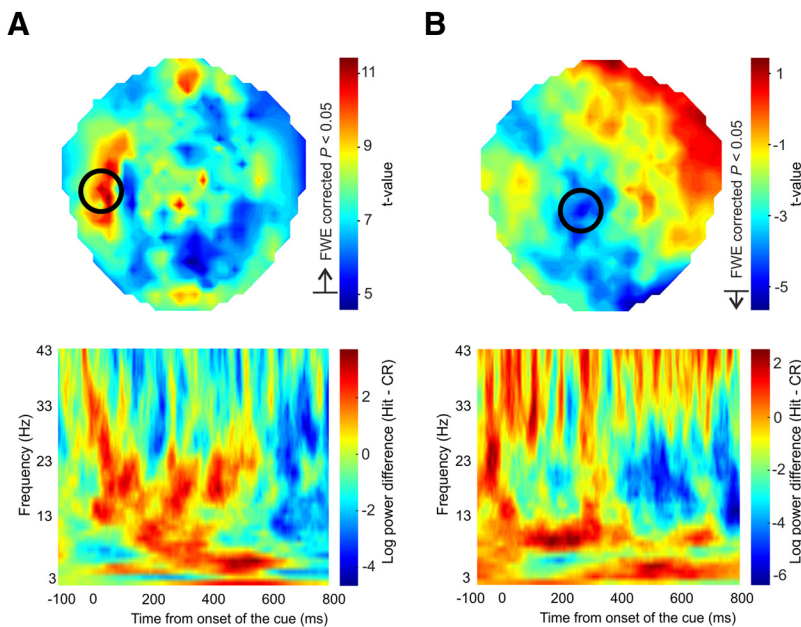


Figure 4. Power differences between recognition hits (Hits) and CRs at 400–550 ms. **A**, Top: Topographic distribution of t -values of the Hits > CRs contrast at 3 Hz, where the difference peaked in the theta (3–8 Hz) band. Bottom: Time-frequency representation of log power differences between Hits and CRs (Hits – CRs) at a left temporal channel where the difference peaked (highlighted with the black circle). **B**, Top: t -value map for the difference in the beta (23–25 Hz) band, peaking at 23 Hz. Bottom: TF representation of the log power differences at a central channel (marked with black circle) where the difference peaked. In **A** and **B**, color bars illustrate the range of t -values in the top plots and the arrows show the FWE-corrected t -value threshold (set at $p < 0.05$). The color bars in other plots show the range of log power differences between Hits and CRs.

results show replay associated with successful retrieval of scene/face associations (an index of recollection), we did not have a sufficient number of trials to assess whether replay is also absent in familiarity; that is, when the items (words) were recognized but scene/face associations could not be recollected.

Recollection is critically dependent on the hippocampal formation (Vargha-Khadem et al., 1997) and we have shown previously that source memory performance in our experimental paradigm is dependent on hippocampal integrity and source memory performance in this task is correlated with hippocampal volume (Horner et al., 2012). This brain structure is capable of pattern completing memory representations of event details in response to a partial memory cue (Norman, 2010; Kumaran and McClelland, 2012). In our study, the partial memory cue was the word that was presented in isolation during retrieval and pattern completion involved retrieving the image paired with it at encoding. To date, the timing of hippocampus-dependent pattern completion has remained unclear because, thus far, the time course of reinstating representations of single events could not be tracked. Our results now show that such a reinstatement is quite rapid and occurs within 446 to 513 ms after the onset of a memory cue (Figs. 2C). Given that such associative retrieval of scene information is hippocampus dependent (Yonelinas et al., 2002), this finding indicates that hippocampus-dependent pattern completion processes in response to partial memory cues are more rapid than many previous ERP/ERF studies of recollection have implied (Düzel et al., 2001; Addante et al., 2012). In these studies, recollection was associated with ERP/ERF components emerging between 500 and 800 ms.

