

Neural Activity in the Medial Temporal Lobe Reveals the Fidelity of Mental Time Travel

James E. Kragel,^{1,2} Neal W. Morton,¹ and Sean M. Polyn¹

¹Department of Psychology and ²Neuroscience Graduate Program, Vanderbilt University, Nashville, Tennessee 37240

Neural circuitry in the medial temporal lobe (MTL) is critically involved in mental time travel, which involves the vivid retrieval of the details of past experience. Neuroscientific theories propose that the MTL supports memory of the past by retrieving previously encoded episodic information, as well as by reactivating a temporal code specifying the position of a particular event within an episode. However, the neural computations supporting these abilities are underspecified. To test hypotheses regarding the computational mechanisms supported by different MTL subregions during mental time travel, we developed a computational model that linked a blood oxygenation level-dependent signal to cognitive operations, allowing us to predict human performance in a memory search task. Activity in the posterior MTL, including parahippocampal cortex, reflected how strongly one reactivates the temporal context of a retrieved memory, allowing the model to predict whether the next memory will correspond to a nearby moment in the study episode. A signal in the anterior MTL, including perirhinal cortex, indicated the successful retrieval of list items, without providing information regarding temporal organization. A hippocampal signal reflected both processes, consistent with theories that this region binds item and context information together to form episodic memories. These findings provide evidence for modern theories that describe complementary roles of the hippocampus and surrounding parahippocampal and perirhinal cortices during the retrieval of episodic memories, shaping how humans revisit the past.

Key words: Computational modeling; episodic memory; fMRI; free recall; MTL

Introduction

The capabilities of the human memory system rival any internet search engine in terms of the flexibility and specificity of search. With a moment of thought, one can search through memories of the recent or distant past, retrieving a particular memory from many thousands. The memories one retrieves during search vary in fidelity. When a memory emerges that is rich in detail regarding the time and place of its original occurrence, the strong subjective sense that one is revisiting past experience inspired Tulving (1993) to call this phenomenon “mental time travel.”

Brain structures in the medial temporal lobe (MTL) are critically involved in the retrieval of memories of one’s past experience, i.e., episodic memories (Milner et al., 1998). Distinct subregions of MTL cortex have distinct functional properties. Perirhinal cortex, which is located on the anterior portion of the parahippocampal gyrus (PHG), becomes active during retrieval of objects and their features (Davachi et al., 2003; Haskins et al., 2008; Staresina and Davachi, 2010). In contrast, parahippocam-

pal cortex, on the posterior extent of the PHG, is engaged during processing of scene context (Hayes et al., 2007; Litman et al., 2009), spatial attributes (Burgess et al., 2002), and nonspatial contextual attributes (Aminoff et al., 2007; Bar et al., 2008). Activity in this region is also sensitive to the temporal context in which a visual scene is presented (Turk-Browne et al., 2012). One view of MTL function (Diana et al., 2007) describes parahippocampal cortex as encoding and retrieving contextual information, while perirhinal cortex mediates item-specific processing (Davachi, 2006; Mayes et al., 2007; Wixted and Squire, 2011). Both item and contextual information are thought to converge in hippocampus, allowing this region to represent objects in a specific temporal sequence (Hsieh et al., 2014) and support temporal memory (Ezzyat and Davachi, 2014).

While these theories outline the functional properties of the MTL, they do not specify the cognitive operations underlying neural signals from these brain regions, nor do they describe how these signals relate to memory search. Retrieved context models (Howard and Kahana, 2002; Polyn and Kahana, 2008) describe memory search in terms of two interacting cognitive processes: temporal reinstatement (TR) and item retrieval. Temporal reinstatement underlies mental time travel: when a particular studied item is remembered, the memory system reactivates a temporal code specifying when the study event occurred relative to other events in the episode. Reactivation of a high-fidelity code makes memories that were formed near that point in time more accessible. The temporal reinstatement operation is followed by a recall competition to determine which memories will be retrieved.

Received Aug. 13, 2014; revised Dec. 5, 2014; accepted Dec. 30, 2014.

Author contributions: J.E.K., N.W.M., and S.M.P. designed research; J.E.K. and N.W.M. performed research; J.E.K., N.W.M., and S.M.P. contributed unpublished reagents/analytic tools; J.E.K. analyzed data; J.E.K. and S.M.P. wrote the paper.

This research was supported by a Vanderbilt Discovery Grant and by National Science Foundation Grant 1157432 to S.M.P. We thank Joshua McCluey for assistance with data collection and initial analyses of the data.

The authors declare no competing financial interests.

Correspondence should be addressed to Sean M. Polyn, Vanderbilt University, Department of Psychology, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37240. E-mail: sean.polyn@vanderbilt.edu.

DOI:10.1523/JNEUROSCI.3378-14.2015

Copyright © 2015 the authors 0270-6474/15/352914-13\$15.00/0

A successful competition yields a remembered item, while a failed competition leads to termination of memory search.

To test whether subregions of the MTL are involved in these cognitive operations, we developed a set of neurally informed models of memory search. These models were designed to predict the sequence of recalls made by a participant in a free recall task, while neural activity was measured using functional MRI. In the TR model, neural signal at the time of a recall event controls the fidelity with which the temporal code associated with the retrieved memory is reactivated. In the retrieval success (RS) model, neural signal controls the balance between success and failure in the recall competition. In this approach, the computational model bridges between neural signal and human behavior, providing a framework to test hypotheses regarding the computational mechanisms supported by specific MTL structures.

Materials and Methods

Participants and behavioral task

Twenty native English speakers (13 female) between 18 and 35 years of age participated in the study, which was conducted in accordance with the procedures approved by the Vanderbilt University Institutional Review Board, after giving informed consent. Participants received compensation of \$20/h, with the opportunity to additionally earn up to \$10 based upon task performance.

Each participant performed 12 trials, spread across two scanning sessions, with 6 trials per session. On each trial, a list of 24 words was presented serially. On half of the trials, list learning was followed by a free-recall test. On the remainder of trials, a source recognition task was used to test memory. The order of the trials was determined pseudorandomly within each session. In the present work, we examine only the free-recall trials. As each word was presented, the participant performed an orienting task, in which they made a button press indicating either the size or animacy of the word, using an MRI-compatible response box. A prior study (Polyn et al., 2012) characterizing task-sensitive patterns of neural activity in these data may be consulted for other methodological details.

For each studied item, the participant first saw a pretrial cue (duration, 0.7 s) indicating which judgment to make, followed by a fixation cross (duration, 0.3 ± 0.1 s) and the study word (duration, 2.5 s). Items were separated by an interstimulus interval containing a fixation cross (duration, 0.5–5 s), with timing optimized to increase efficiency in the estimation of the response to each judgment type (Dale, 1999). A retrieval precue (duration, 2 s) followed the final item, indicating whether the current trial involved free recall or source recognition. Finally, a row of asterisks and an auditory tone (duration, 0.5 s) indicated the start of the recall period (duration, 75 s). During this time, subjects were instructed to recall as many items as they could from the most recent list, in any order. Vocal responses were recorded with a scanner-safe microphone (Resonance Technologies), and were scored using *PyParse* (Solway et al., 2010) and *Penn TotalRecall*.

Image acquisition

Imaging was performed using a 3 T Philips Intera Achieva magnet at the Vanderbilt University Institute of Imaging Science. Functional data were collected using an interleaved gradient echo T2*-weighted pulse sequence with BOLD contrast (repetition time = 2000 ms, echo time = 30 ms, flip angle = 75°, voxel size = $3.0 \times 3.0 \times 3.6$ mm, field of view = 192 mm). Thirty oblique slices, oriented parallel to the AC–PC plane were collected during functional scanning. Whole-brain structural scans were acquired using an MP-RAGE sequence (repetition time = 2500 ms, echo time = 4.38 ms, flip angle = 8°, voxel size = $1.0 \times 1.0 \times 1.0$ mm, field of view = 256 mm).

Image processing

The first four functional images of each session were discarded to allow the scanner signal to equilibrate. Preprocessing was carried out using the SPM8 software package, as follows: the remaining functional volumes

from each run were realigned to the first functional scan to correct for head motion. The mean of this realigned series was computed and coregistered to the T1 structural image. The T1 image was then segmented (i.e., separated into gray matter, white matter, and CSF), and normalized to Montreal Neurological Institute (MNI) stereotaxic space using the unified segmentation approach (Ashburner and Friston, 2005). Voxels were resampled to 3 mm isotropic and smoothed spatially, using an 8 mm FWHM Gaussian kernel.

Construction of regions of interest

BOLD time series were extracted from the MTL, as well as a set of nearby temporal fusiform regions previously shown to play a role in the retrieval of temporal sequences (Ekstrom and Bookheimer, 2007; Lehn et al., 2009). Extrahippocampal regions of interest (ROIs), including anterior and posterior regions of the parahippocampal gyrus and temporal fusiform cortex, were constructed using the Harvard-Oxford atlas. Hippocampus was defined as all hippocampal subregions (i.e., the subicular complex, dentate gyrus, and cornu ammonis), as well as the entorhinal cortex, from the Jülich Atlas probabilistic maps (Amunts et al., 2005; Eickhoff et al., 2005). A 25% threshold was applied to all probabilistic maps, resulting in an ROI of 2024 voxels within the MTL and surrounding cortical regions. We created an additional set of ROIs that partitioned both MTL cortex (eight ROIs) and hippocampus (six ROIs) along their anterior–posterior axes. Each ROI was up to 6 mm in length, limited by the anterior–posterior extent of the MTL. ROI placement respected the anatomical boundary between perirhinal cortex and parahippocampal cortex (Insausti et al., 1998).

A neurally informed computational model of memory search

Overview of the general approach. We used a computational model of memory search as part of an analytic framework designed to determine whether the neural signal from the MTL contained information that could improve the ability of the model to predict the sequence of responses made by participants during the recall task. This model was based on the Context Maintenance and Retrieval (CMR) model of free recall (Polyn et al., 2009). We created a set of model variants to test specific hypotheses regarding the functional properties of the neural signal recorded from different regions of the MTL.

The analysis framework is presented schematically in Figure 1. Figure 1a describes the neural signals incorporated into the model. We recorded BOLD time series from voxels within a priori anatomical ROIs, as participants performed a memory search task. The BOLD signal at the time of each recall event was collected and normalized across the events in a given trial.

Figure 1b demonstrates how a set of model variants is used to infer what cognitive operations are reflected in BOLD signal changes during memory search. For a given neural signal, two neurally informed model variants were created. The behavior and predictions of each model variant are determined by a set of model parameters, each of which corresponds to a different cognitive mechanism. For each neurally informed model, the degree of engagement of a particular cognitive mechanism was controlled by a neural signal sampled at the level of individual recall events.

