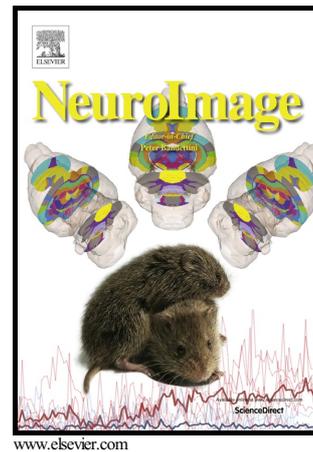


Author's Accepted Manuscript

Similar patterns of neural activity predict memory function during encoding and retrieval

James E. Kragel, Youssef Ezzyat, Michael R. Sperling, Richard Gorniak, Gregory A. Worrell, Brent M. Berry, Cory Inman, Jui-Jui Lin, Kathryn A. Davis, Sandhitsu R. Das, Joel M. Stein, Barbara C. Jobst, Kareem A. Zaghloul, Sameer A. Sheth, Daniel S. Rizzuto, Michael J. Kahana



PII: S1053-8119(17)30254-9
DOI: <http://dx.doi.org/10.1016/j.neuroimage.2017.03.042>
Reference: YNIMG13917

To appear in: *NeuroImage*

Received date: 18 November 2016
Revised date: 22 February 2017
Accepted date: 20 March 2017

Cite this article as: James E. Kragel, Youssef Ezzyat, Michael R. Sperling, Richard Gorniak, Gregory A. Worrell, Brent M. Berry, Cory Inman, Jui-Jui Lin, Kathryn A. Davis, Sandhitsu R. Das, Joel M. Stein, Barbara C. Jobst, Kareem A. Zaghloul, Sameer A. Sheth, Daniel S. Rizzuto and Michael J. Kahana, Similar patterns of neural activity predict memory function during encoding and retrieval *NeuroImage*, <http://dx.doi.org/10.1016/j.neuroimage.2017.03.042>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Similar patterns of neural activity predict memory function during encoding and retrieval

James E. Kragel,¹ Youssef Ezzyat,¹ Michael R. Sperling,² Richard Gorniak,³ Gregory A. Worrell,⁴ Brent M. Berry,⁴ Cory Inman,⁵ Jui-Jui Lin,⁶ Kathryn A. Davis,⁷ Sandhitsu R. Das,⁷ Joel M. Stein,⁸ Barbara C. Jobst,⁹ Kareem A. Zaghoul,¹⁰ Sameer A. Sheth,¹¹ Daniel S. Rizzuto,^{1*} Michael J. Kahana^{1*}

¹Department of Psychology, University of Pennsylvania, Philadelphia PA 19104, USA

²Department of Neurology, ³Department of Radiology, Thomas Jefferson University Hospital, Philadelphia PA 19107, USA

⁴Department of Neurology, Mayo Clinic, Rochester MN 55905, USA

⁵Department of Neurosurgery, Emory School of Medicine, Atlanta GA 30322, USA

⁶Department of Neurosurgery, University of Texas Southwestern, Dallas TX 75390, USA

⁷Department of Neurology, ⁸Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia PA 19104, USA

⁹Department of Neurology, Dartmouth Medical Center, Lebanon NH 03756, USA

¹⁰Surgical Neurology Branch, National Institutes of Health, Bethesda MD 20814, USA

¹¹Department of Neurosurgery, Columbia University Medical Center, New York NY 10032, USA

*These authors contributed equally to this work.

Running title: Memory Encoding and Retrieval

Corresponding Author:

Michael J. Kahana

University of Pennsylvania, Department of Psychology
Suite 263

Stephen A. Levin Building

Philadelphia, PA 19104

e-mail: kahana@psych.upenn.edu

phone: (215) 746-3501

fax: (215) 746-6848

Conflict of Interest: The authors declare no competing financial interests.

Acknowledgments: We thank Blackrock Microsystems for providing neural recording equipment. This work was supported by the DARPA Restoring Active Memory (RAM) program (Cooperative Agreement N66001-14-2-4032), as well as National Institutes of Health grant MH55687. We are indebted to all patients who have selflessly volunteered their time to participate in our study. The views, opinions, and/or findings contained in this material are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

1 **Abstract**

2 Neural networks that span the medial temporal lobe (MTL), prefrontal cortex, and posterior cortical regions are
3 essential to episodic memory function in humans. Encoding and retrieval are supported by the engagement of both
4 distinct neural pathways across the cortex and common structures within the medial temporal lobes. However, the
5 degree to which memory performance can be determined by neural processing that is common to encoding and retrieval
6 remains to be determined. To identify neural signatures of successful memory function, we administered a delayed
7 free-recall task to 187 neurosurgical patients implanted with subdural or intraparenchymal depth electrodes. We
8 developed multivariate classifiers to identify patterns of spectral power across the brain that independently predicted
9 successful episodic encoding and retrieval. During encoding and retrieval, patterns of increased high frequency activity
10 in prefrontal, MTL, and inferior parietal cortices, accompanied by widespread decreases in low frequency power
11 across the brain predicted successful memory function. Using a cross-decoding approach, we demonstrate the ability
12 to predict memory function across distinct phases of the free-recall task. Furthermore, we demonstrate that classifiers
13 that combine information from both encoding and retrieval states can outperform task-independent models. These
14 findings suggest that the engagement of a core memory network during either encoding or retrieval shapes the ability
15 to remember the past, despite distinct neural interactions that facilitate encoding and retrieval.

16 **Keywords:** iEEG, MVPA, free recall, episodic memory

17 Introduction

18 Episodic memory requires both encoding operations that translate experiences into durable memories and retrieval
19 operations that reactivate memories when prompted by a retrieval cue (Polyn and Kahana, 2008). Many neurocognitive
20 models of episodic memory account for these processes using similar architectures, with distinct neural systems
21 facilitating memory encoding and retrieval (Nyberg et al., 1996; Lepage et al., 1998; Kim, 2015). Neuroimaging
22 studies provide evidence for these models, revealing dissociations between the neural correlates of episodic encoding
23 and retrieval within the hippocampus (Zeineh et al., 2003; Eldridge et al., 2005), as well as functional networks
24 spanning prefrontal, medial temporal, lateral temporal, and parietal cortical regions (Daselaar et al., 2009; Kim et al.,
25 2010). In contrast to neural substrates that are specific to encoding and retrieval, structures within the hippocampal
26 formation (Small et al., 2001; Stark and Okado, 2003; Prince et al., 2005) in addition to widespread cortical regions
27 (Sederberg et al., 2007) contribute to both episodic encoding and retrieval. Understanding the relative contributions
28 of neural processes engaged during encoding and retrieval is critical to understanding the mechanisms that ultimately
29 determine whether information can be successfully retrieved at a later point in time. The goal of the present work is to
30 determine whether memory performance is best predicted by distinct neural processes that occur during the encoding
31 and retrieval of episodic information, or common electrophysiological states that may reflect the engagement of core
32 mnemonic processes.

33 Theoretical models of episodic memory propose that successful memory retrieval involves the reinstatement of representations
34 that were present during the initial experience of an event (McClelland et al., 1995; O'Reilly and Rudy, 2001).
35 Using multivariate pattern analysis, it has been repeatedly demonstrated that patterns of cortical activation that are
36 present during the encoding of an event are recapitulated during retrieval (Staresina et al., 2012; Ritchey et al., 2013;
37 Wing et al., 2014). While considerable progress has been made linking the reinstatement of neural activity to the
38 contents of retrieval (for a review, see Danker and Anderson 2010), the neural states that facilitate the encoding and
39 reactivation of this content may differ. Multivariate classification techniques allow for the estimation of latent cognitive
40 states, such as attentive states that improve task performance (Rosenberg et al., 2015), and incorporate information
41 from neural populations across the brain. By training a multivariate classifier to identify neural signals that predict
42 successful memory function, we can determine the relationship between the processes that mediate episodic encoding

43 and retrieval.

44 If the success of episodic memory is determined by the recapitulation of the same processes that were engaged
45 during encoding (Kolers, 1973; Tulving and Thompson, 1973; Kolers and Roediger, 1984), one would predict that
46 engagement of neural systems that are generally predictive of memory success (e.g., the hippocampus; Prince et al.
47 2005) should be able to identify when either successful encoding or retrieval operations occur. In contrast, if the
48 function of neural systems that are specific to encoding or retrieval determine the fate of studied information, then
49 classifiers trained to predict the success of memory formation should fail to identify when successful retrieval occurs.
50 Recent electrophysiological investigations of human memory have demonstrated that shifts in the power spectrum,
51 specifically increased high frequency activity (HFA) with concomitant decreases in low frequency activity (LFA)
52 provide informative features for predicting successful memory formation (Burke et al., 2014a; Long et al., 2014;
53 Greenberg et al., 2015) and retrieval (Burke et al., 2014b). These measures reflect a combination of oscillatory dynamics
54 and broadband changes in HFA indexing the firing rate of local neuronal populations (Manning et al., 2009; Miller
55 et al., 2014; Burke et al., 2015). Incorporating these measures of neural activity into multivariate classifiers enables the
56 development of biomarkers of successful memory function, based on the dynamics of distributed neuronal populations.

57 To determine whether the same neural mechanisms support the encoding and retrieval of episodic memories and to
58 determine the degree to which these measures predict memory performance, we examined intracranial electroencephalographic
59 (iEEG) data from 187 patients who performed a free-recall task while they were undergoing clinical monitoring for
60 the surgical treatment of their drug-resistant epilepsy. To classify successful encoding and retrieval states, we fit
61 L2-penalized regression models to spectral activity recorded during the encoding of verbal stimuli, or in the moments
62 preceding successful recall. Using a cross-classification approach, we tested whether the same changes in spectral
63 power predict successful memory formation and retrieval. Finally, we estimated general models of episodic memory
64 function by constructing a joint classifier that incorporated information from both encoding and retrieval phases of
65 the task. These analyses aimed to uncover the degree to which latent mnemonic states can be measured from brain
66 activity, and the degree to which variation in patterns of spectral power common to both encoding and retrieval can
67 predict memory function.

68 **Materials and Methods**

69 *Participants*

70 187 patients with medication-resistant epilepsy underwent neurosurgical procedures to implant intracranial electrodes
71 (subdural, depth, or both) to determine epileptogenic regions. Data reported were collected at Dartmouth Medical
72 Center (Hanover, NH), Emory University Hospital (Atlanta, Georgia), Hospital of the University of Pennsylvania
73 (Philadelphia, PA), Mayo Clinic (Rochester, MN), Thomas Jefferson University Hospital (Philadelphia, PA), Columbia
74 University Medical Center (New York, NY), and University of Texas Southwestern Medical Center (Dallas, TX). Prior
75 to data collection, the research protocol was approved by the institutional review board at each hospital. Informed
76 consent was obtained from either the participant or their guardians. Hemisphere dominance was determined by
77 assessing handedness, Wada test (Wada and Rasmussen, 1960), or fMRI data collected during a verb generation
78 task. Previous publications utilizing a subset of these data (88 patients) have characterized the electrophysiological
79 signatures of successful memory encoding and retrieval (Burke et al., 2014a,b; Long et al., 2014); however, the present
80 analyses describing the relationship of successful encoding and retrieval are novel. Deidentified raw data and analysis
81 code used in the present analyses is available at http://memory.psych.upenn.edu/Electrophysiological_
82 [Data.](http://memory.psych.upenn.edu/Electrophysiological_)

83 *Free-recall task*

84 Each subject performed a variant of the delayed free-recall task in which they studied a list of words with the intention
85 to commit the items to memory. The task was performed at the bedside on a laptop, using PyEPL software (Geller
86 et al., 2007). Analog pulses were sent to available recording channels to enable alignment of experimental events (e.g.,
87 stimulus presentation) with the recorded iEEG signal.