Our data now indicate that within ~500 ms, a hippocampal pattern completion process and the ensuing cortical reinstatement must have been completed. Such a rapid timing is consistent with more recent electromagnetic data from patients with

bilateral hippocampal lesions showing that memory cues such as those used here can initiate hippocampus-dependent retrieval of contextual information within 350 ms (Horner et al., 2012). It is also consistent with recent intracortical recordings in humans showing associative recognition effects at ~400 ms in the perirhinal cortex, just after an earlier hippocampus response at 250 ms (Staresina et al., 2012a). It should be noted that under circumstances in which the retrieval of visual associations may not rely on recollection and is possibly unconscious, reactivation can be observed even earlier (Waldhauser et al., 2012; Wimber et al., 2012). In these two recent studies, memory representations of simple visual associations (color or frequency) appeared to be reactivated by visual cues at ~100–300 ms. To the extent that these reactivations tap into hippocampus-dependent memory processes, these studies would raise the possibility that hippocampus-dependent reactivation of simple visual associations may occur earlier than we have observed here for our more complex scene/face stimuli.

There are two caveats to consider with regard to our conclusions. Our MVPA is based on neural oscillations that are most likely to be largely cortical in origin, and our MVPA analyses therefore likely detect the reactivation of a cortical pattern rather than the retrieval mechanism that would necessarily precede (or trigger) that reactivation (e.g., pattern completion in the hippocampus). As long as these retrieval mechanisms that trigger memory reactivation are inaccessible in relation to MVPA analyses, there remains uncertainty as to whether the reactivation that we observe is a direct consequence of retrieval processing (which we could also refer to as “ecphory”; Tulving et al., 1983) or results from additional postretrieval processing and includes mental imagery. Furthermore, our interpretations of our MVPA findings have focused on hippocampal mechanisms because we have thoroughly established a tight hippocampal dependence of our task in a previous study (Horner et al., 2012). However, this link remains necessarily indirect because we cannot conclusively determine at which stage or time the hippocampus may have been involved.

In summary, our results suggest that hippocampus-dependent pattern completion processes can lead to a reinstatement of the early neural representations of experienced events akin to a visual “snapshot.” Therefore, the memory engram (Dudai, 2012; Liu et al., 2012) stored in the hippocampus must be sufficiently precise to enable the conservation of cortical event representations formed during very early stages of encoding. Encoding processes, despite their prolonged nature, appear capable of faithfully conserving initial representations of events without actively maintaining them in their early representation pattern. We believe that the method of decoding neural representations at encoding and retrieval with high temporal resolution and determining which representations are conserved and subsequently replayed can provide a new approach for future investigations of these mechanisms.