On each trial of the experiment, participants produced a series of responses, which were coded as recall events. An example recall sequence is presented, in which the participant recalled six items before terminating recall. This was coded as a series of seven events, one for each of the recalled items, and a termination event. For each recall event, the model produced a likelihood score: the estimate of the probability of that particular response being generated by the model, given the particular set of recall events leading up to that response. These probability scores were collected in a vector $\mathbf{p}_{\text{event}}$. The fitness of a model, given a particular set of parameters, was then determined as follows:

$$L = - \sum \log(\mathbf{p}_{\text{event}}). \quad (1)$$

By this equation, a smaller value of L (the negative natural logarithm of the maximum likelihood for each model) corresponds to a better fitting model (i.e., one that makes more accurate predictions regarding the

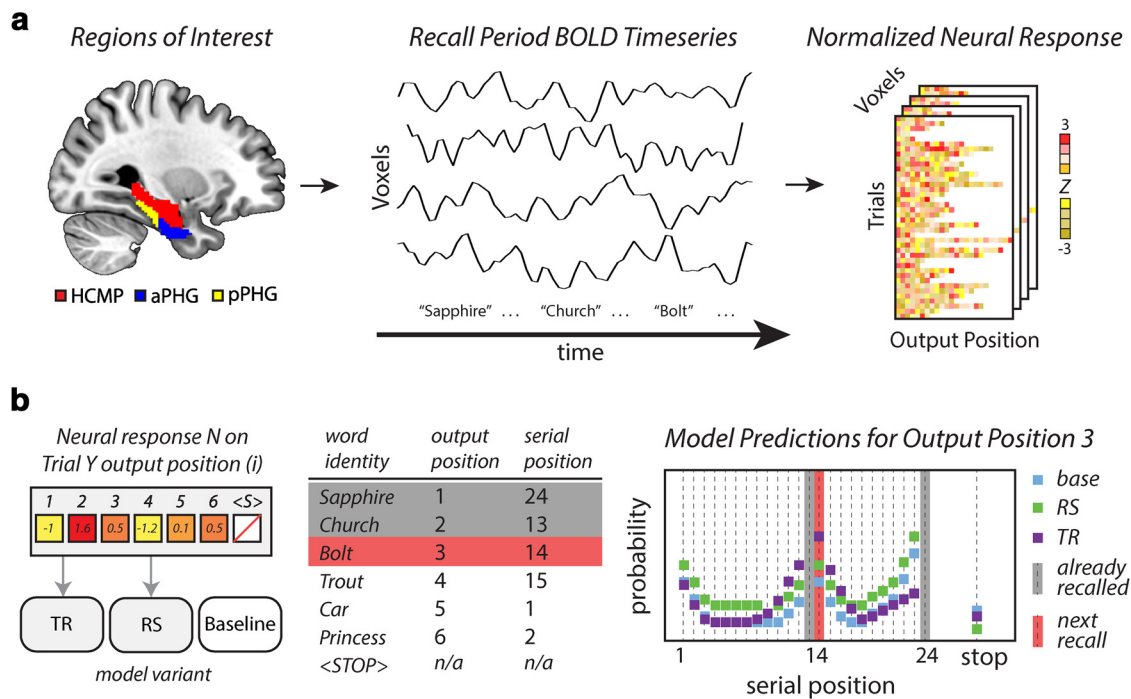


Figure 1. Schematic overview of the computational analysis framework. **a**, BOLD signal from a priori regions of interest within the MTL, including hippocampus (HCMP; red), anterior PHG (aPHG; blue), posterior PHG (pPHG; yellow), and temporal fusiform cortex (anterior and posterior regions not shown) is recorded while participants recall items from memory. The BOLD time series is sampled at the onset of vocalization of a given recall event, and neural responses are normalized within the trial. **b**, For a given neural signal, three models are created. In the TR and RS models, this neural signal controls the value of a specific parameter on a recall-by-recall basis. The parameters of the baseline model are not controlled by neural signal. A hypothetical recall sequence and the associated neural signal for seven recall events (including recall termination) demonstrate the predictions of each model. Each model determines a set of probability estimates for a given recall event, corresponding to the likelihood that each studied item is recalled (serial position), as well as the likelihood that no more items are recalled (stop). Given the increased MTL signal during the second retrieval event, the TR model (purple squares) predicts that nearby items are more likely to be recalled next. The RS model (green squares) predicts a general increase in the likelihood that any studied item will be retrieved, relative to the baseline model (blue squares). The red bar indicates the item that was actually recalled. The likelihood estimates made by each model are aggregated across all recall events to determine the overall likelihood score associated with each model.

recall sequences). For each model variant, an optimization procedure was used to estimate the set of parameters that maximized the likelihood that the model generated the observed behavioral data.

The predictive power of the neurally informed model variants were compared with the predictions of a baseline version of the model (which was not informed by neural data). If the BOLD signal from a particular region was informative about a model mechanism, this would improve the ability of the neurally informed model to predict recall sequences, resulting in a significant improvement to the global likelihood score associated with that model.

We additionally created a generative version of the baseline model, which generates synthetic recall sequences for a given set of parameters. In this version of the model, the behavior of the model during recall is stochastically determined by the probability of each possible recall event (i.e., recall of a list item or termination), rather than the observed recall sequences. The generative model was used to demonstrate the ability of the model to capture important summary statistics characterizing recall behavior.

Basic structure of the model. The CMR model is best described as a simplified neural network, in which two representational layers interact with one another via two sets of associative weighted connections. Each layer contains a set of elements with activation states that can vary between 0 and 1. The first layer contains a feature-based representation of studied material (**f**), and the second layer contains a contextual representation (**c**), where **f** and **c** are column vectors, each containing 50 elements. To simulate the presentation or retrieval of an individual item, the activation of **f** changes to represent that item. As a simplifying assumption, item representations in **f** correspond to a unit vector, with a single element in **f** activated. As such, 24 elements in **f** correspond to the individual items from the study list. The remaining elements are used to simulate the activation of task-irrelevant information during the inter-

stimulus intervals of the study period (24 elements) or during the end-of-list retention interval (1 element). The final element is used to initialize context at the start of each list. The dimensionality of **c** is constrained to match that of **f**. The activation patterns in both **c** and **f** are constrained to be of unit length.

These layers are connected by the following two associative matrices: one that allows featural representations to update the current state of context (the item-to-context matrix, M^{FC}), and another that allows context to reactivate item features during memory search (the context-to-item matrix, M^{CF}). Each of these matrices is initialized with a set of pre-experimental associations, corresponding to associations formed before the experimental session; learning processes create experimental associations as the list progresses.

The diagonal elements of M^{FC} allow an item representation **f** to influence **c**, when the item is first studied. A model parameter, γ , controls the relative strength of pre-experimental and experimental associations; these diagonal elements are initialized at $1 - \gamma$, with other elements initialized to zero. Elements of M^{CF} representing associations between context and individual list items are initialized at a uniform value determined by the model parameter α ; these pre-experimental associations are meant to represent a participant's knowledge of relationships between the studied items.

Simulating the study period. The study period of each trial was simulated as a sequence of 49 events, as follows: 24 study events, which simulated the presentation of a particular item; and 25 disruption events (one preceding each study event and one following the final study event). The equations below describe how the state of the model changes over the course of these events. We use the subscript $t - 1$ to indicate when a particular term represents the influence of a prior state of the model, where t indexes the set of 49 events. We have omitted the t subscript for terms representing the current state of the model.

To simulate presentation of an item during study, encoding processes activate the representation of that item, \mathbf{f}_i , where i indexes the serial position of the item on the list. This information projects through \mathbf{M}^{FC} to influence the contextual representation. An integrative mechanism governs context updating, such that the updated state of the contextual representation is a function of both the incoming information and the prior state of context, as follows:

$$\mathbf{c} = \rho \mathbf{c}_{t-1} + \beta \mathbf{c}^{\text{IN}}, \quad (2)$$

where

$$\mathbf{c}^{\text{IN}} = \mathbf{M}^{\text{FC}} \mathbf{f}, \quad (3)$$

and is normalized to unit length before updating context. Here, β is a model parameter controlling the amount of item information that is integrated into the current state of context, and ρ is a scaling parameter that ensures that \mathbf{c} is of unit length, as follows:

$$\rho = \sqrt{1 + \beta^2[(\mathbf{c}_{t-1} \cdot \mathbf{c}^{\text{IN}})^2 - 1]} - \beta(\mathbf{c}_{t-1} \cdot \mathbf{c}^{\text{IN}}). \quad (4)$$

During encoding, β is determined by the free parameter β_{enc} ; below, β will take on other values to simulate contextual updating during interitem intervals and recall.

After context is updated, an episodic learning process creates associative structures binding the current state of \mathbf{f} to \mathbf{c} , according to the following:

$$\mathbf{M}^{\text{FC}} = \mathbf{M}_{t-1}^{\text{FC}} + \gamma \mathbf{c} \mathbf{f}^{\text{T}} \quad (5)$$

$$\mathbf{M}^{\text{CF}} = \mathbf{M}_{t-1}^{\text{CF}} + \phi_i \mathbf{f} \mathbf{c}^{\text{T}}, \quad (6)$$

where γ controls the strength of experimental associations, and ϕ_i is a scaling factor that increases context-to-item associations for early list positions to capture the primacy effect (i.e., the enhanced recall performance associated with the first few list positions). Primacy scaling is determined by the following equation:

$$\phi_i = \phi_s e^{-\phi_d(i-1)} + 1, \quad (7)$$

where ϕ_s and ϕ_d are free parameters, and i indexes the serial position of the item.

Preliminary simulations suggested that this primacy gradient did not adequately describe the primacy effect in the behavioral data. As such, we incorporated a secondary primacy mechanism, which was inspired by prior theoretical work (Laming, 1999). By this mechanism, at the start of the recall period, the model partially reactivates the contextual state associated with the start of the study list. This process follows the contextual updating rule described in Equation 2, using the model parameter β_s to determine the fidelity of reactivation of start-of-list context. We found that a version of the model with both primacy mechanisms (primacy gradient and start-of-list contextual retrieval) was able to account well for the behavioral variance associated with the primacy effect.

The model also allows the contextual representation to be perturbed by task-irrelevant information during the interstimulus intervals (ISI) and the end-of-list retention interval (RI), as described in prior work (Sederberg et al., 2008). A representation of the task-irrelevant information is activated on \mathbf{f} and is allowed to update the contextual representation, following Equation 2, with the exception that the rate of integration is controlled by either β_{isi} or β_{ri} . We assume that no associative learning takes place during these intervals (i.e., \mathbf{M}^{FC} and \mathbf{M}^{CF} are not updated) and that these task-irrelevant representations do not enter the recall competition described below.