88 The recall task consisted of three distinct phases: encoding, delay, and retrieval. During encoding, lists of 12 words
89 were presented in the native language (either English or Spanish) of the subject. Words were selected at random,
90 without replacement, from a pool of nouns (<http://memory.psych.upenn.edu/WordPools>). Word presentation
91 lasted for a duration of 1600 ms, followed by a blank inter-stimulus interval (ISI) of 750 to 1000 ms (see Fig 1a).
92 Presentation of word lists was followed by a 20 s post-encoding delay. Subjects performed an arithmetic task during

93 the delay in order to disrupt memory for end-of-list items. Math problems of the form $A+B+C=??$ were presented
94 to the participant, with values of A, B, and C were set to random single digit integers. Responses were made on
95 a keypad, with presentation of additional math problems following each response (i.e., a self-paced task). After the
96 delay, a row of asterisks, accompanied by an 800 Hz auditory tone, was presented for a duration of 300 ms to signal the
97 start of the recall period. Subjects were instructed to recall as many words as possible from the most recent list, in any
98 order during the 30 s recall period. Vocal responses were digitally recorded and parsed offline using Penn TotalRecall
99 (<http://memory.psych.upenn.edu/TotalRecall>). Subjects performed up to 25 recall trials in a single
100 recall session. Across sessions, subjects performed an average of 38 trials in total (range 7 – 150).

101 A subset of patients ($n = 88$) performed a variant (hereafter Experiment 2) of the previously described task. List
102 presentation consisted of a total of 15 items. In addition, a green fixation cross served as a list-cue to signal an
103 upcoming list of words. The list-cue was presented for a duration of 1600 ms, followed by the presentation of a blank
104 screen for 800 to 1200 ms. The ISI in this variant of the task lasted from 800 to 1200 ms in duration. The recall period
105 for this version of the task was 45 s in length.

106 *Electrophysiological recordings and data processing*

107 iEEG signal was recorded using subdural grids and strips (contacts spaced 10 mm apart) or depth electrodes (contacts
108 spaced 5-10 mm apart) using recording systems at each clinical site. iEEG systems included DeltaMed & XITek
109 (Natus), Grass Telefactor, and Nihon-Kohden EEG systems. Signals were sampled at 500, 512, 1000, 1024, or 2000
110 Hz, depending on hardware restrictions and considerations of clinical application. Signals recorded at individual
111 electrodes were converted to a bipolar montage by computing the difference in signal between adjacent electrode pairs
112 on each strip, grid, and depth electrode (Burke et al., 2013). Bipolar signal was notch filtered at 60 Hz with a fourth
113 order 2 Hz stop-band butterworth notch filter in order to remove the effects of line noise on the iEEG signal. Electrodes
114 determined to be within the epileptogenic zone were excluded from analysis.

115 *Anatomical Localization*

116 Anatomical localization of electrode placement was accomplished using independent processing pipelines for depth
117 and surface electrode localization. For data collected as a part of Experiment 1, post-implant CT images were
118 coregistered with presurgical T1 and T2 weighted structural scans with Advanced Normalization Tools (Avants et al.,
119 2008). For patients with MTL depth electrodes, hippocampal subfields and MTL cortices were automatically labeled
120 in a pre-implant, T2-weighted MRI using the automatic segmentation of hippocampal subfields (ASHS) multi-atlas
121 segmentation method (Yushkevich et al., 2015). Subdural electrodes were localized by reconstructing whole-brain
122 cortical surfaces from pre-implant T1-weighted MRIs using Freesurfer (Fischl et al., 2004), and snapping electrode
123 centroids to the cortical surface using an energy minimization algorithm (Dykstra et al., 2012). The localization for
124 data collected in Experiment 2 differed by using FMRIB's linear image registration tool (Jenkinson et al., 2002) for
125 coregistration of CT and structural scans. In addition, MTL depth electrodes that were visible on CT scans were
126 localized to either the hippocampus or PHG by neuroradiologists with expertise in MTL anatomy.

127 Each subject's T1-weighted MRI was additionally registered to both an average T1 constructed from a sample of
128 101 patients, facilitating group-level comparisons of subdural electrodes on the cortical surface. Registration to MNI
129 space further enabled group analysis of subdural and depth electrodes near MTL structures. Using this approach, we
130 achieved extensive coverage across the cortical surface (Fig. 1b, left), and the MTL (Fig. 1b, right).

131 In order to rule out the possibility that classification of successful recall, which involves the vocalization of a verbal
132 response, was being informed by neural processing in cortical regions involved in the planning and production of
133 a vocal response, we constructed classifiers using an anatomically restricted set of neural features (Fig. 1b), by
134 excluding electrodes localized within inferior frontal gyrus (IFG) pars triangularis, IFG pars orbitalis, precentral gyrus,
135 postcentral gyrus, paracentral gyrus, superior temporal gyrus, the bank of the superior temporal sulcus, transverse
136 temporal gyrus, and supramarginal gyrus as defined in the Desikan-Kiliany atlas (Desikan et al., 2006).

137 *Spectral power*

138 To compute spectral power during word encoding, we applied the Morlet wavelet transform (wave number 5) to all
139 bipolar electrode EEG signals from the onset to the offset of stimulus presentation, across 50 logarithmically spaced

140 frequencies from 3 to 180 Hz. Spectral power during recall was estimated from 500 ms preceding the onset of response
141 vocalization, for correct recalls and recall errors alike. We also computed spectral power during unsuccessful periods
142 of memory search, defined as 500 ms epochs in which the onset of any vocalization did not occur in the following
143 2000 ms. All events within the recall period (i.e., correct recalls and unsuccessful search) were required to be free
144 of vocalization onsets in the preceding 2000 ms, to account for potential differences in response production across
145 event types. Power estimates were log transformed and down sampled to 50 Hz. To avoid edge artifacts, we included
146 buffers of 1000 ms surrounding events of interest during the computation of spectral power. For multivariate decoding
147 analyses, features (i.e., log-transformed power from a specific electrode and frequency band) were standardized using
148 the mean and standard deviation estimated from all training samples, for each individual session. Test data were
149 normalized using session means and standard deviations estimated from the training data.

150 By focusing on patterns of neural activity in the moments preceding individual retrieval events, we excluded recalls
151 that were preceded by the onset of any additional vocalization or the recall period within 2000 ms. This duration
152 was determined in order to allow estimation of low frequency power without contamination of signal from response
153 production earlier in the recall period. While this approach decreased power in the analysis of retrieval data, it allowed
154 for low frequency features, which have been previously demonstrated to predict successful recall (Burke et al., 2014b),
155 to inform classification of memory states. This procedure resulted in the removal of an average of 53.8% of recall
156 events per subject (range 6.3% to 92.7%), which had an inter-response time of 1198 ± 15.4 msec (mean \pm SEM).
157 As a result, some subjects (21 from Experiment 1, and 19 from Experiment 2) did not have a sufficient number of
158 recall events to perform classification. For a given analysis, each subject was required to have at least 20 observations
159 per condition. For the purposes of classification analyses, we treated unique montages (resulting from reimplants, or
160 changes in recording electrodes) of individual patients as a single subject. This approach resulted in 202 subjects for
161 analysis of encoding period data, and 162 subjects for analysis of retrieval period data.

162 *Pattern classification*

163 To identify neural features that were associated with latent memory states during the task, we trained L2-penalized
164 logistic regression classifiers to distinguish between spectral power associated with memory success and failure.
165 During encoding, power patterns were averaged across the 1600 ms stimulus duration at each bipolar electrode,

166 resulting in a set of features to train the encoding classifier. For retrieval analysis, features were constructed from
 167 the average spectral power in the 500 ms preceding response vocalization of either correct recalls or periods of
 168 unsuccessful memory search. While we sought to identify neural markers of successful retrieval by comparing
 169 successful and unsuccessful memory search epochs, these events are confounded by when they occur in the recall
 170 period. As a result, changes in neural processing that distinguish between these two periods (e.g., an indicator of
 171 decreased attention at the end of the recall period) may limit our ability to detect memory-related signals. To overcome
 172 this issue, we selected unsuccessful memory search events that were matched to the onset of individual recall events
 173 in separate lists. This resulted in an average of 69.2 ± 4.6 successful, and 169.9 ± 8.8 unsuccessful search events, per
 174 subject.

175 Classifier performance was evaluated using a cross-validation procedure. We implemented an n -fold cross-validation
 176 scheme, where n was equal to the number of completed lists. Data from one list was held out from training to test
 177 the ability of the classifier to generalize to novel data, and held out lists were matched across encoding and retrieval
 178 classifiers. The remaining $n-1$ lists were used to train the parameters of the classifier. We chose the value of the cost
 179 parameter, C , from one of 22 values, spaced from 10^{-4} to 10^6 . For each subject, C was determined by the selecting the
 180 value that maximized the performance of the classifier as assessed by computing the area under the receiver operating
 181 characteristic curve using list-based cross validation in the remaining sample of subjects. This approach enabled
 182 optimization of the cost parameter on an independent set of data. We trained the classifier weights by minimizing a
 183 loss function over \mathbf{w} for a set of n training events:

$$\min_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{w}}{2} + C \sum_{i=1}^n \log(1 + e^{-y_i \mathbf{w}^T \mathbf{x}_i}), \quad (1)$$

184 where \mathbf{x}_i is a set of features (i.e., the pattern of spectral power across all electrodes during encoding or retrieval),
 185 and y_i is the corresponding class label for each event. Class labels during encoding were defined as 1 for recalled
 186 items, and -1 for forgotten items. During retrieval, correct recalls were labeled 1 and unsuccessful search periods
 187 were labeled -1 . A bias term was included by appending a feature that was fixed to 1 for each observation. The
 188 loss function was weighted for each observation proportionally to the number of observations in each class, to prevent

189 class imbalance from influencing classification results. Additional classification models were constructed from both
190 encoding and retrieval data, scaling the relative influence of each task phase by weighting retrieval observations by a
191 scaling parameter, log spaced from 10^{-5} to 10^5 . We performed L2-logistic regression as implemented by the liblinear
192 package (Fan et al., 2008).

193 In addition to constructing separate classifiers to identify successful memory states during either the encoding or
194 retrieval periods of the task, we developed joint classifiers that were trained on observations from both encoding and
195 retrieval phases of the task. In addition to being weighted proportionally to the imbalance of positive and negative
196 classes, the loss function was weighted by a scaling parameter, w_{enc}/w_{rec} which controlled the relative contribution
197 of encoding and retrieval observations to training of the classifier. We examined generalization of this classifier to
198 both encoding and recall periods of the task over a range of 15 log-spaced values from $10^{-1.4}$ to $10^{1.4}$. This range of
199 parameter values allowed use parametrically module the influence of different task phases on memory performance.

