

Cued Memory Retrieval Exhibits Reinstatement of High Gamma Power on a Faster Timescale in the Left Temporal Lobe and Prefrontal Cortex

Robert B. Yaffe,¹ Ammar Shaikhouni,² Jennifer Arai,¹ Sara K. Inati,³ and Kareem A. Zaghloul¹

¹Surgical Neurology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892, ²Department of Neurological Surgery, Ohio State University Wexner Medical Center, Columbus, Ohio 43210, and ³Office of the Clinical Director, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892

Converging evidence suggests that reinstatement of neural activity underlies our ability to successfully retrieve memories. However, the temporal dynamics of reinstatement in the human cortex remain poorly understood. One possibility is that neural activity during memory retrieval, like replay of spiking neurons in the hippocampus, occurs at a faster timescale than during encoding. We tested this hypothesis in 34 participants who performed a verbal episodic memory task while we recorded high gamma (62–100 Hz) activity from subdural electrodes implanted for seizure monitoring. We show that reinstatement of distributed patterns of high gamma activity occurs faster than during encoding. Using a time-warping algorithm, we quantify the timescale of the reinstatement and identify brain regions that show significant timescale differences between encoding and retrieval. Our data suggest that temporally compressed reinstatement of cortical activity is a feature of cued memory retrieval.

Key words: compression; iEEG; memory; reinstatement; retrieval; timescale

Significance Statement

We show that cued memory retrieval reinstates neural activity on a faster timescale than was present during encoding. Our data therefore provide a link between reinstatement of neural activity in the cortex and spontaneous replay of cortical and hippocampal spiking activity, which also exhibits temporal compression, and suggest that temporal compression may be a universal feature of memory retrieval.

Introduction

Successfully remembering an event involves mentally reexperiencing both the details of what happened and the context within which the original event occurred (Tulving, 1972). During recall of a memory, distributed patterns of neural activity are similar to those observed when the memory was initially formed, thus allowing the neural representation of the previous experience to be recovered (Manning et al., 2011; Yaffe et al., 2014). This memory-related reinstatement of activity has been demonstrated in

both oscillatory and neuroimaging changes in the human cortex (Johnson and Rugg, 2007; Johnson et al., 2009; Danker and Anderson, 2010; Manning et al., 2011; Staresina et al., 2012; Yaffe et al., 2014) and in single-unit activity in the human hippocampus (Howard et al., 2012; Miller et al., 2013). Indeed, hippocampal activity may serve to bind together the distributed patterns of cortical activity specific to a given memory to facilitate the reinstatement of those representations (Sederberg et al., 2007a; Kahn et al., 2008; Preston and Eichenbaum, 2013).

Despite converging evidence linking human cortical reinstatement to successful memory retrieval, the temporal dynamics of this process have not been well investigated. One possibility is that, although the patterns of activity are similar, the timescale is faster during retrieval than encoding. This would allow for the reinstated activity occurring during recall to play out more quickly than the associated activity during encoding, which must necessarily occur in real time. This hypothesis closely parallels the concept of replay in hippocampal neurons, in which task-related firing patterns reoccur spontaneously, although at a faster timescale (Diba and Buzsáki, 2007; Karlsson and Frank, 2009; Carr et al., 2011). Both forward and

Received Dec. 13, 2016; revised Feb. 27, 2017; accepted March 2, 2017.

Author contributions: R.B.Y. and K.A.Z. designed research; R.B.Y., S.K.I., and K.A.Z. performed research; R.B.Y., A.S., and J.A. analyzed data; R.B.Y. and K.A.Z. wrote the paper.

This work was supported by the intramural research program of the National Institute for Neurological Disorders and Stroke. We thank all patients who have selflessly volunteered their time to participate in this study.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Kareem A. Zaghloul, Surgical Neurology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health Building 10, Room 3D20, 10 Center Drive, Bethesda, MD 20892-1414. E-mail: kareem.zaghloul@nih.gov.

DOI:10.1523/JNEUROSCI.3810-16.2017

Copyright © 2017 the authors 0270-6474/17/374472-09\$15.00/0

reverse replay of sequences of single-unit spiking activity have been observed in rodent hippocampal place cells during navigation (Lee and Wilson, 2002; Diba and Buzsáki, 2007; Karlsson and Frank, 2009; Gupta et al., 2010), and spontaneous yet faster replay of task-related activity has also been reported in single units in rodent (Euston et al., 2007) and primate cortex (Hoffman and McNaughton, 2002). Although it is unknown whether spontaneous replay is causally related to memory retrieval, replay that occurs when an animal is at rest and planning future trajectories (Karlsson and Frank, 2009) suggests a link between spontaneous replay and memory-related reinstatement.

Here, we leverage the temporal precision afforded by intracranial EEG (iEEG) to investigate the timescale of reinstatement of human cortical activity during cued retrieval in a verbal paired associates task. We have previously shown that cued retrieval is associated with reinstatement of distributed patterns of spectral power present during encoding (Yaffe et al., 2014). We extend this work here by specifically asking whether reinstatement of cortical activity during retrieval is faster than the activity present during encoding, a principal feature of spontaneous replay. During encoding, we examine the progression of high gamma (62–100 Hz) power, a surrogate marker for cortical activation (Manning et al., 2009; Burke et al., 2015), and demonstrate that a similar progression occurs on a faster timescale during retrieval. In addition, we identify the anatomic regions where we observe similar activity between encoding and retrieval, and quantify the difference in the timescales of activity. Our findings suggest that temporal compression of neural activity in the human cortex may also be a feature of cued memory retrieval.

Materials and Methods

Participants. Thirty-four participants (18 males; age 34.6 ± 12.9 years, mean \pm SD) with medication-resistant epilepsy underwent a surgical procedure in which platinum recording contacts were implanted subdurally on the cortical surface as well as deep within the brain parenchyma. For each participant, the clinical team determined the placement of the contacts so as to best localize epileptogenic regions. Of the 34 participants, 15 received electrodes bilaterally on both hemispheres, 10 received electrodes unilaterally on the right, and 9 received electrodes unilaterally on the left. Twenty-three participants had a seizure focus that was successfully localized and surgically resected. Of these participants, the seizure focus was located in the left temporal lobe (9 participants), the right temporal lobe (9 participants), the left frontal lobe (1 participant), the right frontal lobe (3 participants), or the right parietal lobe (1 participant). Eleven participants did not have a localizable seizure focus. The full-scale IQ of the participants was 102.6 ± 15.8 . Thirty participants were right-handed, 2 were left-handed, and 2 were ambidextrous. The Institutional Review Board approved the research protocol, and informed consent was obtained from the participants and their guardians. These data were initially collected and analyzed for previous publications (Yaffe et al., 2014; Greenberg et al., 2015; Haque et al., 2015).