Simulating the recall period. On each trial, the participant made a series of responses, which were coded as a series of recall events and a termination event (Fig. 1*b*). Repetitions and recall errors (i.e., intrusions) were rare during this task. To simplify the dynamics of the model, we removed recall events corresponding to repetitions of already recalled items (46 events across 120 recall trials), as well as prior and extralist intrusions (a total of 37 events).

For each recall event, the context representation \mathbf{c} is projected through the associative weight matrix \mathbf{M}^{CF} , determining the relative support (\mathbf{s}) for each individual item representation in \mathbf{f} , as follows:

$$\mathbf{s} = (\mathbf{M}^{\text{CF}} \mathbf{c}). \quad (8)$$

Here, \mathbf{s} is a vector in which each element corresponds to the support for a given studied item. If a participant were to recall an item that was assigned a zero probability in the recall competition, the likelihood metric used to assess model performance would assign a log-likelihood of negative infinity to that model. To avoid this possibility, we introduced the assumption that there is always some minimal support for any item (for the present simulations, this was set to 10^{-6}). Once \mathbf{s} is created, a recall competition is simulated. The probability of the model recalling each of the not-yet-recalled study items (p_i , where i indexes serial position) and the probability of recall termination (p_{stop}) are calculated.

The p_{stop} term is determined by an exponential function based on the ratio of support for items that have not yet been recalled to items that have been recalled, as follows:

$$p_{\text{stop}} = \xi_s + e^{-\xi_d s_{\text{nr}}/s_r}. \quad (9)$$

Here, ξ_d is a model parameter controlling the rate at which p_{stop} rises with output position, and s_r and s_{nr} are scalar values corresponding to the summed support for the already recalled items and the not-yet-recalled items, respectively. For the present simulation work, ξ_s was fixed to a value of 10^{-3} and corresponds to the probability that recall terminates before any items are retrieved. The ξ_d parameter is controlled by neural signal in the RS version of the model, which is described below.

We then set the support values for already recalled items to zero and calculate the probabilities associated with the recall of each of the remaining studied items (p_i), as follows:

$$p_i = (1 - p_{\text{stop}}) \frac{s_i}{N}, \quad (10)$$

$$\sum_k s_k$$

where N is the number of items in the list.

At this point, the element of p corresponding to the item that was actually reported by the participant (or p_{stop} , if the participant terminated recall at this point) is appended to the vector $\mathbf{p}_{\text{event}}$ (from Eq. 1). If this is a termination event, the trial is over, and the model is initialized for the next trial. If this is a recall event, the model simulates recall of that item. First, the featural representation of the item is reactivated on \mathbf{f} . The reactivated item representation is then projected through \mathbf{M}^{FC} , which retrieves the temporal code associated with that item. This updates the state of \mathbf{c} , following Equation 2, setting β to β_{rec} . The β_{rec} parameter is controlled by neural signal in the TR version of the model, which is described below.

Incorporating BOLD signal into a model. To test hypotheses regarding potential links between BOLD activity in the MTL and particular model mechanisms, we constructed a number of model variants in which neural signal influences specific model parameters on a recall-by-recall basis. In these models, a given model parameter, θ , is updated for each recall event, according to the following:

$$\theta_{\text{event}} = \theta + \nu N_{\text{event}} \quad (11)$$

where N_{event} represents the neural signal recorded for this recall event (i.e., BOLD signal measured at the onset of vocalization), and ν is a neural scaling parameter determining the influence of the neural signal on the model parameter θ . Thus, θ_{event} represents the value of that model parameter for this particular recall event.

We constructed a TR model in which θ corresponds to β_{rec} (which controls contextual integration during retrieval), and the neural signal is scaled by ν_{TR} . We also constructed an RS model in which θ corresponds to ξ_d (which controls the rate p_{stop} increases as items are recalled), and the neural signal is scaled by ν_{RS} . For the TR model, ν_{RS} was fixed to zero; for the RS model, ν_{TR} was fixed to zero. The baseline model can be considered a restricted variant of either of the neurally informed models, in that

both ν_{TR} and ν_{RS} were fixed to zero. Finally, we created a joint model, which allowed both ν_{TR} and ν_{RS} to take nonzero values.

Parameter estimation techniques

For a given version of the model, we estimated the set of parameters that best allows the model to capture the observed behavioral data, using a particle swarm optimization (PSO) algorithm to search over the parameter space (Eberhart and Kennedy, 1995). This technique creates a set of particles, each of which can be thought of as a point moving through the high-dimensional parameter space, searching for the region of this space that maximizes the fitness of the model. This technique was used to estimate the optimal parameters for the baseline model and was used three times for each voxel containing neural signal (for the TR, RS, and joint models).

A swarm was constructed by creating 40 uniformly distributed random parameter sets, each of which serves as the initial position for one of the particles. The movement of each particle was determined by the velocity vector \mathbf{v} , which was added to the position vector \mathbf{x} to determine the position of the particle on each iteration of the search. The set of positions visited by each particle was saved as the search progressed, allowing the algorithm to keep track of the best-fit position of each particle (\mathbf{b}) and the global best fit over all particle histories (\mathbf{g}).

Three steps were repeated until the average change in fitness over 50 generations did not change by more than a specified tolerance (10^{-7}). For each iteration of the PSO algorithm, (1) the fitness L of the parameter set associated with each particle was evaluated; (2) the individual and global best fitness values and positions were updated; and (3) the velocities and positions of each particle were updated.

The velocity of each particle, \mathbf{v}_p , in the swarm was determined as follows:

$$\mathbf{v}_i(t+1) = w\mathbf{v}_i(t) + c_1r_1[\mathbf{b}_i(t) - \mathbf{x}_i(t)] + c_2r_2[\mathbf{g}(t) - \mathbf{x}_i(t)], \quad (12)$$

where i indexes a given particle, and t indexes the generation of the particle search. The i th particle is represented by its position in a high dimensional search space, with dimensionality determined by the number of parameters to be fit. The value of \mathbf{b}_i is the position of an individual particle that produced the best fitness value over the history of the search, and $\mathbf{g}(t)$ is the best-fitting position in parameter space over all particles. In the current work, the PSO parameters c_1 , and c_2 were fixed at 0.5 and 1.25, respectively. The values of r_1 and r_2 were pseudorandom values drawn from the standard uniform distribution on the interval (0,1) during each generation of the PSO algorithm.

The parameter w controls the inertia of the particles in the swarm, and was updated as follows:

$$w = 0.9 - 0.4(t - 1)/(T - 1), \quad (13)$$

where t is the current generation, and T is the maximum number of generations, which was set to 1000 in the implemented parameter searches. For all of the parameter estimation searches reported here, the algorithm terminated search before reaching T .

Model comparison techniques

Once the optimal set of parameters was estimated for the baseline model and a neurally informed model, a likelihood ratio test (Wilks, 1938) was used to determine whether any improvement in model predictions for the neurally informed model was significant. This produced a test statistic, D (also known as deviance, twice the difference of the log-likelihood of the neural and baseline models), for each voxel within a priori regions of interest. Significance was determined by testing D on a χ^2 distribution with 1 df.

Once D was calculated for all voxels, a Monte Carlo resampling procedure [3dClustSim from AFNI, National Institute of Mental Health, Bethesda, MD (<http://afni.nimh.nih.gov/afni>)] was used to identify significant clusters, correcting for multiple comparisons across the entire search region. First, individual voxels surviving an uncorrected threshold of $p < 0.05$ ($D > 3.8415$) were identified, and a resampling procedure

indicated that a cluster extent of 31 would ensure a cluster-level $p < 0.05$, corrected for multiple comparison within the predefined anatomical ROI.

Aikake's information criterion (AIC) was used to evaluate model fitness across multiple models (Burnham and Anderson, 2004; Wagenmakers and Farrell, 2004). This measure takes model complexity (i.e., number of free parameters) into account, allowing us to compare a neurally informed model to the baseline model. Unlike the likelihood ratio test, this method does not require nested models, and was used to compare the TR, RS, and joint models to one another. For each model, we computed the AIC with a correction for finite samples, as follows:

$$AIC_c = 2L + 2V + \frac{2V(V+1)}{(n-V-1)}, \quad (14)$$

where the L value is for the candidate model (from Eq. 1), V is the number of free parameters, and n is the number of estimated data points. AIC weights were computed to determine the probability that each model (of K competing models) generated the observed data, under the assumption that one of the models generated the data, as follows:

$$w_i AIC = \frac{\exp\left(-\frac{1}{2}\Delta_i AIC\right)}{\sum_{k=1}^K \exp\left(-\frac{1}{2}\Delta_k AIC\right)}. \quad (15)$$

Here, $\Delta_i AIC$ is the difference in AIC_c between a given candidate model (specified by i) and the best-fitting model in the set.

To determine which of the models was most supported for a given cluster of voxels, we created an average model representative of that cluster. The parameter values $\hat{\theta}$ for this representative model were calculated as a weighted average of the parameter values $\hat{\theta}$ for all of the models associated with the voxels in that cluster, as follows:

$$\hat{\theta} = \sum_{i=1}^N w_i AIC \hat{\theta}_i, \quad (16)$$

where $\hat{\theta}_i$ is the set of estimated parameters from voxel i in the set of N voxels in the cluster. This approach was used to calculate the relative fitness of the TR, RS, and joint models for a given cluster.

To perform model comparisons for a given anatomical ROI, we used a bootstrap sampling procedure. For each ROI, informative voxels (i.e., those identified by either the TR or the RS analyses) were selected at random with replacement, allowing for the construction of a distribution of AIC weights for each model type, for each ROI. Bootstrap statistics were used to compare these distributions to one another and to test for linear trends across anatomical ROIs (at $p < 0.05$, Bonferroni corrected for multiple comparisons).