200 To determine whether individual classifiers performed above chance, we first constructed the receiver operating
201 characteristic and computed the area under the curve (AUC) across left-out samples (i.e., across the test set). We
202 implemented a permutation test by constructing a null distribution of classifier performance by permuting the class
203 labels within the training data. Subjects with an observed AUC in the top 95% of the null distribution were considered
204 significant, corresponding to a one-tailed test. To assess group performance, we used a binomial test to determine
205 whether the proportion of subjects with significant classification would exceed the false positive rate of 0.05.

206 In order to rule out confounding effects of serial position and output position, we used linear mixed effects models to
207 determine whether classifier output varied as a function of memory success (i.e., subsequent memory status or recall
208 success). During encoding, we constructed a model with subject treated as a random effect, and memory status, list
209 position, and the interaction between the two treated as fixed effects. Significance was determined by constructing
210 nested models with an fixed effect term constrained to zero. After fitting the restricted model, a parametric bootstrap
211 ($n = 1000$) was performed to determine whether the variance explained by the additional parameter was greater than
212 fitting residual error with an additional parameter. We utilized a similar procedure to evaluate potential confounds in
213 retrieval classifiers, with the fixed effects of memory status and output position.

214 *Identifying neural features predictive of memory success*

215 In order to determine which spectral features (i.e., which frequencies and regions) facilitated discrimination between
 216 successful and unsuccessful memory states, we constructed a forward model of these latent states from our linear
 217 classifiers (Haufe et al., 2014). For each subject, we computed a set of activations, A , based on the learned coefficients
 218 in \mathbf{w} :

$$A_x = \Sigma_{\mathbf{x}} \mathbf{w} \Sigma_{\xi}^{-1}, \quad (2)$$

219 where $\Sigma_{\mathbf{x}}$ is the covariance matrix of the full set of features, with each feature standardized within each experimental
 220 session, and Σ_{ξ} is the covariance matrix of the output of the classifier for each predicted class. For each subject and
 221 frequency, we computed an estimate of the activation at each electrode. Features were aggregated into LFA (< 10 Hz)
 222 and HFA (> 60 Hz) by averaging estimated activity across the frequency dimension. As a preprocessing step to the
 223 estimation of forward models, individual events were scanned for spikes and discarded if the kurtosis value for any
 224 feature exceeded a threshold of 2.5.

225 We projected the associated activations of each subdural electrode to all vertices on the cortical surface within 10mm
 226 of the midpoint of electrode centroids in each bipolar pair. For electrodes localized within the MTL, statistical maps
 227 were constructed for each subject by assigning voxels within a 3mm radius of each bipolar midpoint a value equal
 228 to its corresponding activation (overlapping regions were assigned the average of all activation values within 3mm).
 229 Each subject's activation map was smoothed with a 4 mm gaussian filter before subsequent analysis.

230 Group inference was performed by nonparametric statistical testing (Maris and Oostenveld, 2007). For each test,
 231 we performed a one-sample t-test on estimated levels of activation across subjects. Resultant statistical maps were
 232 thresholded for statistical significance by constructing a null distribution obtained by random sign flipping ($n = 1000$),
 233 and identifying positive and negative thresholds at 2.5th and 97.5th percentiles of the null distribution, ensuring a false
 234 discovery rate of 0.05.

Results

Behavioral Results

To validate patient performance, we examined characteristic behavior on the free-recall task. Subjects recalled an average of 3.26 ± 0.10 (mean \pm SEM) items per list, yielding an overall $24.27 \pm 0.008\%$ of items recalled. By interleaving encoding and recall with a distracting arithmetic task, we observed primacy effects in both recall initiation and overall recall rates (Fig. 2a,b). These findings are consistent with delayed free recall (Glanzer and Cunitz, 1966), and suggest that the neural mechanisms associated with memory performance reflect the encoding and retrieval of long-term memories (Strange et al., 2002). An additional behavioral marker of episodic memory function is the temporal contiguity effect (Kahana, 1996; Sederberg et al., 2010), in which subjects recall items in clusters based on their temporal proximity in the study list. Subjects recalled items with temporal organization (Fig. 2c), indicated by an average temporal factor score of 0.66 ± 0.01 , which reflects significantly greater temporal clustering than chance levels of 0.5 ($t_{191} = 22.46$, $p < 10^{-10}$; see Polyn et al., 2009). These findings indicate normal task performance, despite relatively low levels of overall recall, consistent with impaired memory performance in epilepsy (Hermann et al., 2008).

Multiple factors contribute to performance on the recall task, including the ability to target and retrieve individual items from the previous list. When memory fails due to the inability to retrieve task-relevant information, studied items are omitted from recall. As shown in Figure 2d, the time to first recall was inversely related to the number of items yet to be recalled, and the amount of time between successive recalls increased exponentially in the recall sequences, consistent with search dynamics commonly observed in single-trial free recall (Murdock and Okada, 1970; Rohrler and Wixted, 1994). Of relevance for analysis of neural activity during memory search, inter-response times greater than 2 seconds were distributed throughout the majority of recall sequences, except early on in the recall sequence for subjects with a high rate of recall.

Memory failure during retrieval can also occur when items not studied on the previous list can be endorsed as targets, resulting in false recall. We observed false memories in the form of extra-list intrusions (ELIs), prior list intrusions (PLIs) and repetitions of previously recalled items. Recall errors were primarily ELIs, with an average rate of $0.88 \pm$

260 0.10 ELIs per list. PLIs were observed with less frequency, with an average rate of 0.57 ± 0.03 intrusions per list. The
261 remaining errors were repetitions, which we rarely observed, with an average of 0.17 ± 0.03 of these errors made per
262 list.

263 *Multivariate classification results*

264 In order to determine whether the same neural mechanisms support episodic memory encoding and retrieval, we
265 constructed subject-specific classifiers that were capable of predicting latent cognitive states from patterns of spectral
266 power across multiple electrodes and frequencies. First, we developed independent classifiers of memory success
267 during either the encoding or retrieval of individual items. We tested the degree to which these models generalized
268 across different phases of the task, allowing us to determine whether common neural features predict successful
269 encoding and retrieval. Finally, we constructed classification models that incorporated information from both the
270 encoding and retrieval phases of the task, allowing us to determine which electrophysiological features were specific
271 to either encoding or retrieval, or are common indicators of episodic memory function.

272 *Decoding of successful memory formation and retrieval*

273 We first constructed a classifier to identify patterns of spectral power associated with successful encoding. This
274 classifier provides an item level estimate of the probability that subsequent remembering will occur. To relate the
275 output of this classifier to trial-level variability in recall performance, we separated each subject's encoding events into
276 terciles based on classifier output. By computing the probability of recall for items encoded in each of these terciles,
277 we translated the output of the classifier to a measure of actual memory performance. We observed an $80.53 \pm 2.97\%$
278 difference in the proportion of items recalled in the upper, relative to the lower tercile of classifier output (Fig. 3c).
279 The difference in recall performance was evident across all list positions, although smaller in magnitude for items
280 encoded at the beginning of the list (Figure 3a). We observed significant classification ($p < 0.05$, permutation test)
281 in 168 out of 202 subjects (treating changes in electrode coverage as independent subjects), with an average AUC of
282 0.64 ± 0.005 . The number of subjects with significant classification performance was more than one would expect by
283 chance ($p < 10^{-10}$, binomial test).

284 As the probability of recalling an item decreased across the encoding interval, we conducted a control analysis to rule
285 out the possibility that processes that varied as a function of list position, but not memory specific processing, were
286 primarily informing the classifier. We constructed a linear mixed effects model, predicting the evidence of each item
287 being subsequently remembered (i.e., the exponential term in Equation 1) from both subsequent memory status and
288 the serial position of each item. We tested whether unique variance in predicted memory states (i.e., classifier output)
289 was associated with each factor by comparing fitness of the full model to models in which the effect of each factor
290 was restricted to zero. We observed a significant effect of list position ($\chi^2_1 = 122.59$, $p < 0.001$ bootstrap test) and
291 subsequent memory status ($\chi^2_1 = 188.43$, $p < 0.001$ bootstrap test). These findings confirm that the ability to identify
292 encoding periods with a high probability of encoding success was independent of simple serial position effects that
293 influence recall behavior.

294 Next, we developed classifiers to discriminate between patterns of neural activity during epochs of memory search
295 associated with successful and unsuccessful retrieval (i.e., periods of memory search that did produce any form of
296 recall response). The relationship between successful retrieval and changes in spectral power was observed across
297 different periods of memory search (Figure 3b), with the biggest difference observed later in the recall period. Of the
298 162 classifiers constructed to predict retrieval success for individual subjects, we observed above chance classification
299 ($p < 0.05$, permutation test) in 155, indicative of significant group-level performance ($p < 10^{-10}$, binomial test).
300 Classifier performance for held out lists of items was evaluated using AUC, with average of 0.83 ± 0.01 , indicating
301 robust performance. We tested whether variability in memory states estimated by the retrieval success classifier could
302 be predicted from the period in which an event occurred (early, middle, or late epochs defined by terciles of event
303 onsets within the search period) and whether retrieval was successful or not. We observed significant effects for both
304 search period ($\chi^2_1 = 14.32$, $p < 0.001$ bootstrap test) and retrieval status ($\chi^2_1 = 340.8$, $p < 0.001$ bootstrap test). These
305 findings establish that retrieval success classifiers are informative across the duration of the search period, and are
306 more sensitive as the retrieval period progresses. When sorting memory search epochs into terciles by classifier output
307 (Figure 3c), subjects were $189.3 \pm 5.0\%$ more likely to recall an item when in the upper, compared to the lower tercile
308 of classifier output.

309 These results suggest that memory related processing across cortical and subcortical sites can be used to reliably

310 classify when retrieval will occur during memory search; however, such classification may be informed by neural
311 features predictive of processes that occur subsequent to retrieval itself, including phonological retrieval and the
312 planning of a vocal response. In order to rule out the possibility that the ability to decode successful retrieval was not
313 primarily driven by difference in response-related activity, we constructed additional classifiers that excluded ROIs
314 associated with vocal response production and planning (Hermes et al., 2014). Performance of classifiers constructed
315 from this restricted feature space was observed to be significantly above chance ($p < 0.05$, permutation test) in 151
316 out of 162 subjects. Furthermore, we observed an average AUC of 0.79 ± 0.01 , demonstrating a significant reduction
317 from models with an unrestricted feature space ($t_{161} = -10.3, p < 10^{-18}$). These findings suggest that classification
318 of retrieval states is not driven by response related features alone.

319 *Identifying general neural signatures of successful memory function*

320 Having constructed independent classifiers that were capable of estimating the probability of memory success during
321 either encoding or retrieval, we next sought to determine the degree to which these models reflect common or distinct
322 processes during each phase of the free-recall task. We evaluated the degree to which the previously trained classifiers
323 generalized across different phases of the free-recall task by evaluating classifier AUC on held out lists. As shown in
324 Figure 4a, while decoding of memory states was less reliable when performed across phases of the task, we observed
325 above chance classifier performance using this cross-decoding approach, providing evidence that common neural
326 features predict successful memory function during both encoding and retrieval phases of the task. When we tested
327 the encoding success classifier on periods of memory search associated with successful and unsuccessful retrieval in
328 held out lists of data, we observed an average AUC of 0.68 ± 0.01 . Application of the retrieval success classifier to
329 average patterns of activity during encoding yielded an average AUC of 0.59 ± 0.01 .