References

- Addante RJ, Ranganath C, Olichney J, Yonelinas AP (2012) Neurophysiological evidence for a recollection impairment in amnesia patients that leaves familiarity intact. *Neuropsychologia* 50:3004–3014. [CrossRef Medline](#)
- Axmacher N, Lenz S, Haupt S, Elger CE, Fell J (2010) Electrophysiological signature of working and long-term memory interaction in the human hippocampus. *Eur J Neurosci* 31:177–188. [CrossRef Medline](#)
- Carlson T, Tovar DA, Alink A, Kriegeskorte N (2013) Representational dynamics of object vision: The first 1000 ms. *J Vis* 13 Available at: <http://www.journalofvision.org/content/13/10/1>. Accessed September 9, 2013.
- Dudai Y (2012) The restless engram: consolidations never end. *Annu Rev Neurosci* 35:227–247. [CrossRef Medline](#)
- Düzel E, Vargha-Khadem F, Heinze HJ, Mishkin M (2001) Brain activity evidence for recognition without recollection after early hippocampal damage. *Proc Natl Acad Sci U S A* 98:8101–8106. [CrossRef Medline](#)
- Düzel E, Habib R, Schott B, Schoenfeld A, Lobaugh N, McIntosh AR, Scholz M, Heinze HJ (2003) A multivariate, spatiotemporal analysis of electromagnetic time-frequency data of recognition memory. *Neuroimage* 18:185–197. [CrossRef Medline](#)
- Düzel E, Neufang M, Heinze HJ (2005) The oscillatory dynamics of recognition memory and its relationship to event-related responses. *Cereb Cortex* 15:1992–2002. [CrossRef Medline](#)
- Düzel E, Penny WD, Burgess N (2010) Brain oscillations and memory. *Curr Opin Neurobiol* 20:143–149. [CrossRef Medline](#)
- Fell J, Axmacher N (2011) The role of phase synchronization in memory processes. *Nat Rev Neurosci* 12:105–118. [CrossRef Medline](#)
- Fernández G, Efferen A, Grunwald T, Pezer N, Lehnertz K, Dümpelmann M, Van Roost D, Elger CE (1999) Real-time tracking of memory formation in the human rhinal cortex and hippocampus. *Science* 285:1582–1585. [CrossRef Medline](#)
- Fisch L, Privman E, Ramot M, Harel M, Nir Y, Kipervasser S, Andelman F, Neufeld MY, Kramer U, Fried I, Malach R (2009) Neural “ignition”: enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron* 64:562–574. [CrossRef Medline](#)
- Fries P, Fernández G, Jensen O (2003) When neurons form memories. *Trends Neurosci* 26:123–124. [CrossRef Medline](#)
- Gao Z, Goldstein A, Harpaz Y, Hansel M, Zion-Golombic E, Bentin S (2013) A magnetoencephalographic study of face processing: M170, gamma-band oscillations and source localization. *Hum Brain Mapp* 34:1783–1795. [CrossRef Medline](#)
- Gelbard-Sagiv H, Mukamel R, Harel M, Malach R, Fried I (2008) Internally generated reactivation of single neurons in human hippocampus during free recall. *Science* 322:96–101. [CrossRef Medline](#)
- Hanslmayr S, Staudigl T, Fellner MC (2012) Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Front Hum Neurosci* 6:74. [CrossRef Medline](#)
- Hasselmo ME, Stern CE (2006) Mechanisms underlying working memory for novel information. *Trends Cogn Sci* 10:487–493. [CrossRef Medline](#)
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430. [CrossRef Medline](#)
- Hoffman KL, McNaughton BL (2002) Coordinated reactivation of distributed memory traces in primate neocortex. *Science* 297:2070–2073. [CrossRef Medline](#)
- Horner AJ, Gadian DG, Fuentemilla L, Jentschke S, Vargha-Khadem F, Düzel E (2012) A rapid, hippocampus-dependent, item-memory signal that initiates context memory in humans. *Curr Biol* 22:2369–2374. [CrossRef Medline](#)
- Jafarpour A, Barnes G, Fuentemilla L, Düzel E, Penny WD (2013a) Population Level Inference for Multivariate MEG Analysis. *PLoS One* 8:e71305. [CrossRef Medline](#)
- Jafarpour A, Horner AJ, Fuentemilla L, Penny WD, Düzel E (2013b) Decoding oscillatory representations and mechanisms in memory. *Neuropsychologia* 51:772–780. [CrossRef Medline](#)
- Johnson JD, McDuff SG, Rugg MD, Norman KA (2009) Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. *Neuron* 63:697–708. [CrossRef Medline](#)