Paired associates task. Each patient participated in a paired associates task (see Fig. 1A). Lists were composed of four pairs of common nouns, chosen at random and without replacement from a pool of high-frequency nouns. Words were presented sequentially and appeared in capital letters at the center of the screen. Study word pairs were separated from their corresponding recall cue by a minimum lag of two study or test items. During the study period (encoding), each word pair was preceded by an orientation stimulus (a row of capital X's) that appeared on the screen for 300 ms followed by a blank interstimulus interval (ISI) of 750 ms with a jitter of 75 ms. Word pairs were then presented on the screen for 2500 ms followed by a blank ISI of 1500 ms with a jitter of 75 ms. For each list, four pairs are studied followed by 20 s of a math distractor task in which math problems are presented of the form $A + B + C$, where $A/B/C$ are random integers between 1 and 9. During the test period (retrieval), one randomly chosen word from each study pair was shown, and the participant was asked to recall the other word

from the pair by vocalizing a response into a microphone. Each cue word was preceded by an orientation stimulus (a row of question marks) that appeared on the screen for 300 ms followed by a blank ISI of 750 ms with a 75 ms jitter. Cue words were then presented on the screen for 3000 ms followed by a blank ISI of 4500 ms. Participants could vocalize their response any time during the recall period after cue presentation. Vocalizations were digitally recorded and then manually scored for analysis. Responses were designated as correct, as intrusions, or as passes when no vocalization was made or when the participant vocalized the word “pass.” Intrusion and pass trials were designated as incorrect trials. A single experimental session contained up to 25 lists. For analysis, we designated a single trial as the encoding period for a study word pair and the retrieval period during testing of its corresponding cue.

iEEG recordings. Depending on the amplifier and the discretion of the clinical team, iEEG signals were sampled at 512, 1000, or 2000 Hz. Signals were referenced to a common contact placed subcutaneously, on the scalp, or on the mastoid process. All recorded traces were resampled at 1000 Hz, and a fourth-order 2 Hz stopband Butterworth notch filter was applied at 60 Hz to eliminate electrical line noise. The testing laptop sent ± 5 V analog pulses via an optical isolator into a pair of open lines on the clinical recording system to synchronize the electrophysiological recordings with behavioral events.

We collected electrophysiological data from a total of 2795 subdural and depth recording contacts (PMT, AdTech). Subdural contacts were arranged in both grid and strip configurations with an intercontact spacing of 10 mm. Contact localization was accomplished by coregistering the postoperative CTs with the postoperative MRIs using both FSL Brain Extraction Tool and FLIRT software packages and mapped to both MNI and Talairach space using an indirect stereotactic technique and OsiriX Imaging Software DICOM viewer package. The resulting contact locations were subsequently projected to the cortical surface of a MNI N27 standard brain (Dykstra et al., 2012). Preoperative MRIs were used when postoperative MR images were not available.

We analyzed iEEG data using bipolar referencing to reduce volume conduction and confounding interactions between adjacent electrodes (Nunez and Srinivasan, 2006). We defined the bipolar montage in our dataset based on the geometry of iEEG electrode arrangements. For every grid, strip, and depth probe, we isolated all pairs of contacts that were positioned immediately adjacent to one another; bipolar signals were then found by differencing the signals between each pair of immediately adjacent contacts. The resulting bipolar signals were treated as new virtual electrodes (referred to as electrodes throughout the text), originating from the midpoint between each contact pair. All subsequent analyses were performed using these derived bipolar signals. In total, our dataset consisted of 2959 bipolar electrodes (1314 left hemispheric, 1645 right hemispheric). The locations of all electrodes in all participants are shown on a standardized brain in Figure 1C. Figure 1D shows a smoothed representation of the number of participants that have electrodes in each spatial location.

Data analyses and spectral power. To quantify changes in spectral power during encoding and retrieval, we convolved the bipolar iEEG signals with complex valued Morlet wavelets (wavelet number 6) to obtain magnitude and phase information (Addison, 2002). We squared and log-transformed the magnitude of the continuous-time wavelet transform to generate a continuous measure of instantaneous power. We used 50 logarithmically spaced wavelets ranging from 2 to 100 Hz. During the encoding and retrieval periods, we convolved each wavelet with 5000 ms of iEEG data surrounding the presentation of word pairs or cue words, from 1000 ms before stimulus presentation to 4000 ms after stimulus presentation. We included an additional 2000 ms buffer on either end of the data to account for any edge effects and which was not subsequently analyzed.

We binned the continuous time transforms into 500 ms epochs spaced every 100 ms (80% overlap) and averaged the instantaneous power over each epoch. We labeled each 500 ms epoch with the time that corresponds to the end of the temporal epoch. We z -transformed power values separately for each frequency and for each session using the mean and SD of all 500 ms epochs for that session. For each temporal epoch, we sub-

sequently averaged the z -transformed power across the high gamma (62–100 Hz) frequency band.

Metrics of reinstatement and asymmetry. For every temporal epoch in each encoding and retrieval period, we constructed a feature vector comprised of the average z -scored power for every electrode in each participant. For each encoding temporal epoch, i , and for each retrieval temporal epoch, j , we define feature vectors as follows:

$$\vec{E}_i = [z_1(i) \dots z_L(i)]$$

$$\vec{R}_j = [z_1(j) \dots z_L(j)]$$

where $z_l(i)$ is the z -transformed high gamma power of electrode $l = 1 \dots L$ in temporal epoch i .

To quantify reinstatement on trial n , we calculated the cosine similarity between all encoding and retrieval feature vectors \vec{E}_i and \vec{R}_j for all pairs of encoding and retrieval temporal epochs during that trial (see Fig. 1B). Thus, for each trial, n , we generate a temporal map of reinstatement values by calculating the cosine similarity between each pair of feature vectors as follows:

$$\rho_n(i, j) = \frac{\vec{E}_i \cdot \vec{R}_j}{\|\vec{E}_i\| \|\vec{R}_j\|}$$

where $\rho_n(i, j)$ corresponds to the reinstatement of neural activity in the high gamma frequency band across all electrodes between encoding epoch i and retrieval epoch j during trial n . We average the temporal maps of reinstatement across all correct and incorrect trials separately for each participant.

To determine whether correct trials exhibited significantly greater reinstatement than incorrect trials, we used a nonparametric clustering procedure. We computed the average reinstatement map for each participant; and then, for each encoding–retrieval time pair, we computed a p value and a t statistic using a paired t test comparing the distributions of averaged reinstatement values across participants between trial types. We then permuted the trial labels of the average reinstatement maps 1000 times, and for each permutation, repeated the t test between trial types for each encoding–retrieval pair. In the true case and each of the permuted cases, we identified contiguous clusters of encoding–retrieval time pairs in which the p value was < 0.05 . For each cluster, we computed a cluster statistic by taking the sum of the t statistic across all encoding–retrieval time pairs in that cluster. In this manner, large clusters statistics can arise either from large differences in reinstatement between trial types that extend over a short duration or from weaker differences that persist over longer time periods. In each permutation ($n = 1000$), we saved the largest observed cluster statistic, thus generating an empiric distribution of maximum cluster statistics that would arise by chance. To generate a p value for each cluster observed in the true dataset, we compared the position of the true cluster statistic to the distribution of maximum cluster statistics from the permuted cases. Clusters were determined to be significant if their p value calculated in this manner was < 0.05 .