330 While at first glance our cross-decoding results may suggest that the retrieval success model has greater specificity than
331 the encoding success model, differences in classifier generalization may result from a higher degree of separability in
332 underlying neural states in different phases of the task. We assessed the degree to which model generalization varied
333 as a function of classifier (i.e., encoding or retrieval success) and the phase of the task on which the classifier was
334 tested using a linear mixed effects model. We observed significant effects of classifier type ($\chi^2_2 = 178.69, p < 0.001$
335 bootstrap test), task phase ($\chi^2_2 = 295.25, p < 0.001$ bootstrap test), and a significant interaction between the two

336 ($\chi^2_5 = 177.37$, $p < 0.001$ bootstrap test). A post-hoc test confirmed that the decrease in classifier performance when
337 decoding memory states across phases of the task was greater during the retrieval period ($t_{161} = 11.71$, $p < 10^{-22}$).

338 These findings suggest the presence of common neural predictors of mnemonic success during both encoding and
339 retrieval, in addition to neural features that are specific to each phase of the task. In order to distinguish between
340 these two types of neural activity, we implemented classifiers that were trained on both encoding and retrieval events,
341 and weighted the relative contribution of each phase of the task to learning. As shown in Figure 4b, identification
342 of successful encoding states, as determined by subsequent memory, was modulated by the incorporation of neural
343 activity associated with successful retrieval. When learning was biased towards retrieval processing (i.e., $w_{enc}/w_{ret} <$
344 1), we observed significant ($p < 0.05$, FDR corrected) decreases in prediction of subsequent memory, consistent
345 with our cross-decoding results. Remarkably, performance of joint classifiers were significantly better than the
346 encoding classifier alone at identifying optimal encoding states (Fig. 4b, red line). We next examined the ability
347 of joint classifiers to distinguish between successful and unsuccessful memory search. We observed a significant
348 ($p < 0.05$, FDR corrected) decrease in classifier performance compared to our retrieval classifier (Fig. 4c, blue
349 line) when learning was biased towards patterns of neural activity present during the encoding period. These results
350 demonstrate that the profiles of neural activation that indicate successful retrieval can be leveraged to build a more
351 accurate description of memory states during learning, based on common changes in the local field potential.

352 *Common and distinct neural signatures of encoding and retrieval success*

353 In order to identify which regions of the brain exhibit changes in activity that are generally predictive of successful
354 memory function, we reconstructed patterns of activity that covaried with memory outcomes from joint models with
355 the most extreme weightings (i.e., $w_{enc}/w_{rec} = 0.04$ and $w_{enc}/w_{rec} = 25.12$), and performed conjunction analyses
356 to identify neural features common and specific to the encoding and retrieval models ($p < 0.05$ FDR corrected,
357 permutation procedure for each independent contrast). Increases in HFA and concomitant decreases in low frequency
358 activity (LFA) have been previously identified as neural signatures of successful memory formation using both iEEG
359 in epileptic patients and scalp EEG in healthy individuals (Long et al., 2014). As a results, we focused on these
360 frequency bands when identifying neural features common or specific to episodic encoding and retrieval.

361 As shown in Figure 5a, we observed common increases in HFA within inferior prefrontal, temporal, hippocampal, and
362 parietal cortices. In addition to these common features, the observed differences in classification performance across
363 phases of the task suggest the existence of features that are specific to the encoding and retrieval of list items. We
364 observed significantly increased HFA within lateral occipital regions that was specific to successful encoding (Fig.
365 5b). In contrast to the typical marker of increased HFA supporting memory outcomes, decreased HFA within the right
366 anterior PFC was additionally found to be a marker of successful encoding. Specific to retrieval, bilateral regions
367 of the anterior hippocampus exhibited increased HFA, in addition to bilateral dorsolateral prefrontal cortex, and pre-
368 and postcentral gyri (Fig. 5c). Activation of motor cortex likely reflects movement of the throat and mouth prior to
369 vocalization, as well as potential smearing of signal from the beginning of vocalization.

370 Decreases in low frequency activity commonly predictive of successful encoding and retrieval were observed across
371 multiple cortical regions, including bilateral prefrontal, temporal, and inferior parietal regions (Fig. 6a). Within the
372 MTL, robust decreases in LFA were associated with goodness of memory in left lateralized hippocampus, and bilateral
373 PHG. Changes in LFA specific to encoding were primarily observed in lateral and medial temporal cortices, as well
374 as in additional hippocampal regions (Fig. 6b). LFA effects specific to retrieval were found across widespread cortical
375 sites, including lateral prefrontal, inferior parietal, and visual cortices (Fig. 6c).

376 These findings replicate previous studies that show increased HFA, concurrent with decreased low frequency power
377 across prefrontal, temporal, and hippocampal sites predicts the formation of episodic memories (Burke et al., 2014a;
378 Long et al., 2014). By incorporating these signals into multivariate classifiers, we demonstrate the ability to predict
379 trial-to-trial variability in encoding and retrieval processes based on global brain dynamics.

380 Discussion

381 Through the multivariate classification of iEEG recorded neural activity, we identified neural states that predicted
382 changes in memory encoding and retrieval. By implementing a cross-decoding approach, we tested the correspondence
383 between the processes that occur during encoding and retrieval, and were able to determine a high degree of overlap
384 between neural predictors of successful encoding and recall. At a broad level, our findings suggest that successful

385 encoding and retrieval rely upon the function of common neural substrates, a putative core episodic memory network,
386 and that variability in the activation of this network predicts the mnemonic fate of processed information.

387 *Correspondence between neural processes engaged during encoding and retrieval*

388 Recent approaches to understanding memory function in the human brain have used multivariate classification techniques
389 to characterize neural mechanisms involved in the formation and retrieval of episodic memories (e.g., Kuhl et al., 2012;
390 Kuhl and Chun, 2014). While there is theoretical consensus regarding how the contents of memories are represented
391 in the brain (Rissman and Wagner, 2012), it remains an outstanding question whether activation of common neural
392 pathways can facilitate both episodic encoding and retrieval. Consistent with multiple neuroimaging studies that have
393 compared the formation and retrieval of episodic memories (Zeineh et al., 2003; Eldridge et al., 2005), we observed
394 activation in neural systems that were specific to either the encoding or recall of verbal information. By using a
395 multivariate approach to estimating successful memory processing, we demonstrate that in spite of these dissociations,
396 the same patterns of neural activity observed across multiple recording sites and frequencies were generally predictive
397 of memory function. In doing so, we have identified a putative general memory network spanning the lateral prefrontal
398 cortex, lateral temporal cortex, and MTL whose activation is associated with enhanced memory performance.

399 The common increases in HFA within this network may reflect the engagement of similar cognitive operations across
400 different phases of the recall task, utilizing HFA as a localizer of neuronal activity (Burke et al., 2015). One potential
401 cognitive process that may account for variability in memory performance is goal-direction attention (Corbetta and
402 Shulman, 2002), which would modulate processing of relevant information during the memory task, including the
403 selection of contextually valid information during recall (Cabeza et al., 2008). As an alternative to an attention
404 based account, common changes in spectral power may reflect cognitive operations supporting the maintenance and
405 integration of episodic content (Polyn and Kahana, 2008). Recent work examining the electrophysiological correlates
406 of episodic encoding (Long and Kahana, 2015) argues that increased neural activity within left prefrontal, lateral
407 temporal, and MTL sites reflects processing that supports the formation of episodic memories (i.e., the association of
408 information within a spatiotemporal context), as it predicts subsequent temporal organization of learned information.
409 Our findings build upon this work, and suggest that the operations supported by this network are not specific to the
410 encoding of memories, as they facilitate the retrieval of previously learned content.

411 The features identified by our multivariate models to be predictive of successful memory function match univariate
412 studies of the electrophysiological correlates of episodic encoding (Burke et al., 2014a; Greenberg et al., 2015) and
413 retrieval (Sederberg et al., 2007; Burke et al., 2014b). These studies identified increases in HFA with concomitant
414 decreases in low frequency power across the brain, including recordings from prefrontal and temporal lobe sites, that
415 were indicative of successful memory function. Our results also reveal differences between the neural underpinnings
416 of encoding and retrieval, supporting theoretical models of brain function that emphasize a division of neural resources
417 during each operation. Specific to retrieval, we observed increased HFA within right prefrontal cortex and the MTL,
418 consistent with longstanding neuroimaging findings of asymmetric hemispheric activation during episodic encoding
419 and retrieval (Nyberg et al., 1996). While neural processing may facilitate memory function on a specific phase of
420 the task, such as the observed activation of ventral visual regions during encoding (see also, Burke et al., 2014a),
421 our cross-classification findings demonstrate that similar neural states give rise to successful memory encoding and
422 retrieval. One caveat to this correspondence results from our choice to restrict our analysis to changes in spectral
423 power, leveraging its utility to serve as a marker of local neuronal firing. In addition to changes in levels of local neural
424 activity, functional connectivity between neuronal populations has been shown to predict the formation (Ranganath
425 et al., 2005; Fell et al., 2008) and retrieval (Watrous et al., 2013; Kragel and Polyn, 2015) of episodic memories. This
426 raises the possibility that while changes in spectral activity are conserved across encoding and retrieval, connectivity
427 states that predict performance during encoding and retrieval may differ (Huijbers et al., 2011; Duncan et al., 2014).

428 The models of latent memory states that we have developed can inform the development of closed-loop systems for
429 memory enhancement. Applications of real-time classification have demonstrated the utility in the real-time detection
430 of attentional lapses to enhance learning by manipulating task difficulty (deBettencourt et al., 2015). Attempts to
431 enhance memory through invasive and noninvasive modulation of neural activity (for a review, see Kim et al., 2016)
432 may prove more efficacious by accounting for the state of the brain when stimulation is applied. Recent work
433 has demonstrated that the ability to enhance memory performance via direct brain stimulation during encoding is
434 dependent upon the state of the brain (Ezzyat et al., In press). Given the correspondence between electrophysiological
435 states and proper memory function, our findings suggest that perturbation of this network during retrieval, in addition
436 to encoding, is likely to modulate memory function.

437 *Decoding memory states from patterns of neural activity*

438 A challenge to the study of human memory is to determine the processes that are responsible for both the transformation
439 of experience into lasting memories as well as the processes involved in retrieving encoded information. Neuroscientific
440 endeavors to better elucidate these processes often rely on contrasting markers of neural function based upon the
441 success of an attempt to remember. For example, the subsequent memory paradigm (Paller and Wagner, 2002) has
442 been utilized to reveal the contributions of prefrontal, MTL, and parietal cortex to the process of memory formation
443 in fMRI (Wagner et al., 1998; Davachi et al., 2001; Kim, 2011) and electrophysiological studies of human memory
444 (Sederberg et al., 2007; Long et al., 2014). While these approaches have established a foundation with which to
445 understand the neural processes that underpin human memory, the correspondence between neural states and general
446 memory performance is often overlooked (c.f., Hariri et al., 2003). In contrast, our classification models provide
447 estimates of the probability of memory succeeding or failing based on the observed electrophysiologic state of the
448 brain, effectively providing an estimate of latent cognitive states related to memory processing. This approach enabled
449 us to determine the contributions of processing during encoding and retrieval periods to memory performance.