Results

Behavioral characterization of recall sequences

During the memory search period of the free-recall task, participants produced a series of responses; each response is referred to as a "recall event." Figure 2 (red lines) summarizes the major behavioral phenomena present in these recall sequences in terms of how participants initiated recall, transitioned between recall events, and, finally, terminated search. Figure 2a depicts the probability of initiating recall with each studied item, arranged in their order of presentation. Participants most often initiated recall with the final list position (demonstrating the recency effect), but sometimes started with the first list position (demonstrating the primacy effect), with a small but nonzero probability of initiating recall with some mid-list item. Our analysis of recall transitions focuses on the temporal organization of the response sequences. Figure 2b depicts the probability of transitions of different lengths, with length measured as the positional distance in the study list (i.e., the lag) between any two successively recalled

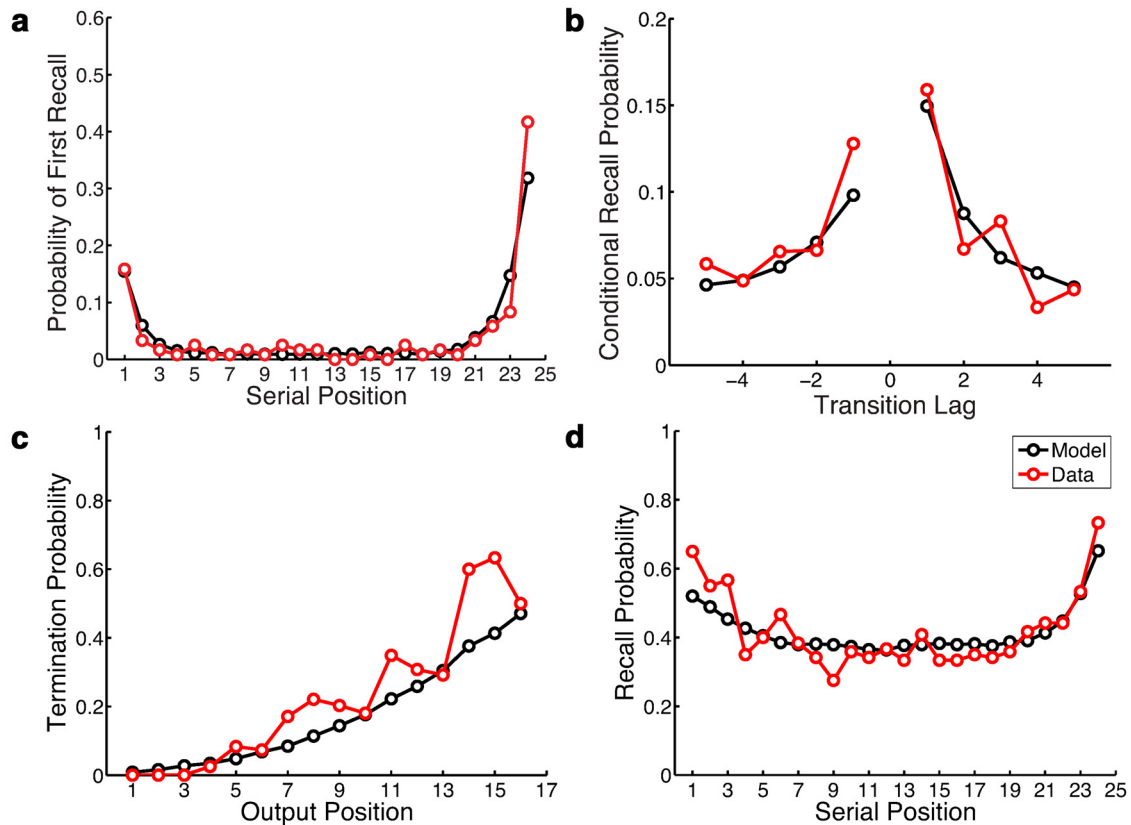


Figure 2. Comparison of summary statistics of behavioral performance (Data; red lines) with predictions of the best-fitting baseline model (Model; black lines). The parameter estimation process found the set of parameters that maximized the ability of the model to predict the likelihood of individual recall events; the model was not explicitly fit to these summary statistics. **a**, Recall initiation, as measured by the probability of first recall by serial position. **b**, Recall transitions, as measured by the probability of transitions in which the two successively recalled items were separated by a particular positional lag. **c**, Recall termination, as measured by the probability of stopping recall as a function of output position for the first 16 output positions. **d**, Overall recall performance, as measured by the probability of recalling an item as a function of its serial position.

items. When a particular item was recalled, there was a strong tendency for the next recalled item to come from a nearby list position, demonstrating the contiguity effect, a nearly ubiquitous phenomenon in free-recall response sequences (Kahana, 1996, 2012; Kahana et al., 2008). The peaks for small values of lag are taken as evidence for the temporal organization of the studied material in memory. Temporal organization is taken as a marker of mental time travel, as this behavior indicates that memories formed near in time to the just-recalled item have become more accessible.

At some point, participants failed to recall any more items, and the recall process terminated (Miller et al., 2012). Figure 2c depicts the probability of recall termination as a function of the number of items recalled (i.e., output position). Finally, Figure 2d depicts the probability that a given studied item appeared anywhere in the recall sequence, demonstrating again the primacy and recency effects, as well as the relatively flat probability of recalling a mid-list item. On average, participants recalled 42% of the studied items before the end of the recall period.

Comparison of neurally informed models to a baseline model

To determine whether the model provides a good account of the core behavioral phenomena of free recall, we developed a generative version of the model, which produces synthetic recall sequences for a given parameter set. Using the parameter settings for the optimized baseline model, we generated a large number of synthetic recall sequences and created model-generated versions of the four critical summary statistics described above. Figure 2 compares these synthetic recall sequences (black lines) to the

Table 1. Model fitness of representative models

Model	Cluster	k	Center of mass (mm)	n	L
TR	Left MTL	167	−28, −27, −21	11	3581.5
TR	Right pPHG	63	25, −32, −18	11	3580.4
RS	Right aMTL	336	28, −6, −32	11	3570.8
RS	Left MTL	554	−28, −16, −27	11	3571.7
RS	Right pPHG	52	17, −31, −10	11	3578.81
Joint	Left MTL	93	−27, −26, −21	12	3576.91
Baseline				10	3586.7

Center of mass values are reported in MNI space. *L* was computed using the average neural signal and representative model for each significant cluster. *n*, number of free parameters; *k*, extent of each cluster; pPHG, posterior PHG; aMTL, anterior MTL.

observed recall sequences (red lines). The correspondence between the black and red lines demonstrates a good qualitative fit to the behavioral dynamics of free recall. A quantitative description of the fitness of the model is provided in Table 1.

The neurally informed TR model allows a signal from a particular MTL voxel to determine the success of the contextual reactivation operation for each recall event. In contrast, the baseline model assumes that the parameter controlling temporal reinstatement (β_{rec}) is stationary over the course of the recall period. The baseline model can capture the average degree of temporal organization in the recall sequences (Fig. 2b), but the addition of a neural signal can improve the predictive power of the model if it indicates, for a particular recall event, the fidelity with which that temporal code is reactivated. That is, a neural signal can potentially allow the model to capture not only the

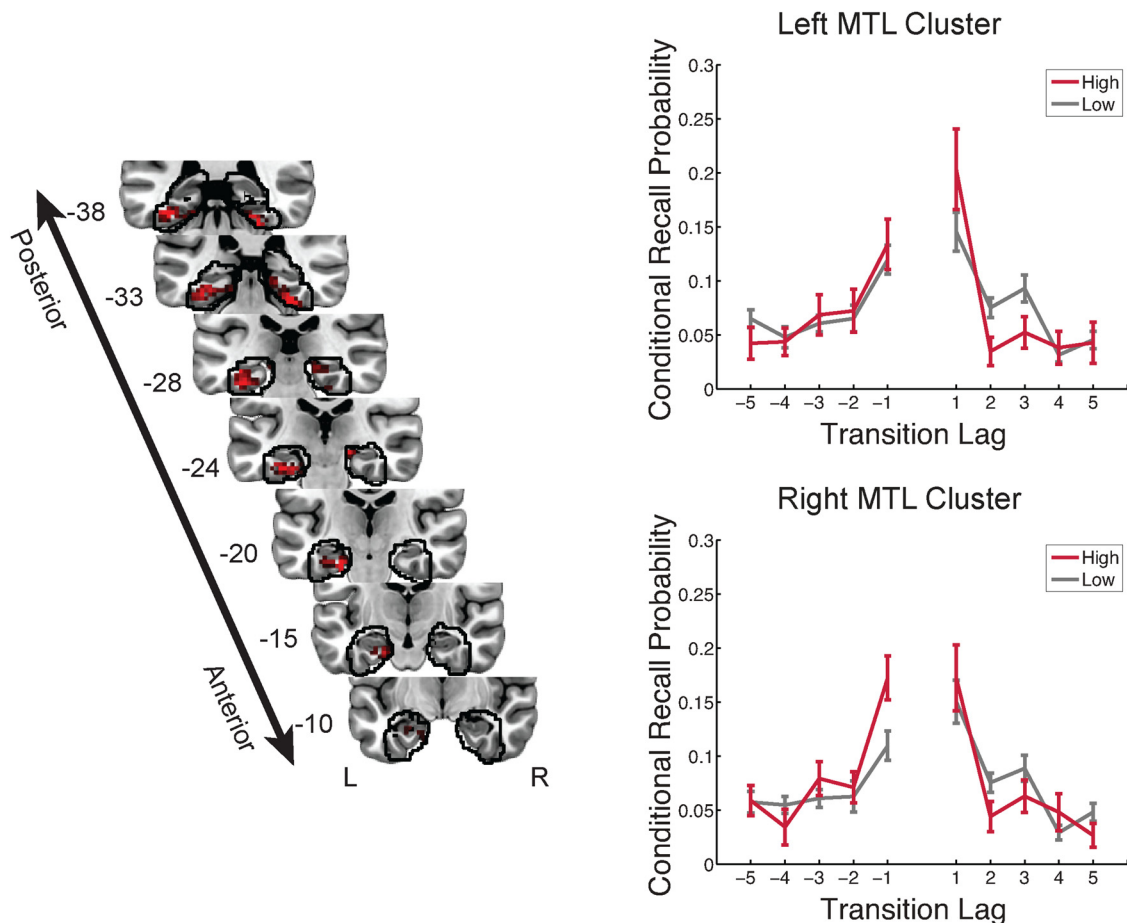


Figure 3. Model evidence for temporal reinstatement processes in the MTL. Left, A deviance map showing clusters of informative voxels ($p < 0.05$, corrected) within anatomically defined regions of interest, outlined in black. Right, Recall events were partitioned on the basis of neural activity in each MTL cluster, and an analysis of temporal organization was carried out separately for each cluster. Periods of high neural activity were associated with increased temporal organization. The y -value indicates the probability (\pm SEM) that the next recalled item will have a particular lag to the just-recalled item (conditional on the availability of that item for recall). The y -coordinate of coronal sections from the group average anatomical image is given in MNI space. R, Right; L, left.

mean level of temporal organization, but also the variability in organization, at the level of individual recall events.