To calculate the reinstatement associated with every individual feature (individual electrode), k , we followed a similar procedure. In this case, we constructed a vector comprised of the z -scored power for that particular feature across all trials, $n = 1 \dots N$. Thus, for every encoding temporal epoch, i , and every retrieval temporal epoch, j as follows:

$$\vec{E}_i = [z_1(i) \dots z_N(i)]$$

$$\vec{R}_j = [z_1(j) \dots z_N(j)]$$

where $z_n(i)$ is the z -transformed power of feature k for trial $n = 1 \dots N$ in temporal epoch i . For N trials, we thus create a vector containing N values for each feature for each time point. We then calculate reinstatement of that particular feature by calculating the cosine similarity between all encoding and retrieval vectors \vec{E}_i and \vec{R}_j for all pairs of encoding and retrieval temporal epochs. Thus, for every feature, k , we generate a reinstatement map, $C_k(i, j)$, corresponding to the reinstatement of that individual feature between encoding epoch i and retrieval epoch j across

N trials. For every participant, we compute the reinstatement maps for each feature for all correct trials.

We computed an asymmetry index for each average temporal map of reinstatement (see Fig. 1B) (Oemisch et al., 2015) using the first 2.5 s following the presentation of the word pairs during encoding and the first 2.5 s following the cue during retrieval. This corresponds to the time period during which words are visually presented on the screen during encoding. Because words are presented for a longer time during retrieval, this is the maximum time period that ensures that any visual responses evoked by word presentation are matched between encoding and retrieval. We define the asymmetry index, AI , as follows:

$$AI = a - b$$

where a and b are the average value of reinstatement during the encoding and retrieval time points in regions above and below the diagonal line, $t_{enc} = t_{ret}$, respectively. Each point along the diagonal line indicates the reinstatement between temporal epochs occurring at identical times during encoding and retrieval following stimulus presentation. Regions above the diagonal line correspond to temporal epochs following the presentation of the cue during retrieval that are paired with temporal epochs occurring later during encoding. Conversely, regions below the diagonal line correspond to retrieval epochs that occur later in time compared with corresponding encoding epochs. Hence, AI reflects the extent to which reinstatement is temporally asymmetric. In this manner, we can calculate AI using reinstatement maps of the distributed pattern of spectral activity across all electrodes together, and separately for each feature.

Time scaling. To quantify the relative timing of encoding and retrieval, we applied a time scaling algorithm to the continuous time series of spectral power for each feature (see Fig. 3B). For each feature, we defined $f(t)$ and $g(t)$ as the continuous time z -scored spectral power between 0.5 and 2.5 s for encoding and retrieval, respectively, averaged across all correct trials and time-locked to the stimulus presentation. We defined a time scaling factor, σ , between -0.5 and 0.5 that yielded the largest Pearson correlation between temporally stretched versions of $f(t)$ and $g(t)$ (see Fig. 3C). By construction, if $\sigma > 0$, we compared the encoding trace, $f(t)$, with a temporally stretched version of the retrieval trace, $g((1 - |\sigma|)t)$. Conversely, if $\sigma < 0$, we compared a temporally stretched version of the encoding trace, $f((1 - |\sigma|)t)$, to the retrieval trace, $g(t)$. The time scaling factor, σ , is therefore symmetric around 0 and informs us as to whether the time course of spectral power is best correlated when temporally stretching either encoding or retrieval.

We identified the optimal time scaling factor, σ , that yielded the corresponding maximum correlation between encoding and retrieval for every electrode separately. Every electrode will have some maximum correlation between encoding and retrieval, even if the continuous time traces are not similar. We therefore generated a distribution of maximum correlations using all electrodes from all participants. We identified the median maximum correlation across all electrodes that results after scaling, which was 0.78. We restricted our subsequent analysis to only those electrodes with a maximum correlation exceeding this threshold. Although some participants had cognitive difficulties due to medication effects or other neurological factors that may have affected their processing speed and response times, those difficulties are present during both encoding and retrieval. Comparing the relative timescales of encoding and retrieval within participants is therefore not confounded by different processing speeds across participants.

Anatomic localization and visualization. To visualize the spatial distribution of similarity, asymmetry indices, and time scaling factors for all electrodes, we created 1441 regions of interest (ROIs) evenly spaced throughout an MNI N27 standard brain. In each participant, we identified all electrodes located within 12.5 mm of each ROI and averaged their values (ρ , AI , or σ) to assign a value to that ROI. We then averaged values across participants within an ROI. For all statistical tests used to identify spatial regions exhibiting significant effects, only ROIs that had electrodes from at least 5 participants were included (Sederberg et al., 2007a, b; Burke et al., 2014b). All other ROIs were excluded from these analyses. For each ROI, the number of participants contributing electrodes to that