450 We observed the greatest correspondence between patterns of neural activity and behavior during successful retrieval,
451 as compared to successful encoding. While it is tempting to interpret these results as indicating that neural processing
452 during retrieval, rather than encoding, plays a greater role in determining the ability to remember (for a theoretical
453 perspective on the role of retrieval processes in remembering, see Tulving, 1974), we do not believe this to be the case.
454 If one assumes our classifiers are influenced by the engagement of multiple cognitive operations (e.g., the maintenance
455 of memory representations during encoding and retrieval), then classifier performance will be determined by the degree
456 to which these processes are differentially engaged during successful or unsuccessful memory function. In this light,
457 improved classification of successful retrieval results from a greater correspondence between the outcome of memory
458 search (i.e., successful recall or failure) and the processes that generate changes in spectral activity during recall.
459 The inability to achieve comparable classifier performance during encoding, as well as the relatively limited cortical
460 regions found to be specific to episodic encoding, results from items that were forgotten despite sufficient processing
461 at the time of encoding. During the free-recall task, wherein subjects are free to determine which cues they use to
462 probe memory, memory can fail due to ineffective self-generated retrieval cues. As it is challenging to determine the

463 retrieval cues used by a subject during a specific recall period (c.f., Polyn et al., 2005), future work should examine
464 the relative contributions of encoding processes when retrieval cues are under experimental control. This could rule
465 out variability in memory performance due to inefficient use of retrieval cues during recall, a phenomenon observed
466 in patients with prefrontal damage (Stuss et al., 1994).

467 An alternative interpretation of our findings is that additional episodic encoding occurs in the moments leading up
468 to recall. Indeed, the act of retrieval is known to enhance learning, with activation of inferior prefrontal and inferior
469 temporal regions (Buckner et al., 2001) as well as structures within the MTL (Stark and Okado, 2003) predicting the
470 degree to which novel recognition probes are learned during retrieval tasks. We believe it is unlikely that the present
471 results solely reflect incidental encoding processes engaged during free recall. It has been demonstrated that processing
472 within the prefrontal cortex and MTL that predicts subsequent remembering is sensitive to the novelty of learned
473 information (Kirchhoff et al., 2000), with MTL activation attenuating in response to the repetition of well encoded
474 stimuli (Turk-Browne et al., 2006). As retrieval of encoded information results from the reactivation of previously
475 formed item representations, processes associated with the encoding of novel stimuli into long-term memory are
476 unlikely to be engaged during memory search (Kragel and Polyn, 2016).

477 *Conclusion*

478 The processes involved in encoding an event into a lasting memory and remembering that event at a later point in
479 time rely upon inherently distinct neural mechanisms. We demonstrated that multivariate models of memory success
480 can generalize across the encoding and recall of verbal information. These findings suggest that in the presence of
481 neural processes that are specific to the encoding and recall of episodic memories, engagement of a putative core
482 memory system generally shapes the ability to remember the past. Determining the contribution of this system to
483 experimental paradigms that involve different aspects of episodic memory (e.g., the spatial and autobiographical nature
484 of remembering) is a critical next step in identifying neural systems that are generally predictive of the ability to
485 remember.

References

- 486
488 Avants BB, Epstein CL, Grossman M, Gee JC. 2008. Symmetric diffeomorphic image registration with
487
489 cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*.
490 12:26–41.
- 491 Buckner R, Wheeler M, Sheridan M. 2001. Encoding processes during retrieval tasks. *Journal of Cognitive*
492 *Neuroscience*. 13:406–415.
- 493 Burke JF, Long NM, Zaghoul KA, Sharan AD, Sperling MR, Kahana MJ. 2014a. Human intracranial high-frequency
494 activity maps episodic memory formation in space and time. *NeuroImage*. 85:834–843.
- 495 Burke JF, Ramayya AG, Kahana MJ. 2015. Human intracranial high-frequency activity during memory processing:
496 Neural oscillations or stochastic volatility? *Current Opinion in Neurobiology*. 31:104–110.
- 497 Burke JF, Sharan AD, Sperling MR, Ramayya AG, Evans JJ, Healey MK, Beck EN, Davis KA, Lucas TH, Kahana MJ.
498 2014b. Theta and high-frequency activity mark spontaneous recall of episodic memories. *Journal of Neuroscience*.
499 34:11355–11365.
- 500 Burke JF, Zaghoul KA, Jacobs J, Williams RB, Sperling MR, Sharan AD, Kahana MJ. 2013. Synchronous and
501 asynchronous theta and gamma activity during episodic memory formation. *Journal of Neuroscience*. 33:292–304.
- 502 Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. 2008. The parietal cortex and episodic memory: an attentional
503 account. *Nature Reviews Neuroscience*. 9:613–625.
- 504 Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews*
505 *Neuroscience*. 3:201 – 215.
- 506 Danker JF, Anderson JR. 2010. The ghosts of brain states past: Remembering reactivates the brain regions engaged
507 during encoding. *Psychological Bulletin*. 136:87–102.
- 508 Daselaar SM, Prince SE, Dennis NA, Hayes SM, Kim H, Cabeza R. 2009. Posterior midline and ventral parietal
509 activity is associated with retrieval success and encoding failure. *Frontiers in Human Neuroscience*. 3:13.
- 510 Davachi L, Maril A, Wagner AD. 2001. When keeping in mind supports later bringing to mind: neural markers of
511 phonological rehearsal predict subsequent remembering. *Journal of Cognitive Neuroscience*. 13:1059–1070.

- 512 deBettencourt MT, Cohen JD, Lee RF, Norman KA, Turk-Browne NB. 2015. Closed-loop training of attention with
513 real-time brain imaging. *Nature Neuroscience*. 18:470–475.
- 514 Desikan R, Segonne B, Fischl B, Quinn B, Dickerson B, Blacker D, Buckner RL, Dale A, Maguire A, Hyman B,
515 Albert M, Killiany N. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI
516 scans into gyral based regions of interest. *NeuroImage*. 31:968–80.
- 517 Duncan K, Tompary A, Davachi L. 2014. Associative Encoding and Retrieval Are Predicted by Functional
518 Connectivity in Distinct Hippocampal Area CA1 Pathways. *Journal of Neuroscience*. 34:11188–11198.
- 519 Dykstra AR, Chan AM, Quinn BT, Zepeda R, Keller CJ, Cormier J, Madsen JR, Eskandar EN, Cash SS.
520 2012. Individualized localization and cortical surface-based registration of intracranial electrodes. *NeuroImage*.
521 59:3563–3570.
- 522 Eldridge L, Engel S, Zeineh M, Bookheimer S, Knowlton B. 2005. A dissociation of encoding and retrieval processes
523 in the human hippocampus. *Journal of Neuroscience*. 25:3280–3286.
- 524 Ezzayat Y, Kragel JE, Burke JF, Levy DF, Lyalenko A, Wanda P, O’Sullivan L, Hurley K, Busygin S, Pedisich I,
525 Sperling MR, Worrell GA, Kucewicz MT, Davis KA, Lucas TH, Inman CS, Lega BC, Jobst BC, Sheth S, Zaghoul
526 K, Jutras M, Stein JM, Das S, Gorniak R, Rizzuto DS, Kahana MJ. In press. Direct brain stimulation modulates
527 encoding states and memory performance in humans. *Current Biology*.
- 528 Fan RE, Chang KW, Hsieh CJ, Wang XR, Lin CJ. 2008. Liblinear: A library for large linear classification. *The*
529 *Journal of Machine Learning Research*. 9:1871–1874.
- 530 Fell J, Ludowig E, Rosburg T, Axmacher N, Elger C. 2008. Phase-locking within human mediotemporal lobe predicts
531 memory formation. *NeuroImage*. 43:410–419.
- 532 Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy
533 D, et al. 2004. Automatically parcellating the human cerebral cortex. *Cerebral Cortex*. 14:11–22.
- 534 Geller AS, Schleifer IK, Sederberg PB, Jacobs J, Kahana MJ. 2007. PyEPL: A cross-platform
535 experiment-programming library. *Behavior Research Methods*. 39:950–58.

- 536 Glanzer M, Cunitz AR. 1966. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*.
537 5:351–360.
- 538 Greenberg JA, Burke JF, Haque R, Kahana MJ, Zaghoul KA. 2015. Decreases in theta and increases in high frequency
539 activity underlie associative memory encoding. *NeuroImage*. 114:257–263.
- 540 Hariri AR, Goldberg TE, Mattay VS, Kolachana BS, Callicott JH, Egan MF, Weinberger DR. 2003. Brain-Derived
541 Neurotrophic Factor val66met Polymorphism Affects Human Memory-Related Hippocampal Activity and Predicts
542 Memory Performance. *Journal of Neuroscience*. 23:6690–6694.
- 543 Haufe S, Meinecke F, Görgen K, Dähne S, Haynes JD, Blankertz B, Bießmann F. 2014. On the interpretation of
544 weight vectors of linear models in multivariate neuroimaging. *NeuroImage*. 87:96–110.
- 545 Hermann B, Seidenberg M, Jones J. 2008. The neurobehavioural comorbidities of epilepsy: can a natural history be
546 developed? *The Lancet Neurology*. 7:151–160.
- 547 Hermes D, Miller KJ, Vansteensel MJ, Edwards E, Ferrier CH, Bleichner MG, van Rijen PC, Aarnoutse EJ, Ramsey
548 NF. 2014. Cortical theta wanes for language. *NeuroImage*. 85:738–748.
- 549 Huijbers W, Pennartz CM, Cabeza R, Daselaar SM. 2011. The hippocampus is coupled with the default network
550 during memory retrieval but not during memory encoding. *PLoS One*. 6:e17463.
- 551 Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimisation for the robust and accurate linear
552 registration and motion correction of brain images. *NeuroImage*. 17:825–841.
- 553 Kahana MJ. 1996. Associative retrieval processes in free recall. *Memory & Cognition*. 24:103–109.
- 554 Kim H. 2011. Neural activity that predicts subsequent memory and forgetting: a meta-analysis of 74 fMRI studies.
555 *NeuroImage*. 54:2446–2461.
- 556 Kim H. 2015. Encoding and retrieval along the long axis of the hippocampus and their relationships with dorsal
557 attention and default mode networks: The HERNET model. *Hippocampus*. 25:500–510.
- 558 Kim H, Daselaar SM, Cabeza R. 2010. Overlapping brain activity between episodic memory encoding and retrieval:
559 Roles of the task-positive and task-negative networks. *NeuroImage*. 49:1045–1054.