We identified two clusters of voxels within the MTL that significantly improved the predictions of the TR model relative to the baseline model. Figure 3 presents a D map, which indicates the predictive power of the TR model relative to the neurally naive baseline model, on a voxel-by-voxel basis. One cluster consisted of a number of voxels in left MTL cortex extending from entorhinal cortex through posterior PHG, as well as posterior temporal fusiform cortex. A second cluster in right MTL extended from the hippocampus through posterior PHG, with the caudal extent of activation reaching posterior temporal fusiform cortex (for more detail, see Table 2). Consistent with the hypothesis that engagement of the parahippocampal cortex during retrieval reflects the reinstatement of contextual information, we observed positive neural scaling factors (ν_{TR}) linking activity in these clusters to the fidelity of temporal reinstatement.

To visualize the behavioral consequences of shifts in MTL activity during memory search, we partitioned the set of recall events into periods of high and low neural signal (split at 0.5 SDs above the mean), using a weighted average of the voxels within each MTL cluster. We separately constructed lag-based conditional response probabilities for these two sets of recall events. Figure 3 (right column) shows an increase in the temporal organization of recall events occurring during periods of increased MTL activity.

Table 2. Peak voxels within a priori regions of interest

Model	Region (hemisphere)	Peak coordinate (mm)	D
Temporal reinstatement	TFC (L)	-33, -37, -17	13.30
	ERC (L)	-18, -19, -29	13.28
	TFC (R)	27, -34, -23	11.57
	HCMP (R)	18, -25, -11	9.13
Retrieval success	aPHG (R)	21, 5, -23	29.52
	aPHG (L)	-18, -1, -23	24.91
	aPHG (R)	36, -10, -29	20.21
	HCMP (L)	-21, -10, -20	14.93
	HCMP (R)	18, -28, -14	11.58

Peak coordinate values are reported in MNI space. TFC, temporal fusiform cortex; ERC, entorhinal cortex; HCMP, hippocampus; aPHG, anterior PHG; R, right; L, left.

The neurally informed RS model allows signal from a particular MTL voxel to influence the relative likelihood that the model will continue to recall studied items, as opposed to terminating recall. Three clusters of voxels were identified ($p < 0.05$, corrected) that improved the predictions of the RS model relative to the baseline model, within a priori regions of interest. Two bilateral clusters were composed mostly of anterior PHG, including perirhinal cortex, the anterior extent of the hippocampus, and anterior parahippocampal cortex. The third cluster spanned right hippocampus and adjacent posterior PHG. The neural signal in these models was associated with a positive neural scaling factor (ν_{RS} ; Table 3); increased neural activity indicated a reduced like-

Table 3. Average model parameters from best fitting models

Parameter	Temporal reinstatement		Retrieval success			Joint
	Right pPHG	Left MTL	Right pPHG	Right aMTL	Left MTL	Left MTL
β_{enc}	0.33 (0.0076)	0.35 (0.018)	0.33 (0.0018)	0.33 (0.002)	0.33 (0.004)	0.34 (0.0083)
β_{rec}	0.82 (0.018)	0.81 (0.016)	0.86 (0.0031)	0.86 (0.0018)	0.86 (0.0021)	0.82 (0.017)
φ_s	1.69 (0.08)	1.48 (0.13)	1.69 (0.031)	1.81 (0.059)	1.75 (0.054)	1.54 (0.069)
φ_d	0.43 (0.0084)	0.42 (0.014)	0.43 (0.0035)	0.44 (0.0072)	0.43 (0.0075)	0.42 (0.0089)
γ	0.23 (0.011)	0.27 (0.018)	0.23 (0.0022)	0.23 (0.0012)	0.23 (0.0017)	0.26 (0.01)
ξ_d	2.37 (0.046)	2.35 (0.047)	2.37 (0.025)	2.50 (0.035)	2.44 (0.048)	2.36 (0.029)
α	0.05 (0.0019)	0.05 (0.0011)	0.05 (0.00023)	0.05 (0.00025)	0.05 (0.0005)	0.05 (0.0011)
ν_{TR}	0.35 (0.014)	0.10 (0.016)				0.09 (0.019)
ν_{RS}			0.35 (0.043)	0.57 (0.042)	0.50 (0.051)	0.26 (0.046)
β_s	0.28 (0.043)	0.28 (0.054)	0.23 (0.003)	0.22 (0.0064)	0.22 (0.006)	0.26 (0.016)
β_{ri}	0.76 (0.14)	0.75 (0.14)	0.82 (0.0035)	0.82 (0.0047)	0.82 (0.0056)	0.78 (0.039)
β_{ipi}	0.90 (0.0027)	0.90 (0.0044)	0.89 (0.00059)	0.89 (0.00038)	0.89 (0.00086)	0.90 (0.0034)

Data are reported as mean values from each cluster (SD). pPHG, Posterior PHG; aMTL, anterior MTL.

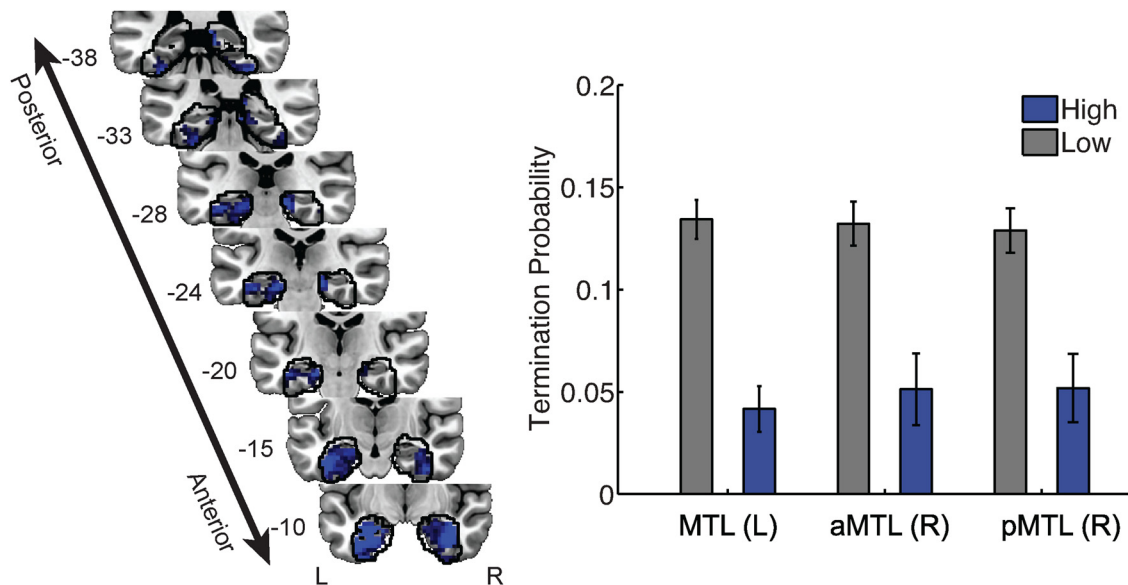


Figure 4. Model evidence for retrieval success processes in the MTL. Left, A deviance map showing clusters of informative voxels ($p < 0.05$, corrected) within anatomically defined regions of interest, outlined in black. Right, Recall events were partitioned on the basis of neural activity in each MTL cluster; periods of high neural activity were associated with a decreased likelihood of recall termination. The y-values correspond to the mean probability (\pm SEM, averaged across participants) that the recall event in question is the final item in the recall sequence. The y-coordinates of each MTL section are given in MNI space. aMTL, anterior MTL; pMTL, posterior MTL.

likelihood that recall will terminate, as shown in Figure 4). These findings are generally consistent with theories in which item of information is reactivated in perirhinal cortex during memory retrieval. If perirhinal activity reflects the reactivation of successfully retrieved studied materials, increased perirhinal activity should indicate that the participant will continue to produce recall responses.

Visual inspection of the maps depicted in Figures 3 and 4 suggests that the neural signal in certain regions was informative for both models. To determine whether, for regions containing neural signal that was informative to both models, these two neurally controlled cognitive mechanisms explain independent sources of variability in the behavioral data, we constructed a joint neural model. In this joint model, a neural signal from a given voxel influenced both the temporal reinstatement and retrieval success mechanisms (with a separate neural scaling parameter for each process). This analysis identified a single cluster of voxels within the left MTL, composed of hippocampus, posterior PHG, and posterior temporal fusiform cortex, which allowed the joint model to perform better than the baseline model (Fig. 5a).

Each of the voxel time courses within this cluster was associated with a positive neural scaling factor for the temporal rein-

statement process and a positive neural scaling parameter for the recall success process. In other words, under this model, increased activity in this region reflects both the high-fidelity reactivation of temporal context, as well as a decreased likelihood that recall will terminate. As in our analyses of the TR and RS models, we partitioned the set of recall events into periods of high and low neural signal (Fig. 5b), using a weighted average of the voxels within the left MTL cluster. Figure 5c shows a clear increase in the temporal organization of recall events occurring during periods of increased MTL activity. Furthermore, we observed a clear decrease in the probability of recall termination for the set of recall events with high MTL activity (Fig. 5d). A bootstrap analysis confirmed that representative models constructed from this cluster favored the joint model over both the RS and TR models alone [mean \pm SEM weighted AIC (wAIC) = 0.68 ± 0.009 , $p < 0.0001$].