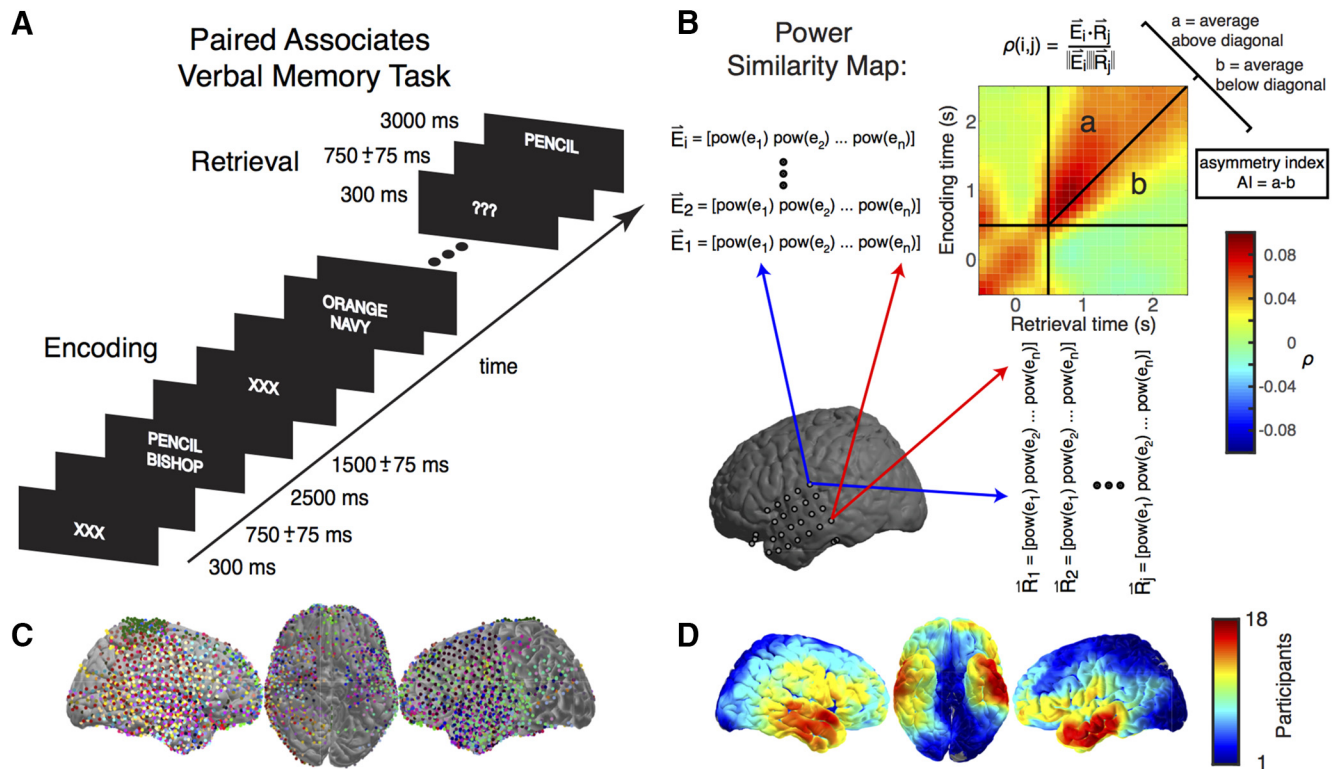


Figure 1. Paired associates task data collection and analysis methods. **A**, Paired associates verbal memory task. Graphic represents timing of task. **B**, Similarity map is constructed by computing the cosine similarity between every brain state vector during encoding and every brain state vector during retrieval. The color of each pixel represents the cosine similarity. Each time point indicates the end of a 500 ms time bin. Consecutive time bins have 80% overlap. Horizontal black line indicates the first-time bin where all the data are from after pair presentation during encoding. Vertical black line indicates the first-time bin where all the data are from after cue presentation during retrieval. Diagonal black line is the $t_{enc} = t_{ret}$ line. **C**, Electrode locations for all participants. Each participant's electrodes are plotted using a distinct color. **D**, Electrode coverage across participants. Color represents number of participants that had electrodes in that region.

ROI is indicated in Figure 1D. We generated cortical topographic plots by assigning each vertex in the 3D rendered image of the standard brain a weighted average of the mean value of each ROI that includes that vertex. Weighted values for each vertex within a single ROI were assigned by convolving the value of the ROI with a 3D Gaussian kernel (radius = 12.5 mm; $\sigma = 4.17$ mm) with center weight 1. The value of each vertex on the standard brain was therefore computed by taking the average of the weighted contributions of all overlapping ROIs. We projected these vertex values onto the standard brain using information from the WFU PickAtlas toolbox (Maldjian et al., 2003). Intensity varied as a function of the value of each ROI and with the SD of the Gaussian kernel, which was used purely as a visualization technique.

To determine whether any anatomic region exhibited similarity (ρ), an AI , or σ , that was significantly different from zero across participants, we used a nonparametric spatial clustering procedure (Maris and Oostenveld, 2007). This procedure identifies contiguous ROIs where the distribution of ρ , AI , or σ across participants significantly deviates from zero while controlling for the family-wise error rate. Briefly, for each ROI, we calculated the true mean ρ , AI , or σ across participants. We then generated 1000 permuted values for each ROI. Because ρ , AI , and σ are symmetric around zero, for each permutation, we randomly multiplied the value for each participant by 1 or -1 , and calculated the mean across participants. We then determined a z score for each true value and each permuted value in each ROI by comparing that value to the distribution of permuted values. For the true data and for each permutation, we identified contiguous spatial clusters of ROIs exhibiting z scores with a magnitude >1.96 (corresponding to a two-tailed p value <0.05). For each cluster, we computed the cluster statistic as the sum of all z -scores in that cluster. In this manner, large magnitude cluster statistics can arise from large deviations in the distributions of ρ , AI , or σ across participants extending over a small spatial region, or moderate deviations that extend over larger regions. We then calculated the exact two-tailed p value for each cluster observed in the true dataset by comparing its cluster statistic to the

distribution of largest cluster statistics drawn from each permutation. Clusters were determined to be significant and corrected for multiple comparisons if their p value calculated in this manner was <0.05 .

Results

Thirty-four participants (18 males; age 34.6 ± 12.9 years, mean \pm SD) with medically refractory epilepsy who underwent surgery for placement of intracranial electrodes for seizure monitoring participated in a verbal paired associates task (Fig. 1A, C, D; see Materials and Methods). Participants studied 224 ± 125 word pairs and successfully recalled $38.0 \pm 20.4\%$ words with a mean response time of 1874 ± 529 ms. On $14.8 \pm 11.2\%$ of trials, participants responded with an incorrect word (intrusions) with a response time of 2773 ± 659 ms. For the remaining $47.2 \pm 18.0\%$ of trials, participants either made no response to the cue word or vocalized the word “pass” with a mean response time of 3374 ± 894 ms.

We were interested in understanding how the time course of neural activity, distributed across all electrode locations, was reinstated between encoding and retrieval. To this end, we first calculated the continuous time changes in spectral power in each electrode in each participant. We focused our analysis on the high gamma (62–100 Hz) frequency band that reflects general cortical activation (Manning et al., 2009; Burke et al., 2015) and that has been implicated in successful memory formation (Yaffe et al., 2014; Burke et al., 2015; Hanslmayr et al., 2016). We used these continuous time traces, locked to the time of word presentation, to generate a temporal map of reinstatement for each trial (Fig. 1B; see Materials and Methods) (Yaffe et al., 2014). Briefly, we constructed a feature vector containing spectral power informa-