- 560 Kim K, Ekstrom AD, Tandon N. 2016. A network approach for modulating memory processes via direct and indirect
561 brain stimulation: Toward a causal approach for the neural basis of memory. *Neurobiology of Learning and Memory*.
- 562 Kirchhoff BA, Wagner AD, Maril A, Stern CE. 2000. Prefrontal-temporal circuitry for episodic encoding and
563 subsequent memory. *Journal of Neuroscience*. 20:6173–6180.
- 564 Kolers PA. 1973. Remembering operations. *Memory & Cognition*. 1:347–355.
- 565 Kolers PA, Roediger HL. 1984. Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*. 23:425–449.
- 566 Kragel JE, Polyn SM. 2015. Functional Interactions Between Large-Scale Networks During Memory Search. *Cerebral*
567 *Cortex*. 25:667–679.
- 568 Kragel JE, Polyn SM. 2016. Decoding Episodic Retrieval Processes: Frontoparietal and Medial Temporal Lobe
569 Contributions to Free Recall. *Journal of Cognitive Neuroscience*. 28:125–139.
- 570 Kuhl BA, Bainbridge W, Chun M. 2012. Neural reactivation reveals mechanisms for updating memory. *Journal of*
571 *Neuroscience*. 32:3453–3461.
- 572 Kuhl BA, Chun MM. 2014. Successful remembering elicits event-specific activity patterns in lateral parietal cortex.
573 *Journal of Neuroscience*. 34:8051–8060.
- 574 Lepage M, Habib R, Tulving E. 1998. Hippocampal PET activations of memory encoding and retrieval: the HIPER
575 model. *Hippocampus*. 8:313–322.
- 576 Long NM, Burke JF, Kahana MJ. 2014. Subsequent memory effect in intracranial and scalp EEG. *NeuroImage*.
577 84:488–494.
- 578 Long NM, Kahana MJ. 2015. Successful memory formation is driven by contextual encoding in the core memory
579 network. *NeuroImage*. 119:332–337.
- 580 Manning JR, Jacobs J, Fried I, Kahana MJ. 2009. Broadband shifts in local field potential power spectra are correlated
581 with single-neuron spiking in humans. *Journal of Neuroscience*. 29:13613–13620.
- 582 Maris E, Oostenveld R. 2007. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience*
583 *Methods*. 164:177–190.

- 584 McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the
585 hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and
586 memory. *Psychological Review*. 102:419–457.
- 587 Miller KJ, Honey CJ, Hermes D, Rao RP, den Nijs M, Ojemann JG. 2014. Broadband changes in the cortical surface
588 potential track activation of functionally diverse neuronal populations. *NeuroImage*. 85:711–720.
- 589 Murdock BB, Okada R. 1970. Interresponse times in single- trial free recall. *Journal of Verbal Learning and Verbal*
590 *Behavior*. 86:263–267.
- 591 Nyberg L, Cabeza R, Tulving E. 1996. Pet studies of encoding and retrieval: The hera model. *Psychonomic Bulletin*
592 *& Review*. 3:135–148.
- 593 O'Reilly RC, Rudy JW. 2001. Conjunctive representations in learning and memory: Principles of cortical and
594 hippocampal function. *Psychological Review*. 108:311 – 345.
- 595 Paller KA, Wagner AD. 2002. Observing the transformation of experience into memory. *Trends in Cognitive Sciences*.
596 6:93–102.
- 597 Polyn SM, Kahana MJ. 2008. Memory search and the neural representation of context. *Trends in Cognitive Sciences*.
598 12:24–30.
- 599 Polyn SM, Natu VS, Cohen JD, Norman KA. 2005. Category-specific cortical activity precedes retrieval during
600 memory search. *Science*. 310:1963–1966.
- 601 Polyn SM, Norman KA, Kahana MJ. 2009. A context maintenance and retrieval model of organizational processes in
602 free recall. *Psychological Review*. 116:129–56.
- 603 Prince S, Daselaar S, Cabeza R. 2005. Neural Correlates of Relational Memory: Successful Encoding and Retrieval
604 of Semantic and Perceptual Associations. *Journal of Neuroscience*. 25:1203–1210.
- 605 Ranganath C, Cohen MX, Brozinsky CJ. 2005. Working memory maintenance contributes to long-term memory
606 formation: neural and behavioral evidence. *Journal of Cognitive Neuroscience*. 17:994–1010.

- 607 Rissman J, Wagner AD. 2012. Distributed representations in memory: Insights from functional brain imaging. Annual
608 Review of Psychology. 63:101–128.
- 609 Ritchey M, Wing EA, LaBar KS, Cabeza R. 2013. Neural Similarity Between Encoding and Retrieval is Related to
610 Memory Via Hippocampal Interactions. Cerebral Cortex. 23:2818–2828.
- 611 Rohrer D, Wixted JT. 1994. An analysis of latency and interresponse time in free recall. Memory & Cognition.
612 22:511–524.
- 613 Rosenberg MD, Finn ES, Constable RT, Chun MM. 2015. Predicting moment-to-moment attentional state.
614 NeuroImage. 114:249–256.
- 615 Sederberg PB, Miller JF, Howard WH, Kahana MJ. 2010. The temporal contiguity effect predicts episodic memory
616 performance. Memory & Cognition. 38:689–699.
- 617 Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, McCarthy DC, Brandt A, Tully MS, Kahana MJ.
618 2007. Hippocampal and neocortical gamma oscillations predict memory formation in humans. Cerebral Cortex.
619 17:1190–1196.
- 620 Small S, Nava A, Perera G, DeLaPaz R, Mayeux R, Stern Y. 2001. Circuit mechanisms underlying memory encoding
621 and retrieval in the long axis of the hippocampal formation. Nature Neuroscience. 4:442–449.
- 622 Staresina BP, Henson RN, Kriegeskorte N, Alink A. 2012. Episodic reinstatement in the medial temporal lobe. Journal
623 of Neuroscience. 32:18150–18156.
- 624 Stark CE, Okado Y. 2003. Making memories without trying: medial temporal lobe activity associated with incidental
625 memory formation during recognition. Journal of Neuroscience. 23:6748–6753.
- 626 Strange BA, Otten LJ, Josephs O, Rugg MD, Dolan RJ. 2002. Dissociable human perirhinal, hippocampal, and
627 parahippocampal roles during verbal encoding. Journal of Neuroscience. 22:523–528.
- 628 Stuss DT, Alexander MP, Palumbo CL, Buckle L, Sayer L, Pogue J. 1994. Organizational Strategies of Patients With
629 Unilateral or Bilateral Frontal Lobe Injury in Word List Learning Tasks. Neuropsychology. 8:355–355.
- 630 Tulving E. 1974. Cue-dependent forgetting. American Scientist. 62:74–82.

- 631 Tulving E, Thompson DM. 1973. Encoding specificity and retrieval processes in episodic memory. *Psychological*
632 *Review*. 80:352–373.
- 633 Turk-Browne NB, Yi DJ, Chun MM. 2006. Linking implicit and explicit memory: common encoding factors and
634 shared representations. *Neuron*. 49:917–927.
- 635 Wada J, Rasmussen T. 1960. Intracarotid Injection of Sodium Amytal for the Lateralization of Cerebral Speech
636 Dominance. *Journal of Neurosurgery*. 17:266–282.
- 637 Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL. 1998.
638 Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science*.
639 281:1188–1191.
- 640 Watrous AJ, Tandon N, Conner CR, Pieters T, Ekstrom AD. 2013. Frequency-specific network connectivity increases
641 underlie accurate spatiotemporal memory retrieval. *Nature Neuroscience*. 16:349–356.
- 642 Wing EA, Ritchey M, Cabeza R. 2014. Reinstatement of Individual Past Events Revealed by the Similarity of
643 Distributed Activation Patterns during Encoding and Retrieval. *Journal of Cognitive Neuroscience*. 27:679–691.
- 644 Yushkevich PA, Pluta JB, Wang H, Xie L, Ding SL, Gertje EC, Mancuso L, Kliot D, Das SR, Wolk DA. 2015.
645 Automated volumetry and regional thickness analysis of hippocampal subfields and medial temporal cortical
646 structures in mild cognitive impairment. *Human Brain Mapping*. 36:258–287.
- 647 Zeineh M, Engel S, Thompson P, Bookheimer S. 2003. Dynamics of the hippocampus during encoding and retrieval
648 of face-name pairs. *Science*. 299:577.

649 **Figure Legends**

650 **Figure 1 Recall task and electrode coverage. a.** Experimental paradigm. On each trial, patients studied a list of
651 words, performed a self paced arithmetic task, and finally recalled the items studied on the most recent list in any order.
652 **b.** Electrode coverage. Left, the number of subjects with a bipolar electrode centers within 10mm of each vertex of the
653 average cortical surface. Right, cross sections along the longitudinal axis of the MTL showing the number of subjects
654 with bipolar electrode coverage localized within hippocampus (Hipp) or parahippocampal gyrus (PHG), extending
655 3mm from electrode centroids. A priori anatomical regions of interest excluded from a subset of analyses are depicted
656 in yellow. The depicted coverage discounts electrodes within the epileptogenic zone. L, left; R, right.

657

658 **Figure 2 Behavioral results. a.** The overall probability that items from each list position were recalled. **b.** The
659 probability that an item from a given list position was the first item recalled, as a function of serial position. **c.** The
660 probability that a recall transition will come from a temporally proximal list position, given it is available for recall. **d.**
661 Inter-response time (measured from onset to onset of response vocalization) as a function of recall sequence length.
662 Error bars represent standard error of the mean.

663

664 **Figure 3 Decoding successful memory formation and retrieval. a.** Probability of encoding success as a function
665 of list position. The top panel depicts the probability of recalling an item as a function of serial position, for each
666 tercile of classifier output. Below, the difference in percent recall change (relative to average percent recall) between
667 the upper and lower tercile of classifier output. **b.** Probability of retrieval success increases across the memory search
668 period. The top panel shows the average proportion of recalls made in epochs decoded across the recall period, sorted
669 into terciles based upon classifier output. Below, the difference in the proportion of recalls made relative to the mean
670 between upper and lower terciles of classifier output. **c.** Overall change in memory performance during encoding
671 (ENC) and retrieval (RET) between high and low classifier terciles. Error bars represent standard error of the mean.

672

673 **Figure 4 Common neural signatures of successful memory function. a.** Classifier performance as a function
674 of task phase. Average area under the curve (AUC) for classifiers trained to detect successful encoding and retrieval
675 processing. Error bars reflect standard error of the mean, across subjects. **b.** Predicting successful encoding using a

676 joint classifier. AUC for joint classifiers applied to encoding period data, with significant ($p < 0.05$, FDR corrected)
677 increases and decreases in performance, compared to the encoding classifier (dashed line) are shown in red and blue,
678 respectively. Shaded area depicts standard error of the mean. **c.** Predicting successful retrieval using a joint classifier.
679 AUC for joint classifiers applied to retrieval period observations, with significant decreases in performance compared
680 to the retrieval classifier alone (dashed line) are depicted in blue ($p < 0.05$, FDR corrected).

681

682 **Figure 5 High frequency activity (HFA) common and specific to successful episodic encoding and retrieval.**

683 **a.** Regions common to successful encoding and retrieval. **b.** Regions with HFA that significantly differs between
684 subsequently remembered and forgotten items, exclusively masking out any regions that exhibit retrieval related
685 effects. **c.** Regions showing significant differences in HFA, specific to episodic retrieval. Significant ($p < 0.05$,
686 FDR corrected) increases and decreases in HFA depicted on the cortical surface (left) and MTL subregions (right) are
687 shown in red and blue, respectively.