Model support varies along the anterior–posterior axis of the MTL

The maps presented in Figures 3 and 4 suggest that the relationship between neural signal and the temporal reinstatement process in the model is stronger in posterior MTL regions, while the

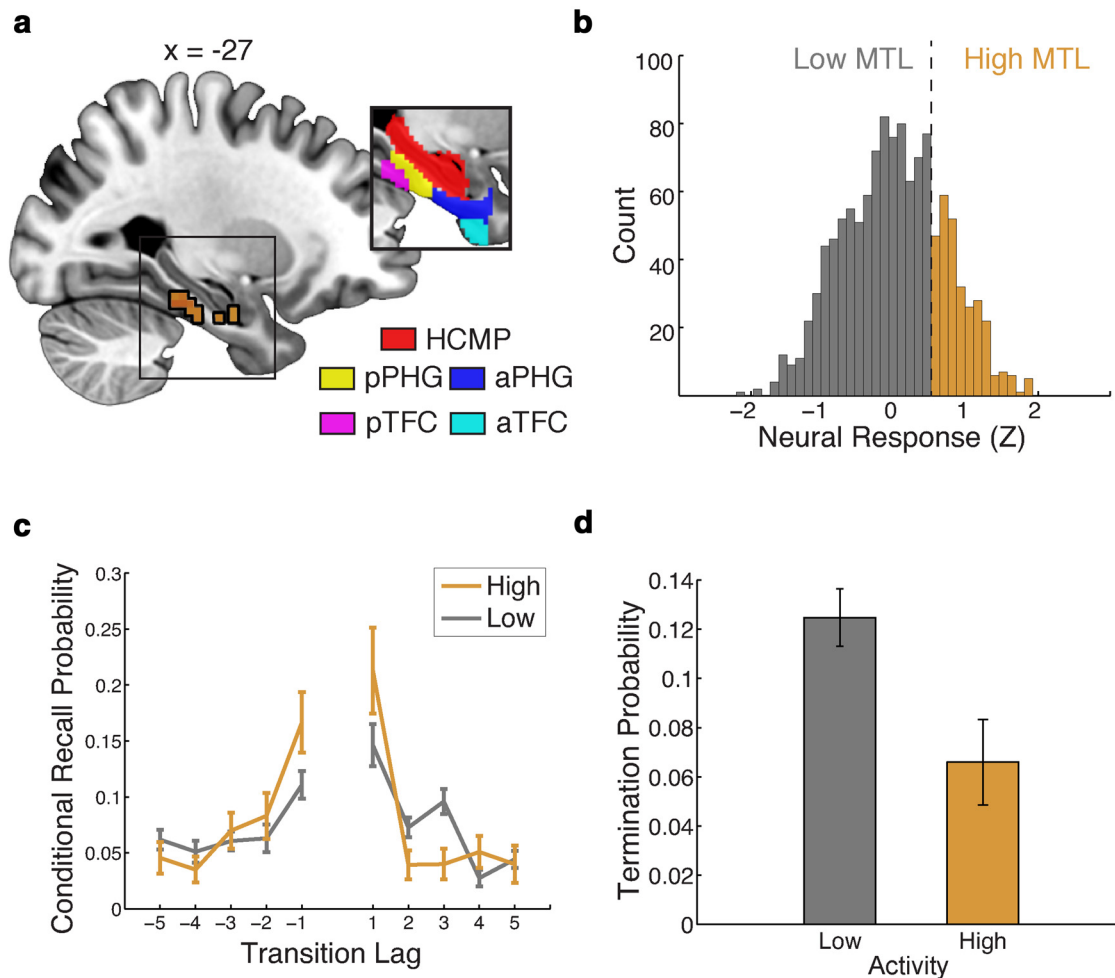


Figure 5. Cortical regions in the left MTL jointly support the temporal reinstatement and retrieval success hypotheses. **a**, A deviance map indicates a single cluster of voxels where the joint model is significantly better than the neurally naive baseline model, in a functionally defined region of interest, outlined in black (voxels identified as informative to both temporal reinstatement and retrieval success). Inset, A priori anatomical regions of interest. HCMP, Hippocampus; aPHG, anterior PHG; pPHG, posterior PHG; TFC, temporal fusiform cortex; aTFC, anterior TFC; pTFC, posterior TFC. The x -coordinates are given in MNI space. **b**, A histogram depicting the distribution of MTL neural responses across all recall events. Neural response is z -score normalized within a trial and is averaged across the voxels in the identified cluster. Responses >0.5 SDs above the mean response (across all trials) are labeled as high-activity recall events. **c**, Recall events were partitioned according to the level of activation in the cluster; periods of high activity were associated with increased temporal organization. **d**, Periods of high neural activity were also associated with a decreased likelihood of recall termination.

relationship between neural signal and retrieval success is stronger in anterior regions. To test whether evidence of a cognitive operation differed along the anterior–posterior axis of the MTL, we carried out a wAIC analysis (Burnham and Anderson, 2004). We computed four wAIC scores corresponding to the four models under consideration (baseline, TR, RS, and joint) for each voxel within anatomical ROIs constructed along the anterior–posterior axis of the MTL (Fig. 6a). For each ROI, we used a bootstrap sampling procedure to generate an estimated average wAIC score for each model. This sampling procedure was restricted to neurally informative voxels (i.e., those identified in the deviance maps of Figs. 3, 4, 5), answering the following question: if a neural signal in this region is informative for at least one of our neural models, which of the neural models is best supported?

The relative support for each of the three neural models is presented for hippocampal ROIs (Fig. 6b) and MTL cortical ROIs (Fig. 6c), with the ROIs arranged in terms of their position along the anterior–posterior axis of the region. Within the hippocampus, RS models were most supported in anterior ROIs; evidence for these models significantly increased in the anterior direction, as revealed by linear regression ($\beta = 6.79 \pm 2.28$, $p < 0.0001$).

This pattern also held across MTL cortical ROIs. Evidence for RS models significantly increased in the anterior direction ($\beta = 14.32 \pm 0.62$, $p < 0.0001$). Within the hippocampus, a significant linear trend in the opposite direction was observed for the TR model, with increasing support in posterior ROIs ($\beta = -11.09 \pm 3.97$, $p < 0.0001$). This pattern also held across MTL cortical ROIs. Evidence for the TR models significantly increased in the posterior direction ($\beta = -25.19 \pm 1.90$, $p < 0.0001$). Support for the joint model did not vary along the anterior–posterior axis in the hippocampus ($\beta = -13.7 \pm 7.55$, $p = 0.14$), although in MTL cortical ROIs there was a reliable increase in support for this model in the posterior direction ($\beta = -26.77 \pm 3.24$, $p < 0.0001$).

For two of the three posterior hippocampal ROIs, and all of the posterior MTL cortical ROIs, the joint model was strongly favored. While we observed an anatomical gradient to these signals, both processes (i.e., temporal reinstatement and retrieval success) were reflected in neural signal from posterior regions. These posterior MTL cortical ROIs spanned parahippocampal cortex and adjacent temporal fusiform cortex.

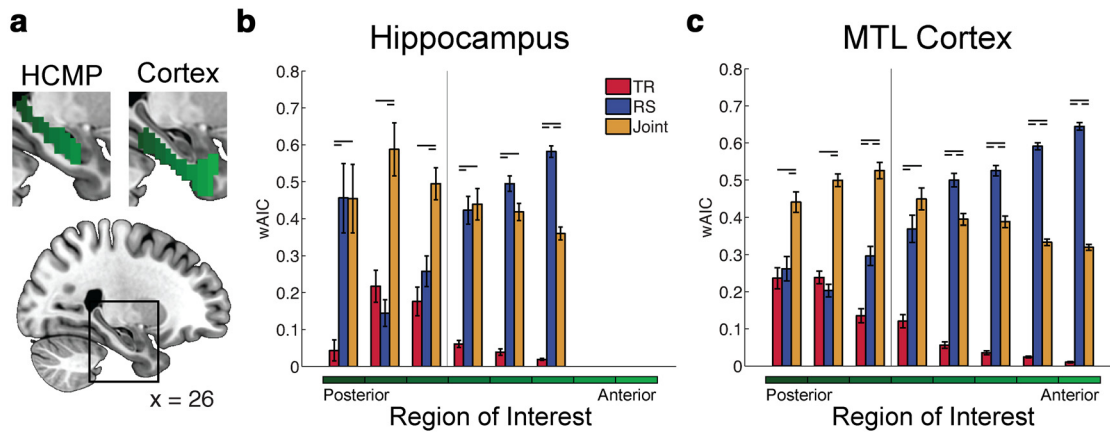


Figure 6. Relative fitness of neural models across MTL subregions. **a**, Anatomically defined ROIs along the anterior–posterior extent of the hippocampus (HCMP; left) and MTL cortex (right). **b**, Average wAIC scores from six anatomically defined hippocampal ROIs. **c**, Model fitness from eight anatomically defined MTL cortical ROIs. Bars indicate the bootstrap-estimated average wAIC scores for neurally informative models within that ROI. Error bars reflect the bootstrap-estimated SEM. Horizontal dashes denote significant pairwise differences ($p < 0.05$, Bonferroni corrected). The vertical line denotes the anatomical boundary between the perirhinal and parahippocampal cortices.

MTL signal is informative at the level of individual trials

As mentioned, our baseline model assumes that the parameters associated with temporal reinstatement and item retrieval are stationary over the recall period. This does not mean that the baseline model predicts that the behavioral markers associated with those parameters are stationary over the recall period. This is clearly indicated in Figure 2c. Despite the recall termination parameter ξ_d being stationary, the model captures the increasing probability of recall termination as a function of output position (due to the dynamics described in Eq. 9). Likewise, the baseline model captures a small decline in temporal organization with output position. According to the baseline model, temporal reinstatement is equivalent across these recall events, but the temporal organization associated with early recall events is amplified by the recency effect, the tendency for the first few recall events to come from the final few serial positions of the study list.

It is worth considering whether a neural signal could shift in such a way that it could be artifactually identified as informative to one of the models, despite not reflecting the cognitive operation embodied by the model. For example, if the baseline model did not perfectly capture the decline in temporal organization from early to later output positions, it is possible that a neural signal that generally decreased over the course of the recall period would be identified as neurally informative for the TR model, as it would allow the model to decrease the amount of temporal reinstatement in later output positions, relative to early output positions. Similarly, one can imagine an equivalent scenario for the RS model, as follows: a neural signal that generally changes its magnitude between early and late output positions could help the model fit the ever increasing tendency toward recall termination with later output positions. We developed a permutation-based analysis of these neurally informed models to address these concerns.

This permutation analysis was designed to rule out the possibility that our analytic framework identified a neural signal reflecting some general process that changes as a function of output position, but does not indicate temporal reinstatement or item retrieval per se. For each representative model associated with a given MTL cluster, we permuted the observed neural signal recorded for a given output position across trials. Permutations were constrained to occur within subjects, and ensuring the variability in BOLD response between subjects did not bias the results

of this analysis. If the observed neural signal contains information that reflects a general process that changes as a function of output position, a permuted signal should be as informative as trial-specific neural data. We observed significantly improved performance in the two representative TR models constructed from left ($p = 0.005$) and right ($p = 0.011$) MTL clusters, relative to models informed by a permuted neural signal. The same was true for the three representative RS models, which were constructed from the right anterior MTL cluster ($p < 0.001$), the left MTL cluster ($p < 0.001$), and the right posterior PHG cluster ($p = 0.009$). These results confirm that trial-level variability in neural signal improved the ability of the neurally informed models to predict recall behavior.

Discussion

To determine the cognitive functions of MTL subregions during memory search, we developed computational models that tested hypotheses regarding the links between complex cognitive operations and BOLD activity recorded during retrieval. We found evidence that distinct MTL structures are associated with different computational mechanisms predicting successful performance and temporal organization during a free-recall task. Our findings implicate the posterior PHG and hippocampus in reinstating the temporal context of prior episodes. This temporal reinstatement is thought to be an important mechanism supporting the subjective reliving of past experience (Tulving, 1993). In contrast, anterior MTL structures, including perirhinal cortex, were agnostic to the temporal structure of retrieved memories, while generally supporting recall of items from the targeted episode.