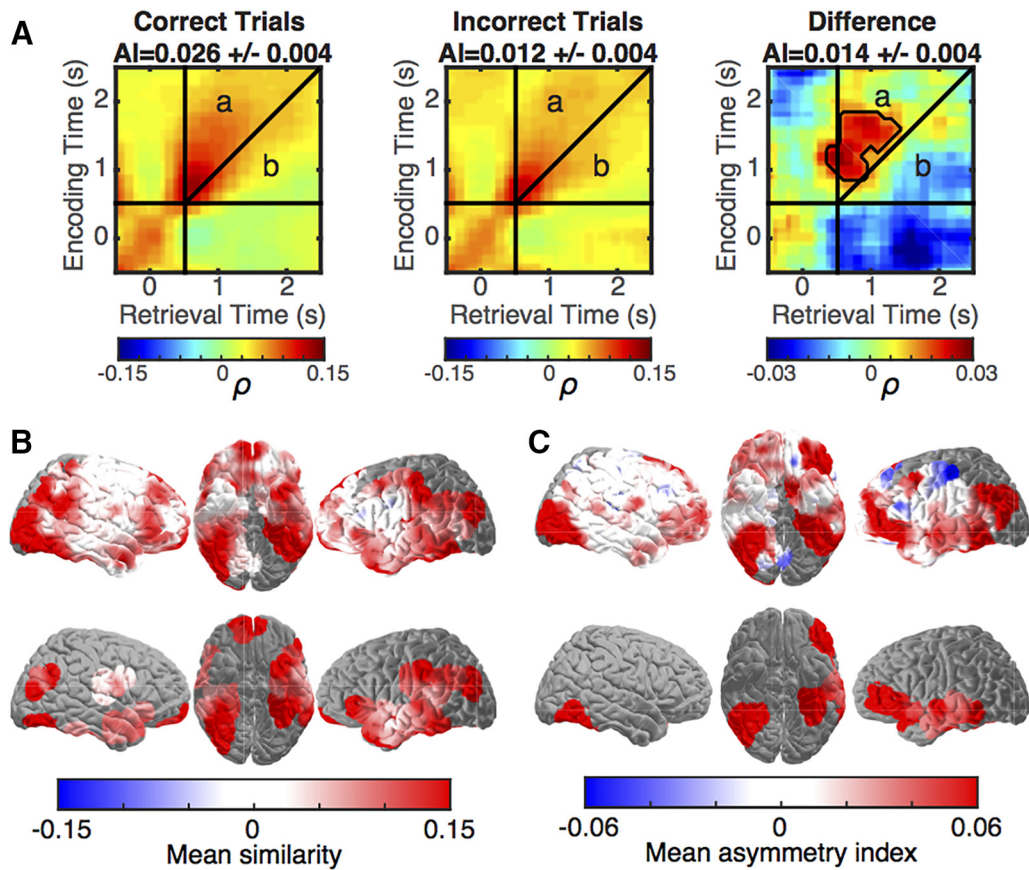


Figure 2. Similarity maps and asymmetry indices. **A**, Similarity maps averaged within participants and then across participants, shown for correct and incorrect trials. The difference between the correct and incorrect trials is also shown. Color represents average cosine similarity or difference in cosine similarity. Asymmetry indices ($a - b$) are shown above each similarity plot. Horizontal black line indicates the first-time bin where all the data are from after pair presentation during encoding. Vertical black line indicates the first-time bin where all the data are from after cue presentation during retrieval. Diagonal black line is the $t_{enc} = t_{ret}$ line. Region outlined in black represents cluster where reinstatement is significantly greater for correct trials ($p < 0.01$). **B**, Top, Mean similarity computed for each electrode and averaged within spatial ROI and across participants. Color represents mean similarity across participants in that region. Colors are only plotted for regions that had electrode coverage in at least 5 participants. Bottom, Mean similarity for regions that belong to a spatial cluster exhibiting a statistically significant effect ($p < 0.01$, nonparametric clustering statistic). **C**, Same as **B** for mean asymmetry indices.

tion from all electrodes in a participant for every 500 ms temporal epoch during the encoding and retrieval periods. We quantified reinstatement between every temporal epoch during encoding and retrieval by calculating the cosine similarity between feature vectors. Thus, for a single trial, we generate a precise temporal map of neural reinstatement between the encoding and retrieval periods.

Each temporal map of reinstatement, averaged across all trials, demonstrates how similar the distributed neural activity during any encoding epoch is to the activity during any retrieval epoch. If both encoding and retrieval exhibit similar patterns of activity in identical temporal epochs, we would expect to see a temporal map of reinstatement with maximum values along the diagonal line (Fig. 1*B*). Any values of reinstatement above the diagonal line indicate the extent to which neural activity during encoding is similar to activity during retrieval occurring at an earlier time period. Conversely, any values of reinstatement below the diagonal line indicate that neural activity during encoding is similar to activity during retrieval occurring at a later time. For each participant, we therefore defined an *AI* as the difference in mean reinstatement for time regions above and below the diagonal line (Fig. 1*B*; see Materials and Methods) (Oemisch et al., 2015). In this manner, distributed neural activity occurring with identical time courses during encoding and retrieval would have an asymmetry index of 0, whereas re-

trieval activity that is similar but faster to activity during encoding will have an asymmetry index >0 .

Reinstatement of high gamma power exhibited an asymmetry index that was significantly >0 across participants for correct trials ($AI = 0.026 \pm 0.004$, $t_{(33)} = 6.60$, $p = 1.7 \times 10^{-7}$, two-tailed t test; Figure 2*A*; for incorrect trials, $AI = 0.012 \pm 0.004$, $t_{(33)} = 2.95$, $p = 0.006$). Correct trials exhibited significantly greater reinstatement than incorrect trials ($p = 0.023$, permutation procedure) and significantly greater temporal asymmetry (mean difference in $AI = 0.014 \pm 0.004$, $t_{(33)} = 3.56$, $p = 0.001$, paired two-tailed t test).

As successful retrieval involves reinstating distributed high gamma power at a faster time scale, we were interested in identifying which brain regions were responsible for this effect (ROIs; see Materials and Methods). We computed a temporal map of reinstatement and resulting asymmetry index for each electrode over all correct trials, and averaged the asymmetry indices from all electrodes within each ROI in each participant (see Materials and Methods). We visualized the average reinstatement and asymmetry indices for each ROI that had electrode coverage in at least 5 participants (Fig. 2*B, C*, top). Many spatially contiguous regions exhibited reinstatement that was significantly >0 (Fig. 2*B*, bottom; $p < 0.01$, permutation procedure; see Materials and Methods). These regions included the left and right ventral and