688

689 **Figure 6 Low frequency activity (LFA) features common and specific to successful episodic encoding and**

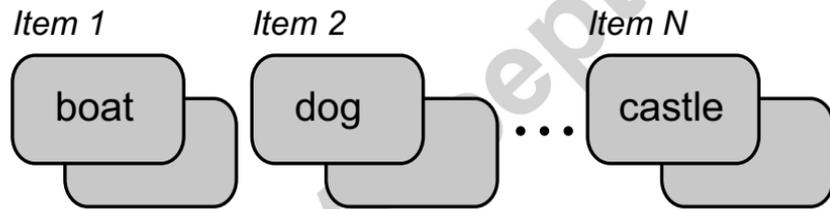
690 **retrieval. a.** Regions exhibiting significant differences in LFA during successful memory processing common to
691 successful encoding and retrieval. **b.** Regions with changes in LFA that significantly differs between subsequently
692 remembered and forgotten items, exclusively masking out any regions that exhibit retrieval related effects are shown.
693 **c.** Regions showing significant differences in LFA with memory success, specific to episodic retrieval. Significant
694 ($p < 0.05$, FDR corrected) increases and decreases in LFA are depicted on the cortical surface (left) and MTL
695 subregions (right) are shown in red and blue, respectively.

696

697

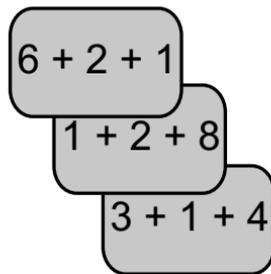
a

Encoding



1600 ms Stimulus ~1000 ms ISI

Delay



~20 sec (self paced)

Retrieval

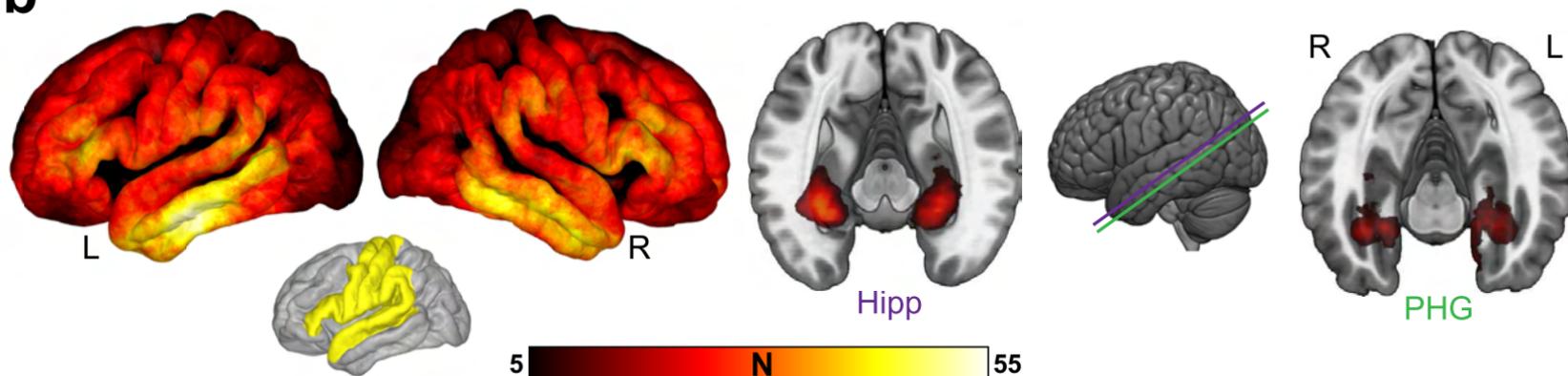
Recall Cue

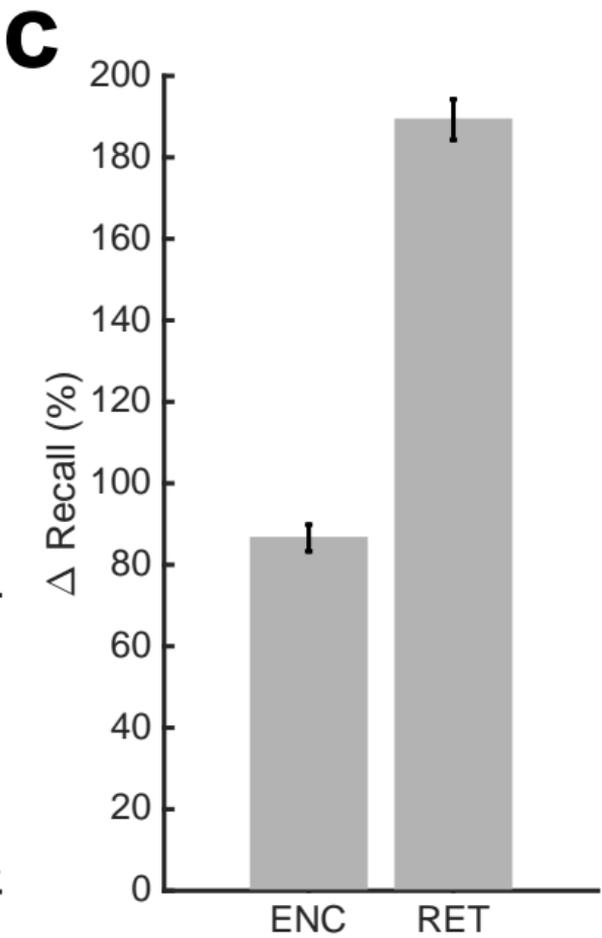
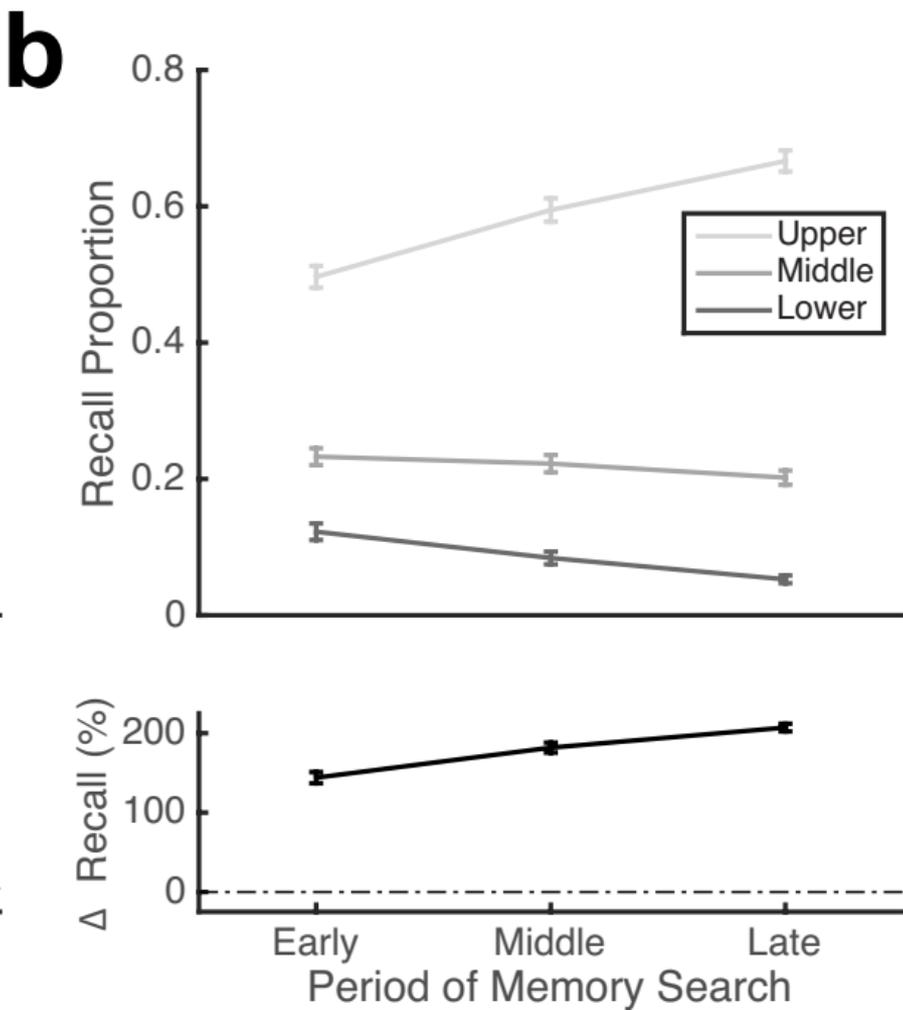
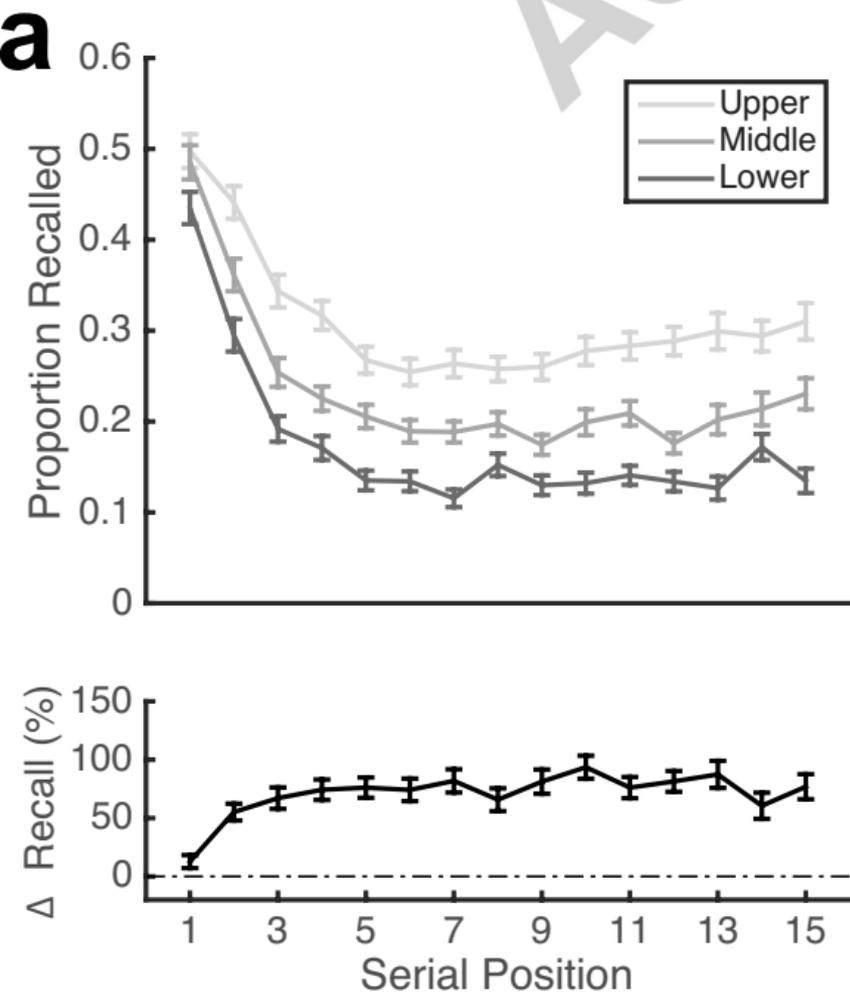


