JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Articles: Behavioral/Cognitive

Differentiation of human medial prefrontal cortex activity underlies long-term resistance to forgetting in memory

Youssef Ezzyat[1], Marika Inhoff[2] and Lila Davachi[3]

[1]*Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104*
[2]*Department of Psychology, University of California, Davis, California 95616*
[3]*Department of Psychology, Columbia University, New York, New York, 10027*

DOI: 10.1523/JNEUROSCI.2290-17.2018

Received: 13 August 2017

Revised: 12 June 2018

Accepted: 12 June 2018

Published: 16 July 2018

Author contributions: Y.E., M.C.I., and L.D. designed research; Y.E. and M.C.I. performed research; Y.E. and M.C.I. analyzed data; Y.E. and L.D. wrote the paper; M.C.I. and L.D. edited the paper.

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

This work was supported by DART Neuroscience and NIH R01 MH074692 to L.D.

Correspondence should be addressed to Contact information for the corresponding author: Lila Davachi, Ph.D., Department of Psychology, Columbia University, Schermerhorn Hall, 1190 Amsterdam Ave #406, New York, NY 10027, Email: ld24@columbia.edu, Phone: 212-854-1754

Cite as: J. Neurosci ; 10.1523/JNEUROSCI.2290-17.2018

Research Article

Differentiation of human medial prefrontal cortex activity underlies long-term resistance to forgetting in memory

Youssef Ezzyat [1], Marika Inhoff [2] and Lila Davachi [3]

[1] Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

[2] Department of Psychology, University of California, Davis, California 95616

[3] Department of Psychology, Columbia University, New York, New York, 10027

Running title: MPFC pattern differentiation and memory retrieval

Contact information for the corresponding author:
Lila Davachi, Ph.D.
Department of Psychology
Columbia University
Schermerhorn Hall
1190 Amsterdam Ave #406
New York, NY 10027

Email: ld24@columbia.edu
Phone: 212-854-1754

This work was funded by DART Neuroscience and NIH R01 MH074692 to L.D.
Pages: 42
Tables: 0
Figures: 6

Keywords: medial prefrontal cortex, hippocampus, temporal context, pattern separation, episodic memory retrieval, fMRI, consolidation

Abstract

It is well known that distributing study events over time leads to better memory over long timescales, compared to massing study events together. One explanation for such long-term resistance to forgetting is that distributed study leads to neural differentiation in memory, which supports retrieval of past experiences by disambiguating highly similar memory representations. Neuroanatomical models of episodic memory retrieval propose that the hippocampus and medial prefrontal cortex (MPFC) work together to enable retrieval of behaviorally appropriate memories. However, it is not known how representations in these regions jointly support resistance to forgetting long after initial learning. Using fMRI, we measured differentiation in retrieved memory representations following an extended delay in male and female human participants. After one week, word-object associations were better remembered if studied across two days (*overnight*), allowing associations to be learned in distinct temporal contexts, compared to learning within a single day (*same day*). MPFC retrieval patterns showed differentiation for *overnight* relative to *same day* memories, while hippocampal patterns reflected associative retrieval success. *Overnight* memory differentiation in MPFC was higher for associative than item memories and higher than differentiation assessed over a brain-wide set of retrieval-active voxels. The memory-related difference in MPFC pattern differentiation correlated with memory success for *overnight* learning and with hippocampal—MPFC functional connectivity. These results show that learning information across days leads to differentiated MPFC memory representations, reducing forgetting after one week, and suggest this arises from persistent interactions between MPFC and hippocampus.

66 **Significance Statement**

67 Neural activity in both the hippocampus and medial prefrontal cortex (MPFC) has been linked to
68 memory-related representations, but prior work has not examined how these representations sup-
69 port episodic memory retrieval over extended timescales that are characteristic of everyday re-
70 trieval. We show that differentiation in MPFC activity one week after encoding is higher for re-
71 trieved information learned across two days compared to within a single day. In hippocampus,
72 differentiation was greater for detailed memory retrieval but was not influenced by whether in-
73 formation had been learned over one or two days. Differentiation in MPFC predicted behavioral
74 robustness to forgetting and was correlated with hippocampal—MPFC connectivity. The results
75 suggest that context-based differentiation supports robust long-term memory via persistent
76 MPFC—hippocampal interactions.
77
78

Introduction

Humans have the ability to retrieve episodic memories long after the initial experience, but not all memories maintain their fidelity over time. Behavioral and computational modeling work (Estes, 1955; Heusser et al., 2016; Howard & Kahana, 2002; Polyn et al., 2009a; Polyn et al., 2009b) suggests that context-based differentiation at encoding facilitates later memory retrieval. Memory in these studies is typically assessed following a relatively short delay, leaving open the question of how context-based differentiation at encoding might support retrieval over the longer timescales characteristic of everyday episodic retrieval.