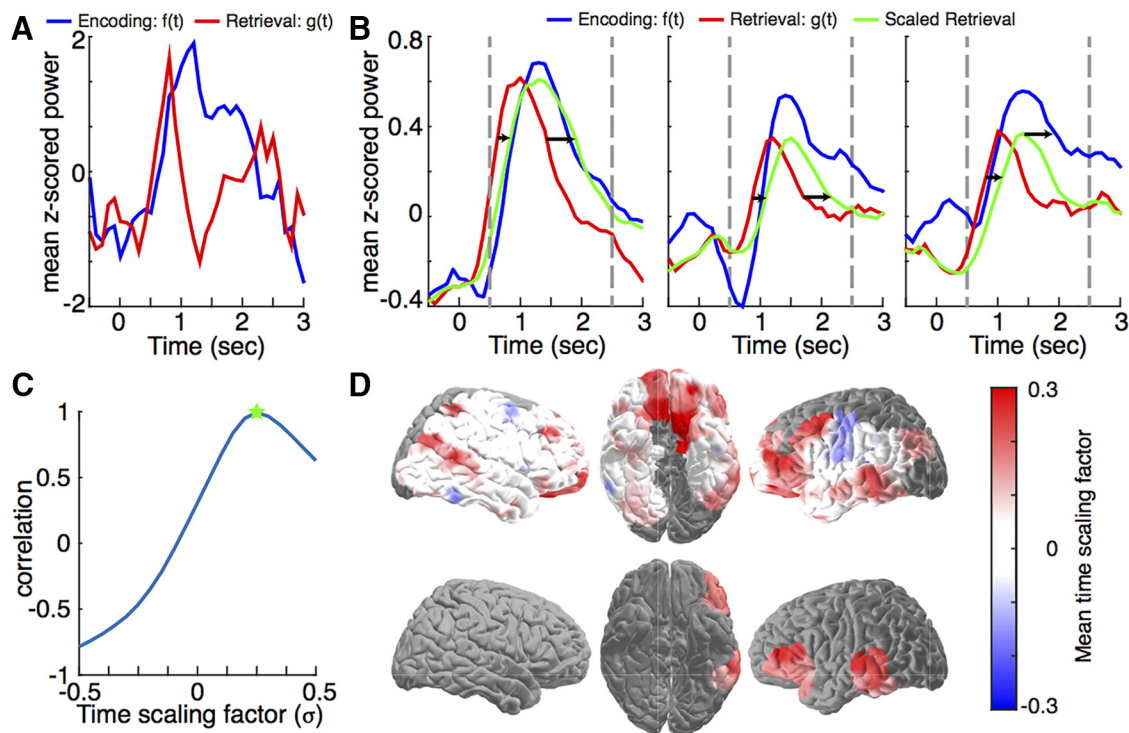


Figure 3. Time scaling analysis. **A**, Data from a single trial from an electrode in the anterior superior temporal gyrus. Mean z-scored high gamma power is plotted for encoding in blue and retrieval in red. **B**, Data from single electrodes from (left to right) the anterior superior temporal gyrus, the left ventrolateral prefrontal cortex, and the left middle temporal gyrus. Mean z-scored high gamma power is plotted for encoding in blue and retrieval in red. Green trace represents the retrieval trace scaled to best match the encoding trace. Vertical gray dashed lines indicate the time range over which the correlation was computed to optimize the time scaling factor. **C**, The correlation between $f(t)$ and stretched $g(t)$ for $\sigma > 0$ and the correlation between $g(t)$ and stretched $f(t)$ for $\sigma < 0$ plotted as a function of σ for the example shown in **B**, left. Green star indicates the maximum correlation, which corresponds to a σ of 0.25 in this case. **D**, Top, Mean time scaling factors for electrodes that achieved at least the median maximum correlation when the time scaling algorithm was applied. Color represents the mean time scaling factor across participants for that region. Colors are only plotted for regions that had electrodes included in this analysis from at least 5 participants. Bottom, Mean time scaling factors for regions that belong to a spatial cluster exhibiting a statistically significant effect ($p < 0.01$, nonparametric clustering statistic).

lateral temporal lobes, inferior frontal gyri, and occipital lobes. Spatially contiguous regions with asymmetry indices across participants that exhibited a significant deviation from zero included areas of the left temporal lobe and ventrolateral prefrontal cortex (Fig. 2C, bottom; $p < 0.01$, permutation procedure; see Materials and Methods). All of the identified regions were characterized by a mean asymmetry index across participants that was >0 , suggesting that spectral power is reinstated in these regions at a faster time scale during successful retrieval.

We next used a time scaling algorithm to determine the extent to which the continuous time change in spectral power for each electrode would need to be stretched during either encoding or retrieval so as to maximize the correlation between them (see Materials and Methods). In individual single trials, we observed that the changes in high gamma power during retrieval appeared to occur more quickly during retrieval than during encoding (Fig. 3A). For each electrode, therefore, we computed the average continuous time change in spectral power in the high gamma frequency band over all correct trials (Fig. 3B, individual electrode examples). A positive time scaling factor, σ , indicates that retrieval time needs to be stretched to match the time course of activity during encoding, and that therefore retrieval activity in that electrode occurs faster than activity during encoding (Fig. 3C). As all electrodes will exhibit some maximum correlation in activity between encoding and retrieval after scaling the time courses, we restricted our analysis to only those electrodes demonstrating a temporal correlation between encoding and retrieval that exceeded a minimum threshold. We set this threshold as the

median value of maximum correlations achieved across all electrodes in all participants after time scaling (see Materials and Methods). In this manner, we only included electrodes exhibiting similar changes in spectral activity during encoding and retrieval.

We calculated the average time scaling factor in each ROI containing electrodes that met this requirement in at least 5 participants (Fig. 3D, top). Most regions demonstrated either a positive time scaling factor or one that was close to zero. We identified regions within spatially contiguous clusters that exhibited a statistically significant effect across participants (Fig. 3D, bottom; $p < 0.01$, permutation procedure; see Materials and Methods). In the left temporal lobe and prefrontal cortex, high gamma power exhibits a faster time course during successful retrieval compared with encoding.

To understand the patterns of high gamma power that were responsible for our main findings, we identified all ROIs that were within clusters exhibiting both significant asymmetry and time scaling across participants (Figs. 2C, 3D). Two regions contained at least two spatial ROIs that met this criterion: the left ventrolateral prefrontal cortex and the left lateral temporal lobe. These regions included electrodes located in Brodmann areas 10, 11, 20, 21, 22, 37, 38, 41/42, 45, 46, and 47. We calculated the mean temporal maps of reinstatement for each electrode within each region across all correct trials and averaged these maps within and then across participants to calculate an asymmetry index for each region (Fig. 4). In both cases, the average temporal map of reinstatement demonstrated high gamma activity during

retrieval that occurred faster and was similar to activity during encoding.

Discussion

Our data demonstrate that patterns of high gamma power present during encoding are reinstated at a faster timescale during successful cued memory retrieval. We found that both distributed patterns of high gamma power across multiple brain regions and individual neural features exhibit similar yet faster dynamics when participants successfully retrieve verbal episodic memories. Our data therefore suggest that the temporal compression observed both in hippocampal and cortical spiking activity during spontaneous replay is also a feature of memory related reinstatement in the human cortex.

There is a growing body of evidence that successful memory retrieval involves reinstatement of neural activity (Johnson and Rugg, 2007; Johnson et al., 2009; Danker and Anderson, 2010; Manning et al., 2011; Staresina et al., 2012; Miller et al., 2013; Yaffe et al., 2014). Our data build upon these studies by demonstrating that reinstatement exhibits one of the principal features of replay, temporal compression (Lee and Wilson, 2002; Diba and Buzsáki, 2007; Euston et al., 2007; Karlsson and Frank, 2009; Gupta et al., 2010). Replay of spiking activity at a faster timescale has been well established in the hippocampus (Carr et al., 2011). Reinstatement of activity in the cortex is hypothesized to be initiated and coordinated by the hippocampus (Yonelinas, 2002; Norman and O'Reilly, 2003; Johnson et al., 2009). By demonstrating that cortical reinstatement also occurs on a faster timescale, our data lend support to this hypothesis by suggesting that cortical reinstatement and hippocampal replay may be related.

We found that the greatest temporal compression during retrieval involved high gamma activity in the left ventrolateral prefrontal cortex and left temporal lobe. High gamma activity has been linked with cortical activation (Manning et al., 2009; Burke et al., 2015), and these regions are involved in the encoding and retrieval of verbal memories (Burke et al., 2014a, b; Greenberg et al., 2015). Our data therefore suggest that the asymmetry observed when examining the reinstatement of distributed patterns of high gamma power across the brain is largely due to temporal compression in brain regions involved in memory formation.

One important difference between our data and studies examining replay is that temporal compression during replay has primarily been observed in the sequence of firing activity in single-unit hippocampal place cells during spatial navigation (Lee and Wilson, 2002; Diba and Buzsáki, 2007; Karlsson and Frank, 2009; Gupta et al., 2010). In our task, the sequence that is being replayed is a series of memory processing steps. Participants process the visually presented text, interpret that text, and activate associative memory processes. As such, although we find clear evidence that this process is faster during retrieval, there are not clearly defined stimuli and responses, making the interpretation of the replayed sequence more difficult.