The hypothesis that particular MTL subregions control the fidelity of mental time travel was strongly supported by the data. Two clusters of voxels carried a signal whose recall-by-recall fluctuations improved the predictive power of the temporal reinstatement model relative to both the retrieval success model and a baseline model. Multiple MTL structures, including right hippocampus, bilateral parahippocampal cortex, and surrounding temporal fusiform cortex contributed to this model. When these regions showed increased activity for a particular recall event, there was an increased likelihood that the next recalled item would come from a nearby point in time (Fig. 3). These results are consistent with theories implicating parahippocampal cortex in the maintenance of contextual information (Davachi, 2006; Ran-

ganath and Ritchey, 2012), studies demonstrating recruitment of the parahippocampal cortex and hippocampus during the retrieval of temporal information (Ekstrom and Bookheimer, 2007), and studies implicating MTL structures in the representation of temporal information in a variety of memory tasks (Jenkins and Ranganath, 2010; Manning et al., 2011; Howard et al., 2012; Ezzayat and Davachi, 2014). These findings are in accord with the proposed role of the MTL in providing temporal structure to one's experience (Levy, 1996; Wallenstein et al., 1998; Howard et al., 2005; Eichenbaum, 2013).

It has been proposed that the hippocampus contains associative structures that are responsible for binding item and contextual information during the formation of episodic memories (McClelland et al., 1995; Howard et al., 2005; Diana et al., 2007; Polyn and Kahana, 2008). During retrieval, patterns of activity present during encoding are reinstated within cortex (Danker and Anderson, 2010). Studies using multivoxel pattern analysis (MVPA) have shed light on how reinstated patterns of neural activity reflect the attributes of memory (Polyn et al., 2005; Lewis-Peacock and Postle, 2008; Johnson et al., 2009), and hippocampal activity has been implicated in the success of this reinstatement process (Staresina et al., 2013; Bosch et al., 2014; Gordon et al., 2014). The analytic framework presented here raises the possibility of linking neural representations characterized with MVPA to the cognitive operations of retrieved-context models of memory search. For example, Polyn et al. (2012) identified patterns of neural activity during list learning that were sensitive to the encoding task used on a particular item. These task-sensitive neural representations were reinstated during memory search. One can imagine an extension to the current analytic framework in which the fidelity of the neural reinstatement of task information (measured with MVPA) is used to control a model parameter indicating how strongly the source characteristics of a given item are represented in the contextual representation (Polyn et al., 2009). As in the current work, support for this neural–cognitive linking hypothesis would take the form of improved behavioral predictions. In this case, an informative neural signal would improve the ability of the model to predict variability in source organization during memory search, providing information regarding the representation of source attributes in the MTL. This approach may also prove valuable in relating distributed neural representations of semantic structure (Haxby et al., 2001; Kriegeskorte et al., 2008) to the behavioral dynamics of memory search (Socher et al., 2009; Manning et al., 2012; Morton et al., 2013).

The hypothesis that MTL subregions indicate the success of memory search was also strongly supported by the data. Activity in the anterior extent of MTL cortex, including perirhinal cortex, demonstrated the strongest evidence for the retrieval success process during memory search. Fluctuations in these regions during memory search did not indicate whether recalled materials will be temporally organized; however, when activity in these regions declined, participants were more likely to terminate the search. This is consistent with prior work implicating perirhinal cortex in the associative retrieval of item details (Düzel et al., 2003; Kirwan and Stark, 2004; Law et al., 2005; Diana et al., 2010). One interpretation of these results is that anterior PHG activation during retrieval reflects the reactivation of item-specific information during memory search. This information could then be projected to executive and attentional systems to determine task relevance, and to linguistic systems for verbal report. Under this interpretation, activity in this region would indicate retrieval success (measured in terms of the verbal reports made by the participant), but

the level of engagement would not necessarily indicate the temporal organization of the recalled materials.

We observed the neural signal in a left MTL region, spanning hippocampus and adjacent entorhinal and parahippocampal cortices, that was informative for both the retrieval success and temporal reinstatement mechanisms of our computational framework. A joint model, in which both of these mechanisms were influenced by neural signal, was favored over the other models when using signal from this region (Fig. 5). One possibility is that these two cognitive processes produce independent neural signals, and hemodynamic activity in this region simply reflects a superposition of these signals. However, it is also possible that the neural circuitry of the posterior MTL performs a function that is important for both of these cognitive operations. Our findings linking the hippocampus to both computational mechanisms align with theories in which this region plays a domain-general role in the retrieval of associative information (McClelland et al., 1995; Diana et al., 2007). Electrophysiological studies with high temporal resolution support this idea, suggesting that associative information flows from the hippocampus to cortex during memory tasks (Naya and Suzuki, 2011; Staresina et al., 2012a). Future computational studies of MTL function may identify a single mechanism signaled by activity within posterior hippocampus and parahippocampal cortex, arbitrating between these possibilities.

We found that the computational mechanisms defined by the model were differentially engaged along the anterior–posterior axis of the MTL, highlighting the anatomical organization of functional signals within this region. Temporal reinstatement effects showed stronger engagement in the posterior hippocampus and PHG, and recall success effects showed stronger engagement in the anterior hippocampus and PHG (Fig. 6). A number of studies have examined how the functional properties of the MTL shift as one moves along this anterior–posterior axis. Neuroimaging studies of cued recall suggest that the representational content of retrieved information is important, with retrieval of visual object and scene information driving perirhinal and parahippocampal cortices, respectively (Staresina et al., 2011, 2012b, 2013). A shift in representational selectivity has also been observed in hippocampus; while posterior hippocampus is selectively engaged by scene information, anterior hippocampus seems to be agnostic to the representational content of studied material (LaRocque et al., 2013; Liang et al., 2013; Huffman and Stark, 2014). It has been suggested that the selective engagement of parahippocampal cortex for scene information indicates a more general role in contextual processing, with scenes engaging mechanisms involved in encoding spatial context (Davachi, 2006). This is consistent with a number of studies relating parahippocampal cortical activity to source context (Davachi et al., 2003), nonspatial contextual attributes (Aminoff et al., 2007), and temporal context (Turk-Browne et al., 2012; Hsieh et al., 2014). Our study provides converging evidence for the idea that posterior regions of the MTL (namely, posterior hippocampus and parahippocampal cortex) are involved in contextual processing, and our model raises the possibility that context-sensitive neural signals in this region are shaped by the cognitive mechanisms underlying temporal reinstatement.

The findings provided by our neurally informed computational model of memory search represent a significant advance in the ability to understand the contributions of the MTL to human memory. Directly relating neural signals to computational mechanisms has led to advances in a number of cognitive domains, including visual search (Purcell et al., 2012), reinforcement learn-

ing (Daw et al., 2011), object categorization (Nosofsky et al., 2012), and problem solving (Rosenberg-Lee et al., 2009). While a number of recent studies suggest the promise of this approach in the domain of episodic memory (Manning et al., 2011; Polyn et al., 2012; Turner et al., 2013; Polyn and Sederberg, 2014), our understanding of the cognitive processes engaged in memory search remains tentative. A number of prior studies have used GLM-based approaches to make inferences about the functional contributions of MTL regions during memory search (Long et al., 2010; Shapira-Lichter et al., 2012; Kragel and Polyn, 2013). These inferences rely on the logic of cognitive subtraction, which involves the partitioning of behavioral events into distinct groups and relies on the assumption that a particular cognitive operation is differentially engaged for the two groups. In and of itself, the GLM does not provide evidence regarding the plausibility of this differential engagement, or even whether the proposed cognitive operation is sufficient to account for the observed behavior. By incorporating well-specified cognitive operations directly into the analytic model, our approach overcomes this limitation. The ability of the model to account for the observed behavioral phenomena (Fig. 2) demonstrates the sufficiency of the cognitive operations used in the model. With the basic form of the model validated, improvement in the predictive power of a neurally informed model allows us to draw inferences regarding the cognitive processes associated with the observed neural signal. Using this approach, we have developed, to our knowledge, the first computational model using activation of the MTL to predict the organization of human memory, allowing us to make inferences about the mechanisms that are sufficient to give rise to mental time travel.

The neurally informed computational models presented here link particular regions of the MTL with the cognitive operations that enable an individual to search through memories of their past experience. Neural signal in the posterior MTL indicated the fidelity of mental time travel, consistent with the proposal that this region specifies the spatiotemporal context of a memory. A neural signal within multiple MTL structures, including the perirhinal cortex, reflected successful item retrieval. Neural signal in the hippocampus was linked to both mental time travel and retrieval success, consistent with the proposal that associative structures in this region support the reactivation of mnemonic representations in cortex. This approach demonstrates how computational models of cognition can be integrated with models of neural activation, providing a unified framework to test specific hypotheses regarding how fluctuations in the neural signal map onto the dynamics of thought.