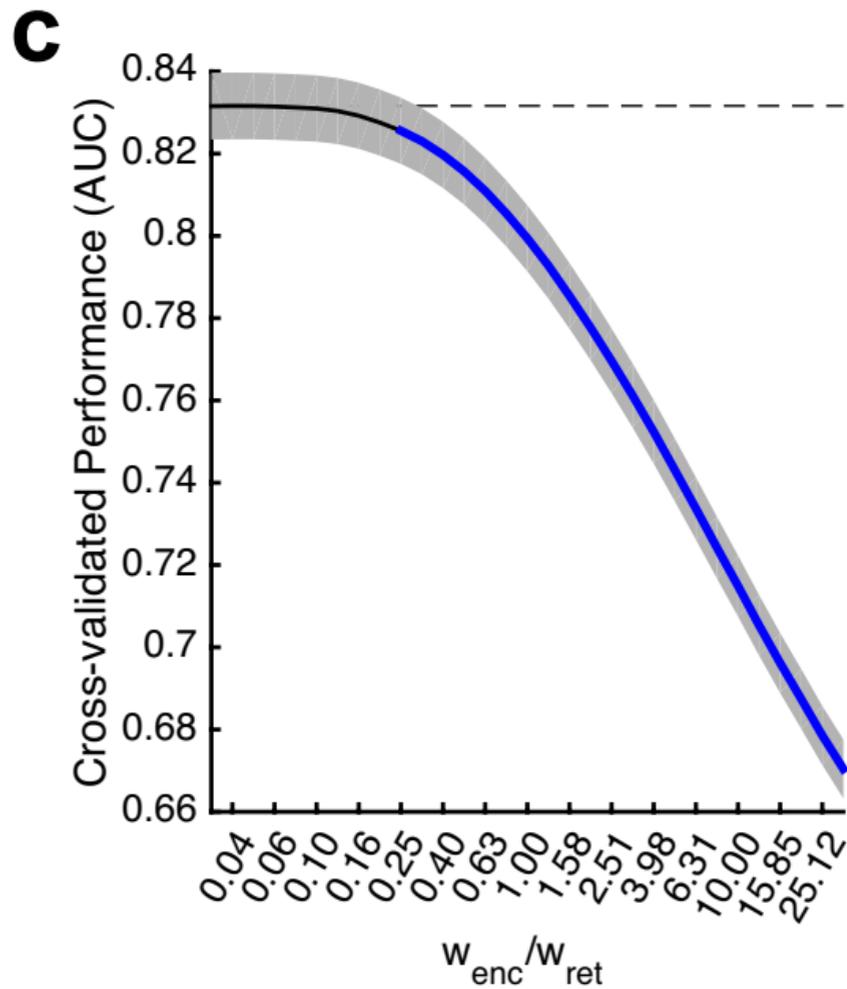
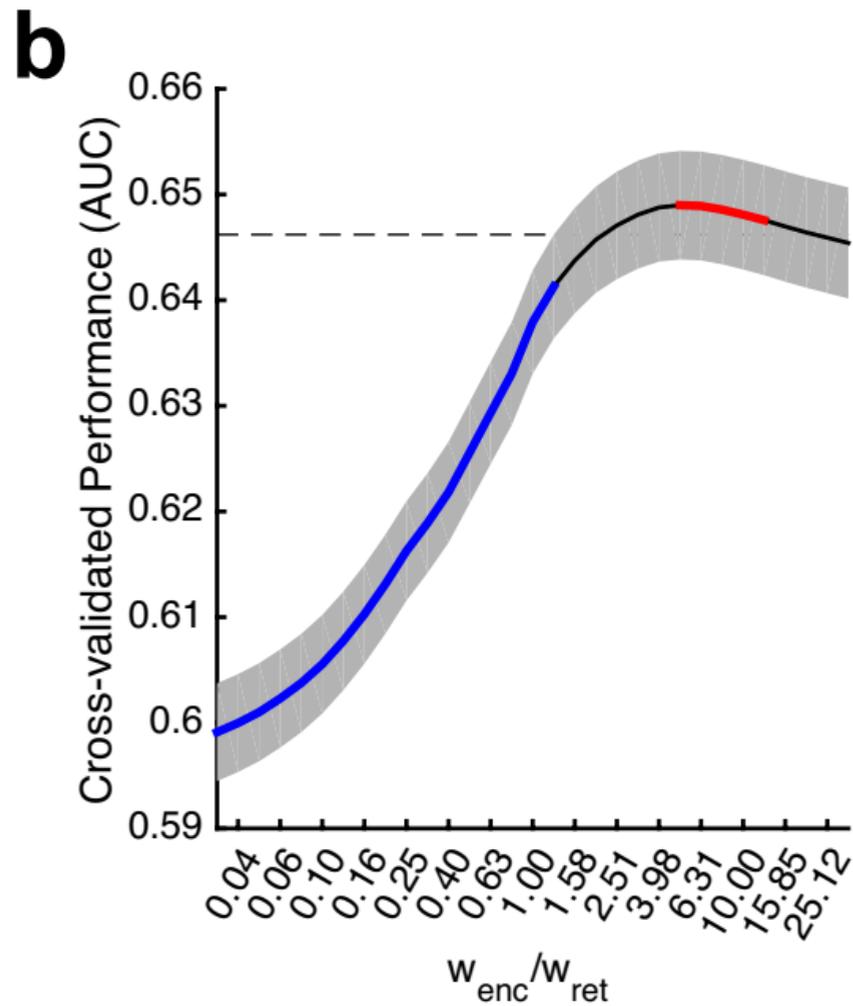
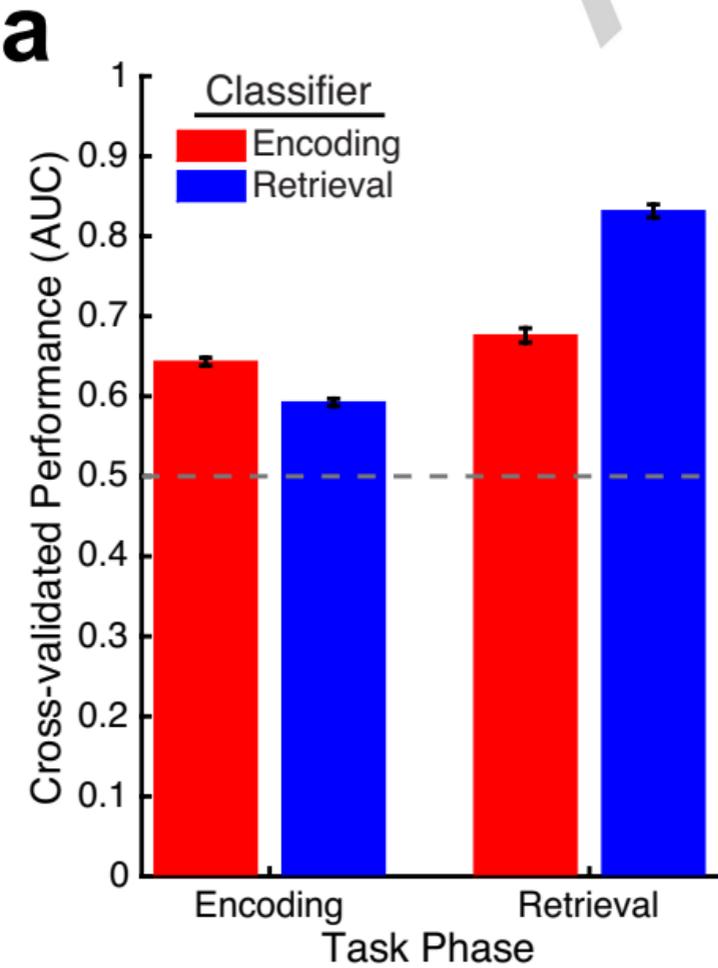
30 sec (LL = 12)

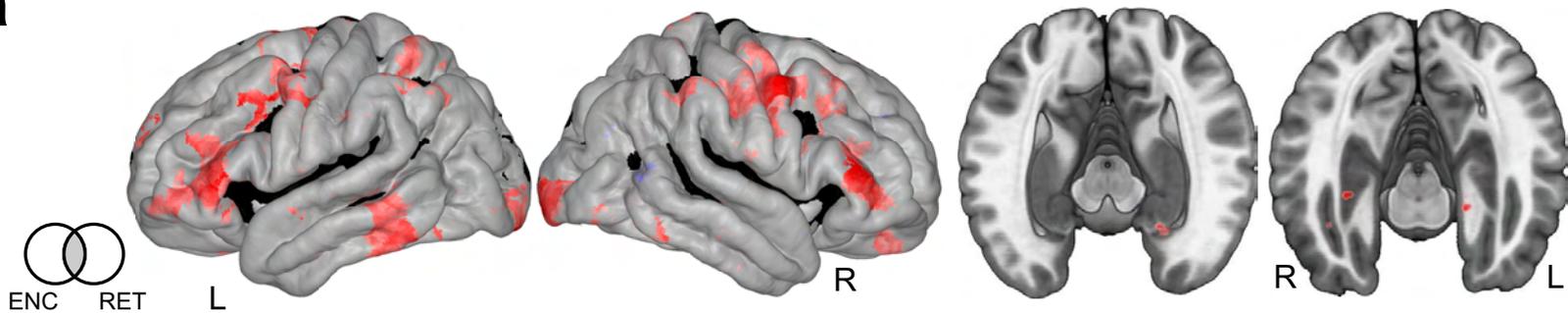
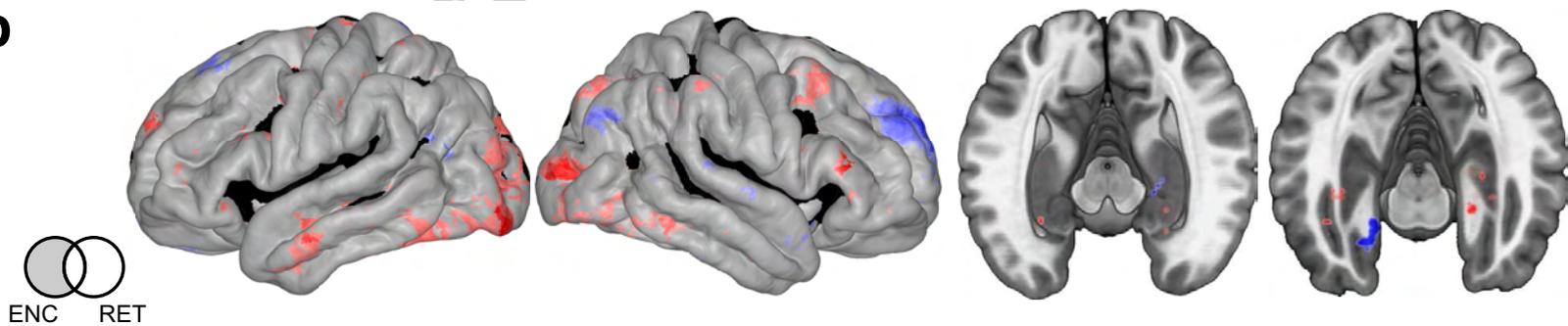
45 sec (LL = 15)

time

b





a**b****c**