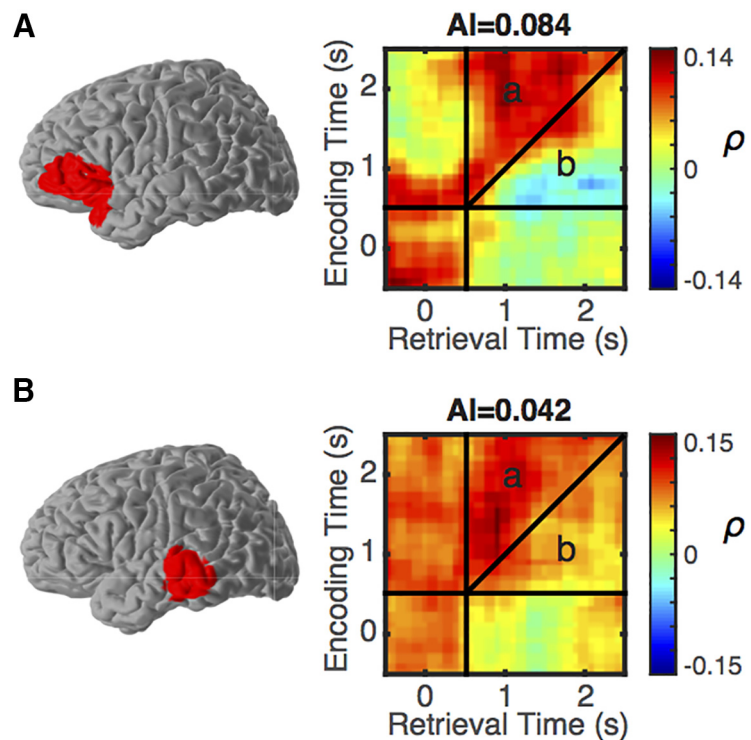


Figure 4. Analysis of specific anatomic regions exhibiting significant effects. **A**, Left, Red represents a region containing at least two spatial ROIs and that was part of a significant cluster in both the asymmetry index analysis (Fig. 2C) and the time scaling factor analysis (Fig. 3D). Right, Mean similarity map for the region indicated in **A**, left. Similarity map is computed using all correct trials for each electrode in the region, averaged across electrodes within a participant and across participants. Color represents average cosine similarity. Asymmetry indices ($a - b$) are shown above the similarity plot. Horizontal black line indicates the first-time bin where all the data are from after pair presentation during encoding. Vertical black line indicates the first-time bin where all the data are from after cue presentation during retrieval. Diagonal black line is the $t_{enc} = t_{ret}$ line. **B**, Same as **A** for additional region satisfying the same criteria as the region in **A**.

Indeed, one possible explanation for the observed temporal compression is that participants are viewing two words during encoding, yet only one word during the cued retrieval. Indeed, while the asymmetry of reinstatement was greater during correct trials, we also measured significant asymmetry during incorrect trials. This may be due to the asymmetry of the processes common to both correct and incorrect trials, such as reading and processing words. These common processes occur on a faster timescale during retrieval. However, if this were solely responsible for the temporal differences between encoding and retrieval, we would expect to observe no difference in the asymmetry of reinstatement between correct and incorrect trials. Instead, the speedup we observe in our data is specifically related to the processes of successful memory encoding and retrieval because we observe significantly greater asymmetry during correct trials.

Consistent with our previous study (Yaffe et al., 2014), we observed significant reinstatement in many brain regions (Fig. 2B). The regions that showed significant asymmetry (Fig. 2C) corresponded to a subset of these regions, and the regions that showed significant time scaling factors (Fig. 3D) were a further subset of the regions showing significant asymmetry. Our results suggest that, in many brain regions, activity during encoding and retrieval occurs on the same timescale, yet in some of these brain regions, retrieval activity is faster. Of note, we observed significant asymmetry, but not significant time scaling, in the ventral aspects of the temporal lobes, regions that are responsible for reading and processing words. Asymmetry without time scaling can occur if the peaks of the time-varying power traces occur at the same time during encoding and retrieval, but activation lasts

longer during encoding. Aligned peaks in activity would result in no difference in time scales between encoding and retrieval as determined by the time scaling algorithm, yet because of the longer-lasting activation during encoding, there will be significant asymmetry. Conversely, we observed both asymmetry and significant time scaling in the prefrontal cortex and lateral temporal lobe. Early and faster activation in these regions may indeed be a functionally important component of memory retrieval (Bowen and Kark, 2016).

Previous evidence suggests that reinstatement of neural activity is driven by specific cortical representations for individual items, although general encoding and retrieval mechanisms are also involved (Yaffe et al., 2014; Danker et al., 2016; Michelmann et al., 2016; Tomparly et al., 2016). In our data, however, measures of reinstatement are derived from multiple signals across multiple brain regions, and so the specific neural signature of an individual item or trial may be obscured. Furthermore, the absence of specific behavioral sequences in our task makes our data unable to identify evidence of temporal compression in individual trials. Hence, direct comparisons between temporal compression observed in our data and that observed during replay of sequences of place cell firing are difficult to interpret and would require further investigation of cortical reinstatement during individual trials.

This may also partly account for the differences in temporal compression observed in our data compared with earlier work on replay. Previous studies of hippocampal and cortical replay of single-unit spiking activity report a speedup of a factor of ≥ 6 (Euston et al., 2007; Carr et al., 2011). In our data, however, we report retrieval processes that are only ~ 1.5 times as fast as during encoding. This difference may be related to our analysis of trial averaged data. Moreover, this difference may also be related to the types of neural signals being compared. Whereas here we examine the reinstatement of high gamma power, reflecting aggregate neural activity averaged over larger time bins and larger regions of the cortex, studies of replay have focused on the activity of individual neurons. While temporal compression studies have focused on spiking neurons, however, a potential application of our approach could be to examine the temporal compression of oscillatory activity in local field potentials recorded in animals and humans. This could provide further insight into the nature of temporal compression.

Together, our data provide evidence that, during successful retrieval, distributed patterns of high gamma activity are reinstated at a faster timescale than they originally occurred during encoding. Our data therefore provide a possible link between previous studies of replay and cortical reinstatement because both processes exhibit evidence of temporal compression. Indeed, behaviorally, the retrieval of a memory appears to occur at a faster timescale than the original episode. For example, an individual can recall a path they traversed often in much less time than it took to traverse the path (Bonasia et al., 2016). Our data suggest a cortical mechanism for how this may occur, and that temporal compression during retrieval may be related to that observed during replay.