References

- Aminoff E, Gronau N, Bar M (2007) The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb Cortex* 17:1493–1503. [CrossRef Medline](#)
- Amunts K, Kedo O, Kindler M, Pieperhoff P, Mohlberg H, Shah NJ, Habel U, Schneider F, Zilles K (2005) Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anat Embryol (Berl)* 210:343–352. [CrossRef Medline](#)
- Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26:839–851. [CrossRef Medline](#)
- Bar M, Aminoff E, Schacter DL (2008) Scenes unseen: the parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *J Neurosci* 28:8539–8544. [CrossRef Medline](#)
- Bosch SE, Jehee JF, Fernández G, Doeller CF (2014) Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *J Neurosci* 34:7493–7500. [CrossRef Medline](#)
- Burgess N, Maguire EA, O'Keefe J (2002) The human hippocampus and spatial and episodic memory. *Neuron* 35:625–641. [CrossRef Medline](#)
- Burnham KP, Anderson DR (2004) Multimodel inference understanding AIC and BIC in model selection. *Sociol Methods Res* 33:261–304. [CrossRef](#)
- Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109–114. [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](#)
- Davachi L (2006) Item, context and relational episodic encoding in humans. *Curr Opin Neurobiol* 16:693–700. [CrossRef Medline](#)
- Davachi L, Mitchell JP, Wagner AD (2003) Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci U S A* 100:2157–2162. [CrossRef Medline](#)
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215. [CrossRef Medline](#)
- Diana RA, Yonelinas AP, Ranganath C (2007) Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends Cogn Sci* 11:379–386. [CrossRef Medline](#)
- Diana RA, Yonelinas AP, Ranganath C (2010) Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *J Cogn Neurosci* 22:1808–1818. [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](#)
- Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp 39–43. Piscataway, NJ: IEEE.
- Eichenbaum H (2013) Memory on time. *Trends Cogn Sci* 17:81–88. [CrossRef Medline](#)
- Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25:1325–1335. [CrossRef Medline](#)
- Ekstrom AD, Bookheimer SY (2007) Spatial and temporal episodic memory retrieval recruit dissociable functional networks in the human brain. *Learn Mem* 14:645–654. [CrossRef Medline](#)
- Ezzyat Y, Davachi L (2014) Similarity breeds proximity: pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron* 81:1179–1189. [CrossRef Medline](#)
- Gordon AM, Rissman J, Kiani R, Wagner AD (2014) Cortical reinstatement mediates the relationship between content-specific encoding activity and subsequent recollection decisions. *Cereb Cortex* 24:3350–3364. [CrossRef Medline](#)
- Haskins AL, Yonelinas AP, Quamme JR, Ranganath C (2008) Perirhinal cortex supports encoding and familiarity-based recognition of novel associations. *Neuron* 59:554–560. [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](#)
- Hayes SM, Nadel L, Ryan L (2007) The effect of scene context on episodic object recognition: parahippocampal cortex mediates memory encoding and retrieval success. *Hippocampus* 17:873–889. [CrossRef Medline](#)
- Howard MW, Kahana MJ (2002) A distributed representation of temporal context. *J Math Psychol* 46:269–299. [CrossRef](#)
- Howard MW, Fotedar MS, Datey AV, Hasselmo ME (2005) The temporal context model in spatial navigation and relational learning: toward a common explanation of medial temporal lobe function across domains. *Psychol Rev* 112:75–116. [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](#)
- Hsieh LT, Gruber MJ, Jenkins LJ, Ranganath C (2014) Hippocampal activity patterns carry information about objects in temporal context. *Neuron* 81:1165–1178. [CrossRef Medline](#)
- Huffman DJ, Stark CE (2014) Multivariate pattern analysis of the human medial temporal lobe revealed representationally categorical cortex and representationally agnostic hippocampus. *Hippocampus* 24:1394–1403. [CrossRef Medline](#)
- Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitkänen A (1998) MR volumetric analysis of the human

- entorhinal, perirhinal, and temporopolar cortices. *Am J Neuroradiol* 19:659–671. [Medline](#)
- Jenkins LJ, Ranganath C (2010) Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *J Neurosci* 30:15558–15565. [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](#)
- Kahana MJ (1996) Associative retrieval processes in free recall. *Mem Cognit* 24:103–109. [CrossRef Medline](#)
- Kahana MJ (2012) *Foundations of human memory*, Ed 1. New York: Oxford UP.
- Kahana MJ, Howard MW, Polyn SM (2008) Associative retrieval processes in episodic memory In: *Cognitive psychology of memory*, Vol 2, *Learning and memory: a comprehensive reference* (Roediger IHL, ed), pp 467–490. Oxford: Elsevier.
- Kirwan CB, Stark CE (2004) Medial temporal lobe activation during encoding and retrieval of novel facename pairs. *Hippocampus* 14:919–930. [CrossRef Medline](#)
- Kragel JE, Polyn SM (2013) Functional interactions between large-scale networks during memory search. *Cereb Cortex*. Advance online publication. Retrieved January 8, 2014. doi:10.1093/cercor/bht258. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60:1126–1141. [CrossRef Medline](#)
- Laming D (1999) Testing the idea of distinct storage mechanisms in memory. *Int J Psychol* 34:419–426. [CrossRef](#)
- LaRocque KF, Smith ME, Carr VA, Withoft N, Grill-Spector K, Wagner AD (2013) Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. *J Neurosci* 33:5466–5474. [CrossRef Medline](#)
- Law JR, Flanery MA, Wirth S, Yanike M, Smith AC, Frank LM, Suzuki WA, Brown EN, Stark CE (2005) Functional magnetic resonance imaging activity during the gradual acquisition and expression of paired-associate memory. *J Neurosci* 25:5720–5729. [CrossRef Medline](#)
- Lehn H, Steffenach HA, van Strien NM, Veltman DJ, Witter MP, Håberg AK (2009) A specific role of the human hippocampus in recall of temporal sequences. *J Neurosci* 29:3475–3484. [CrossRef Medline](#)
- Levy WB (1996) A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus* 6:579–590. [CrossRef Medline](#)
- Lewis-Peacock JA, Postle BR (2008) Temporary activation of long-term memory supports working memory. *J Neurosci* 28:8765–8771. [CrossRef Medline](#)
- Liang JC, Wagner AD, Preston AR (2013) Content representation in the human medial temporal lobe. *Cereb Cortex* 23:80–96. [CrossRef Medline](#)
- Litman L, Awipi T, Davachi L (2009) Category-specificity in the human medial temporal lobe cortex. *Hippocampus* 19:308–319. [CrossRef Medline](#)
- Long NM, Oztekin I, Badre D (2010) Separable prefrontal cortex contributions to free recall. *J Neurosci* 30:10967–10976. [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](#)
- Manning JR, Sperling MR, Sharan A, Rosenberg EA, Kahana MJ (2012) Spontaneously reactivated patterns in frontal and temporal lobe predict semantic clustering during memory search. *J Neurosci* 32:8871–8878. [CrossRef Medline](#)
- Mayes A, Montaldi D, Migo E (2007) Associative memory and the medial temporal lobes. *Trends Cogn Sci* 11:126–135. [CrossRef Medline](#)
- McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* 102:419–457. [CrossRef Medline](#)
- Miller JF, Weidemann CT, Kahana MJ (2012) Recall termination in free recall. *Mem Cognit* 40:540–550. [CrossRef](#)
- Milner B, Squire LR, Kandel ER (1998) *Cognitive neuroscience and the study of memory*. *Neuron* 20:445–468. [CrossRef Medline](#)
- Morton NW, Kahana MJ, Rosenberg EA, Baltuch GH, Litt B, Sharan AD, Sperling MR, Polyn SM (2013) Category-specific neural oscillations predict recall organization during memory search. *Cereb Cortex* 23:2407–2422. [CrossRef Medline](#)
- Naya Y, Suzuki WA (2011) Integrating what and when across the primate medial temporal lobe. *Science* 333:773–776. [CrossRef Medline](#)
- Nosofsky RM, Little DR, James TW (2012) Activation in the neural network responsible for categorization and recognition reflects parameter changes. *Proc Natl Acad Sci U S A* 109:333–338. [CrossRef Medline](#)
- Polyn SM, Kahana MJ (2008) Memory search and the neural representation of context. *Trends Cogn Sci* 12:24–30. [CrossRef Medline](#)
- Polyn SM, Sederberg PB (2014) Brain rhythms in mental time travel. *Neuroimage* 85:678–684. [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](#)
- Polyn SM, Norman KA, Kahana MJ (2009) A context maintenance and retrieval model of organizational processes in free recall. *Psychol Rev* 116:129–156. [CrossRef Medline](#)
- Polyn SM, Kragel JE, Morton NW, McCluey JD, Cohen ZD (2012) The neural dynamics of task context in free recall. *Neuropsychologia* 50:447–457. [CrossRef Medline](#)
- Purcell BA, Schall JD, Logan GD, Palmeri TJ (2012) From salience to saccades: multiple-alternative gated stochastic accumulator model of visual search. *J Neurosci* 32:3433–3446. [CrossRef Medline](#)
- Ranganath C, Ritchey M (2012) Two cortical systems for memory-guided behaviour. *Nat Rev Neurosci* 13:713–726. [CrossRef Medline](#)
- Rosenberg-Lee M, Lovett MC, Anderson JR (2009) Neural correlates of arithmetic calculation strategies. *Cogn Affect Behav Neurosci* 9:270–285. [CrossRef Medline](#)
- Sederberg PB, Howard MW, Kahana MJ (2008) A context-based theory of recency and contiguity in free recall. *Psychol Rev* 115:893–912. [CrossRef Medline](#)
- Shapira-Lichter I, Vakil E, Glikmann-Johnston Y, Siman-Tov T, Caspi D, Paran D, Hendler T (2012) Inside out: a neuro-behavioral signature of free recall dynamics. *Neuropsychologia* 50:2245–2256. [CrossRef Medline](#)
- Socher R, Gershman SJ, Perotte AJ, Sederberg PB, Blei DM, Norman KA (2009) A Bayesian analysis of dynamics in free recall. In: *Advances in neural information processing systems* (Bengio Y, Schuurmans D, Lafferty J, Williams CKI, Culotta A, eds). Cambridge, MA: MIT.
- Solway A, Geller AS, Sederberg PB, Kahana MJ (2010) Pyparse: a semiautomated system for scoring spoken recall data. *Behav Res Methods* 42:141–147. [CrossRef Medline](#)
- Staresina BP, Davachi L (2010) Object unitization and associative memory formation are supported by distinct brain regions. *J Neurosci* 30:9890–9897. [CrossRef Medline](#)
- Staresina BP, Duncan KD, Davachi L (2011) Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. *J Neurosci* 31:8739–8747. [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](#)
- Staresina BP, Cooper E, Henson RN (2013) Reversible information flow across the medial temporal lobe: the hippocampus links cortical modules during memory retrieval. *J Neurosci* 33:14184–14192. [CrossRef Medline](#)
- Tulving E (1993) What is episodic memory? *Curr Dir Psychol Sci* 2:67–70. [CrossRef](#)
- Turk-Browne NB, Simon MG, Sederberg PB (2012) Scene representations in parahippocampal cortex depend on temporal context. *J Neurosci* 32:7202–7207. [CrossRef Medline](#)
- Turner BM, Forstmann BU, Wagenmakers EJ, Brown SD, Sederberg PB, Steyvers M (2013) A Bayesian framework for simultaneously modeling neural and behavioral data. *Neuroimage* 72:193–206. [CrossRef Medline](#)
- Wagenmakers EJ, Farrell S (2004) AIC model selection using akaike weights. *Psychon Bull Rev* 11:192–196. [CrossRef Medline](#)
- Wallenstein GV, Eichenbaum H, Hasselmo ME (1998) The hippocampus as an associator of discontinuous events. *Trends Neurosci* 21:317–323. [CrossRef Medline](#)
- Wilks SS (1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann Math Stat* 9:60–62. [CrossRef](#)
- Wixted JT, Squire LR (2011) The medial temporal lobe and the attributes of memory. *Trends Cogn Sci* 15:210–217. [CrossRef Medline](#)