- Kahn I, Davachi L, Wagner AD (2004) Functional-neuroanatomic correlates of recollection: implications for models of recognition memory. *J Neurosci* 24:4172–4180. [CrossRef Medline](#)
- Kilner JM, Kiebel SJ, Friston KJ (2005) Applications of random field theory to electrophysiology. *Neurosci Lett* 374:174–178. [CrossRef Medline](#)
- Kuhl BA, Rissman J, Chun MM, Wagner AD (2011) Fidelity of neural reactivation reveals competition between memories. *Proc Natl Acad Sci U S A* 108:5903–5908. [CrossRef Medline](#)
- Kumaran D, McClelland JL (2012) Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychol Rev* 119:573–616. [CrossRef Medline](#)
- Lega BC, Jacobs J, Kahana M (2012) Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus* 22:748–761. [CrossRef Medline](#)
- Litvak V, Mattout J, Kiebel S, Phillips C, Henson R, Kilner J, Barnes G, Oostenveld R, Daunizeau J, Flandin G, Penny W, Friston K (2011) EEG and MEG data analysis in SPM8. *Comput Intell Neurosci* 2011:852961. [CrossRef Medline](#)
- Liu J, Harris A, Kanwisher N (2002) Stages of processing in face perception: an MEG study. *Nat Neurosci* 5:910–916. [CrossRef Medline](#)
- Liu X, Ramirez S, Pang PT, Puryear CB, Govindarajan A, Deisseroth K, Tonegawa S (2012) Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature* 484:381–385. [CrossRef Medline](#)
- Marr D (1971) Simple memory: a theory for archicortex. *Philos Trans R Soc Lond B Biol Sci* 262:23–81. [CrossRef Medline](#)
- McCarthy G, Puce A, Belger A, Allison T (1999) Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cereb Cortex* 9:431–444. [CrossRef Medline](#)
- Norman KA (2010) How hippocampus and cortex contribute to recognition memory: revisiting the complementary learning systems model. *Hippocampus* 20:1217–1227. [CrossRef Medline](#)
- Osipova D, Takashima A, Oostenveld R, Fernández G, Maris E, Jensen O (2006) Theta and gamma oscillations predict encoding and retrieval of declarative memory. *J Neurosci* 26:7523–7531. [CrossRef Medline](#)
- Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes retrieval during memory search. *Science* 310:1963–1966. [CrossRef Medline](#)
- Puce A, Allison T, McCarthy G (1999) Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cereb Cortex* 9:445–458. [CrossRef Medline](#)
- Ritchey M, Wing EA, Labar KS, Cabeza R (2013) Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cereb Cortex* 23:2818–2828. [CrossRef Medline](#)
- Rossion B, Caharel S (2011) ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception. *Vision Res* 51:1297–1311. [CrossRef Medline](#)
- Staresina BP, Fell J, Do Lam AT, Axmacher N, Henson RN (2012a) Memory signals are temporally dissociated in and across human hippocampus and perirhinal cortex. *Nat Neurosci* 15:1167–1173. [CrossRef Medline](#)
- Staresina BP, Henson RN, Kriegeskorte N, Alink A (2012b) Episodic reinstatement in the medial temporal lobe. *J Neurosci* 32:18150–18156. [CrossRef Medline](#)
- Tulving E (1985) How many memory systems are there? *American Psychologist* 40:385–398. [CrossRef](#)
- Tulving E, Voi MEL, Routh DA, Loftus E (1983) Ecphoric processes in episodic memory [and Discussion]. *Philos Trans R Soc Lond B* 302:361–371. [CrossRef](#)
- Vapnik V (2000) *The nature of statistical learning theory*. New York: Springer.
- Vargha-Khadem F, Gadian DG, Watkins KE, Connelly A, Van Paesschen W, Mishkin M (1997) Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* 277:376–380. [CrossRef Medline](#)
- Waldhauser GT, Johansson M, Hanslmayr S (2012) Alpha/Beta oscillations indicate inhibition of interfering visual memories. *J Neurosci* 32:1953–1961. [CrossRef Medline](#)
- Wimber M, Maaß A, Staudigl T, Richardson-Klavehn A, Hanslmayr S (2012) Rapid memory reactivation revealed by oscillatory entrainment. *Curr Biol* 22:1482–1486. [CrossRef Medline](#)
- Yonelinas AP, Kroll NE, Quamme JR, Lazzara MM, Sauvé MJ, Widaman KF, Knight RT (2002) Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. *Nat Neurosci* 5:1236–1241. [CrossRef Medline](#)