Activity in both the hippocampus and medial prefrontal cortex (MPFC) during encoding has been shown to represent contextual information that supports memory, making each a candidate region for context representation following long delays (Kitamura et al., 2017). In rodents, ensemble activity in hippocampus becomes more differentiated over time in a way that predicts temporal memory for learned information (Mankin et al., 2012; Manns et al., 2007). Similarly, ensembles in MPFC show pattern differentiation with changes in context, although these representations appear to integrate a broader range of contextual changes (Hyman et al., 2012). In humans, blood-oxygen-level dependent (BOLD) fMRI encoding patterns in hippocampus (DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014; Hseih et al., 2014) and MPFC (Jenkins & Ranganath, 2016; Tompary and Davachi 2017) reflect the temporal context of learned information and predict temporal memory. Hippocampal activity also reflects contextual retrieval when memory is tested within one day of learning (Long et al., 2017; Ritchey et al., 2015).

Although activity in both the hippocampus and MPFC reflects memory-related context representations, less is known about how they support memory over delays that extend beyond one day. There is growing evidence that interactions between the hippocampus and MPFC are

critical to memory retrieval. These interactions occur during and following learning of new information (Guise & Shapiro, 2017; Schlichting & Preston 2016; Siapas et al., 2005; Tompary and Davachi 2017), and are thought to reflect processes in which new hippocampal memory representations are integrated into established MPFC cortical networks. This implies that the hippocampal and MPFC contextual representations that support memory should change over time, with MPFC patterns assuming a more important role in supporting detailed memory retrieval over time. While there is evidence that univariate MPFC fMRI activity is greater during retrieval of temporally remote compared to recent memories (Gais et al., 2007; Nieuwenhuis & Takashima, 2011; Sterpenich et al., 2009; Takashima et al., 2006) and more differentiable for remote compared to recent memories (Bonnici et al., 2012) there is little evidence that MPFC patterns reflect contextual information for memories that are all equally remote.

Here, we ask how learning information across contexts leads to neural pattern differentiation in MPFC and hippocampus that supports resistance to forgetting. Each participant was trained to a high criterion (90%) on two separate lists of word-object associations. To manipulate whether associations were learned in the same or distinct contexts, training sessions occurred either across a 24-hour period (*overnight*) or within the same session (*same day*). One week later, participants were scanned using fMRI during cued retrieval to compare the multivariate patterns of retrieved memories that had been learned in the *overnight* and *same day* conditions. To foreshadow the core findings, we observed that learning across a 24-hour period led to more differentiated long-term memory representations in MPFC, which correlated with the behavioral benefit of *overnight* learning over same-day learning. Functional connectivity between MPFC and hippocampus also correlated with the degree of MFPC differentiation, suggesting that context-

based differentiation supports memory through hippocampal-cortical interactions at least one week following learning.

Materials and Methods

Participants

Twenty-two healthy volunteers (15 female; mean age = 23.3, range = 19-29) from the New York University and New York City communities participated in this study for payment. All participants were native English speakers with normal or corrected-to-normal vision. Informed consent for this experiment was obtained in a manner approved by the University Committee on Activities Involving Human Subjects at New York University. Three participants were excluded from analysis due to mechanical issues with the fMRI scanner and all remaining 19 participants were included in behavioral analyses. For the fMRI analyses, two additional participants were excluded due to having fewer than 5 responses in at least one memory condition.

Stimuli

The words used in the experiment were taken from a set of 506 adjectives downloaded from the MRC psycholinguistics database (http://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm). For each participant, 72 words were randomly chosen from the pool to be used in the *overnight*, *same day* and *single session* study conditions; an additional 72 words were randomly chosen for each participant to be used as *new* lures for the Day 7 retrieval test. Within each of the *overnight*, *same day* and *single session* study lists, half of the words were paired with pictures of *natural* objects (e.g. animals, plants) and half were paired with pictures of *manmade* objects (e.g. tools, kitchen gadgets). All object stimuli were taken from an Internet image search. Our randomization procedure ensured that the assignment of words and images to each other and to the overnight/same day/single session conditions was equally likely for all stimuli.

Experimental Design and Statistical Analysis

General Procedure

On Day 1 participants came into the lab and were trained to criterion (90% correct) on a set of word-object pairings (*overnight* pairs). Twenty-four hours later (Day 2) participants returned to the lab and were trained to criterion on a novel set of word-object pairings (*same day* pairs). Following this training on the *same day* pairs, participants were re-trained to criterion (90% criterion, see *Behavioral Procedures* for details on training) on both sets from Day 1 and Day 2 (Figure 1A). Thus, participants were trained to criterion twice on both lists; for the *overnight* list this occurred across two days, for the *same day* list this occurred within the same day. One week later (Day 7) participants came into the lab and were trained to criterion on a final novel set of word-object pairings (*single session* pairs). Immediately following this final training session, participants were brought to the fMRI scanner and were scanned while performing a retrieval test for all the word-object pairings learned across all three days (~15 minute delay between the final study session in the lab and the start of the scanned retrieval test). The results of this retrieval test were analyzed to look for behavioral evidence of differences in memory performance on the basis of study condition (*overnight*/*same day*/*single session*), as well as neural evidence for differentiation in memory representations on the basis of study condition when controlling for memory.