References

- Addison PS (2002) *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. Bristol, UK: Institute of Physics.
- Bonasia K, Blommestejn J, Moscovitch M (2016) Memory and navigation: compression of space varies with route length and turns. *Hippocampus* 26:9–12. [CrossRef Medline](#)
- Bowen HJ, Kark SM (2016) Commentary: episodic memory retrieval functionally relies on very rapid reactivation of sensory information. *Front Hum Neurosci* 10:196. [CrossRef Medline](#)
- Burke JF, Long NM, Zaghoul KA, Sharan AD, Sperling MR, Kahana MJ (2014a) Human intracranial high-frequency activity maps episodic memory formation in space and time. *Neuroimage* 85:834–843. [CrossRef Medline](#)
- Burke JF, Sharan AD, Sperling MR, Ramayya AG, Evans JJ, Healey MK, Beck EN, Davis KA, Lucas TH 2nd, Kahana MJ (2014b) Theta and high-frequency activity mark spontaneous recall of episodic memories. *J Neurosci* 34:11355–11365. [CrossRef Medline](#)
- Burke JF, Ramayya AG, Kahana MJ (2015) Human intracranial high-frequency activity during memory processing: neural oscillations or stochastic volatility? *Curr Opin Neurobiol* 31:104–110. [CrossRef Medline](#)
- Carr MF, Jadhav SP, Frank LM (2011) Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat Neurosci* 14:147–153. [CrossRef Medline](#)
- Danker JF, Anderson JR (2010) The ghosts of brain states past: remembering reactivates the brain regions engaged during encoding. *Psychol Bull* 136:87–102. [CrossRef Medline](#)
- Danker JF, Tomparly A, Davachi L (2016) Trial-by-trial hippocampal encoding activation predicts the fidelity of cortical reinstatement during subsequent retrieval. *Cereb Cortex*. Advance online publication. Retrieved Jun. 10, 2016. [CrossRef](#) [CrossRef Medline](#)
- Diba K, Buzsáki G (2007) Forward and reverse hippocampal place-cell sequences during ripples. *Nat Neurosci* 10:1241–1242. [CrossRef Medline](#)
- Dykstra AR, Chan AM, Quinn BT, Zepeda R, Keller CJ, Cormier J, Madsen JR, Eskandar EN, Cash SS (2012) Individualized localization and cortical surface-based registration of intracranial electrodes. *Neuroimage* 59:3563–3570. [CrossRef Medline](#)
- Euston DR, Tatsuno M, McNaughton BL (2007) Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science* 318:1147–1150. [CrossRef Medline](#)
- Greenberg JA, Burke JF, Haque R, Kahana MJ, Zaghoul KA (2015) Decreases in theta and increases in high frequency activity underlie associative memory encoding. *Neuroimage* 114:257–263. [CrossRef Medline](#)
- Gupta AS, van der Meer MA, Touretzky DS, Redish AD (2010) Hippocampal replay is not a simple function of experience. *Neuron* 65:695–705. [CrossRef Medline](#)
- Hanslmayr S, Staresina BP, Bowman H (2016) Oscillations and episodic memory: addressing the synchronization/desynchronization conundrum. *Trends Neurosci* 39:16–25. [CrossRef Medline](#)
- Haque RU, Wittig JH Jr, Damera SR, Inati SK, Zaghoul KA (2015) Cortical low-frequency power and progressive phase synchrony precede successful memory encoding. *J Neurosci* 35:13577–13586. [CrossRef Medline](#)
- Hoffman KL, McNaughton BL (2002) Coordinated reactivation of distributed memory traces in primate neocortex. *Science* 297:2070–2073. [CrossRef Medline](#)
- Howard MW, Viskontas IV, Shankar KH, Fried I (2012) Ensembles of human MTL neurons “jump back in time” in response to a repeated stimulus. *Hippocampus* 22:1833–1847. [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](#)
- Johnson JD, Rugg MD (2007) Recollection and the reinstatement of encoding-related cortical activity. *Cereb Cortex* 17:2507–2515. [CrossRef Medline](#)
- Kahn I, Andrews-Hanna JR, Vincent JL, Snyder AZ, Buckner RL (2008) Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *J Neurophysiol* 100:129–139. [CrossRef Medline](#)
- Karlsson MP, Frank LM (2009) Awake replay of remote experience in the hippocampus. *Nat Neurosci* 12:913–918. [CrossRef Medline](#)
- Lee AK, Wilson MA (2002) Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron* 36:1183–1194. [CrossRef Medline](#)
- Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19:1233–1239. [CrossRef Medline](#)
- Manning JR, Jacobs J, Fried I, Kahana MJ (2009) Broadband shifts in LFP power spectra are correlated with single-neuron spiking in humans. *J Neurosci* 29:13613–13620. [CrossRef Medline](#)
- Manning JR, Polyn SM, Baltuch GH, Litt B, Kahana MJ (2011) Oscillatory

- patterns in temporal lobe reveal context reinstatement during memory search. *Proc Natl Acad Sci U S A* 108:12893–12897. [CrossRef Medline](#)
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190. [CrossRef Medline](#)
- Michelmann S, Bowman H, Hanslmayr S (2016) The temporal signature of memories: identification of a general mechanism for dynamic memory replay in humans. *PLoS Biol* 14:e1002528. [CrossRef Medline](#)
- Miller JF, Neufang M, Solway A, Brandt A, Trippel M, Mader I, Hefft S, Merkow M, Polyn SM, Jacobs J, Kahana MJ, Schulze-Bonhage A (2013) Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science* 342:1111–1114. [CrossRef Medline](#)
- Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: a complementary learning systems approach. *Psychol Rev* 110:611–646. [CrossRef Medline](#)
- Nunez PL, Srinivasan R (2006) *Electric fields of the brain*. New York: Oxford UP.
- Oemisch M, Westendorff S, Everling S, Womelsdorf T (2015) Interareal spike-train correlations of anterior cingulate and dorsal prefrontal cortex during attention shifts. *J Neurosci* 35:13076–13089. [CrossRef Medline](#)
- Preston AR, Eichenbaum H (2013) Interplay of hippocampus and prefrontal cortex in memory. *Curr Biol* 23:R764–R773. [CrossRef Medline](#)
- Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, Litt B, Brandt A, Kahana MJ (2007a) Gamma oscillations distinguish true from false memories. *Psychol Sci* 18:927–932. [CrossRef Medline](#)
- Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, McCarthy DC, Brandt A, Tully MS, Kahana MJ (2007b) Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cereb Cortex* 17:1190–1196. [CrossRef Medline](#)
- Staresina BP, Henson RN, Kriegeskorte N, Alink A (2012) Episodic reinstatement in the medial temporal lobe. *J Neurosci* 32:18150–18156. [CrossRef Medline](#)
- Tomparay A, Duncan K, Davachi L (2016) High-resolution investigation of memory-specific reinstatement in the hippocampus and perirhinal cortex. *Hippocampus* 26:995–1007. [CrossRef Medline](#)
- Tulving E (1972) Episodic and semantic memory. In: *Organization of memory* (Tulving E, Donaldson W, eds), pp 381–403. San Diego: Academic.
- Yaffe RB, Kerr MS, Damera S, Sarma SV, Inati SK, Zaghoul KA (2014) Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proc Natl Acad Sci U S A* 111:18727–18732. [CrossRef Medline](#)
- Yonelinas AP (2002) The nature of recollection and familiarity: a review of 30 years of research. *J Mem Lang* 46:441–517.