Behavioral Procedures

On Day 1 participants came into the lab and were trained to learn 72 word-object associations (*overnight* pairs). Training began with a study phase in which participants were presented

with all of the word-object pairs and asked to memorize the associations (Figure 1B, left panel; no behavioral response was required). Each study phase trial began with 500 ms of fixation, followed by presentation of the word-object pair for 4500 ms. After the study phase, participants began the training phase in which the words were presented alone and participants had to indicate whether the word had been paired with a 'natural' or 'manmade' object during study, or whether they were unsure (Figure 1B, middle panel). Each training trial began with 500 ms of fixation followed by the presentation of a word for a maximum of 12 s. Once a participant made a natural/manmade choice, feedback was given and the object stimulus was presented alongside the word for 8.5 s. Training was broken into three blocks of 24 trials based on pilot data suggesting faster learning for several smaller blocks compared to a single large block. Once all 24 trials were tested, training began again. If a participant correctly answered the natural/manmade decision on two consecutive training rounds, the pair was dropped from future training rounds. Participants continued with training until they reached 90% of the 24 pairs correct. The training procedure was then repeated for the second and third blocks of 24 pairs. After reaching the 90% criterion on the third and final block, participants performed a final test on all 72 pairs together and were required to reach 90% correct. This was done to ensure that participants still showed robust learning for pairs that had been learned in the first and second blocks. Three participants that performed slightly below the 90% criterion (minimum = 86%) on this final test were re-exposed to the full set of 72 pairs one final time. Participants were then dismissed until the following day.

Participants returned 24 hours later on Day 2 and were trained to 90% criterion on a novel set of 72 word-object pairs (*same day* pairs) using a training paradigm that was identically structured to the training on Day 1. After performing the final test on this new set of pairs, participants performed another training session that included all 72 pairs from Day 1 intermixed with

all 72 pairs from Day 2 (Figure 1A) in a single block. Thus, participants were re-trained to 90% criterion on both the *overnight* and *same day* lists; the critical difference is that *overnight* pairs were first trained on the previous day while the *same day* pairs were first trained earlier in that same session. All participants met or exceeded 90% correct for both *overnight* and *same day* lists, after which they were dismissed until the following week.

One week after the Day 1 session, participants returned to the lab for the Day 7 session. They began this session with another learning phase on a novel set of 72 word-object pairs (*single session* pairs). After reaching criterion on this set, participants were taken to the fMRI scanner for the final testing session. In the scanner, participants were presented with all of the studied words (72 each of *overnight*, *same day*, and *single session*) as well as 72 novel words (Figure 1B, right panel) in a randomly ordered sequence. On each trial, participants were presented with a word and were asked to indicate which type of object the word had been paired with during the learning phases (*Old-Natural*; *Old-Manmade*; *Old-Unsure*; *New*). Participants were given a maximum of 6 s to make their response. Following a response, participants performed an odd/even judgment on a series of numbers for the remainder of the trial, up to a total trial length of 24 s.

Behavioral Analysis

Analysis of the behavioral data focused on the Day 7 retrieval test and was designed to identify memory differences relating to study list condition (*overnight*/*same day*/*single session*). Trials for which participants responded correctly to the *Natural*/*Manmade* decision (e.g. selecting *Old-Natural* when the word was old and had been paired with a 'natural' object) were coded as *associative*; trials for which participants responded incorrectly to the *Natural/Manmade* decision (e.g. selecting *Old-Natural* when the word was old but had been paired with a 'manmade'

object) or responded *Old-Unsure* were coded as *item-only*. Trials for which participants responded *New* when the word was 'old' were coded as *forgotten*. Correctly identified *new* words were coded as correct rejections (CR), and *new* words that were given any *Old* response were coded as false alarms (FA). Overall memory performance was defined as the difference between the hit rate (any *old* response to an old trial) minus the false alarm rate. We defined the *overnight learning benefit* as (*overnight* associative) – (*same day* associative) proportion correct.

FMRI Procedures

As described in the *Behavioral Procedures*, on Day 7 participants first completed the learning phase for the *single session* stimuli in the lab. After completion, participants were brought to the scanner and performed a scanned retrieval test that included all of the learned words (*overnight*, *same day*, and *single session*) as well as novel words (response options: *Old-Natural*; *Old-Manmade*; *Old-Unsure*; *New*). Participants performed a total of 288 retrieval trials split evenly into 12 scanning runs.

FMRI Data Acquisition

Functional imaging was performed using a Siemens Allegra 3T head-only scanner with a custom head coil (NM-011; Nova Medical) located at the Center for Brain Imaging at New York University. Functional data were collected using an echo-planar (EPI) pulse sequence (34 contiguous slices; TR = 2000 ms; 3 mm isotropic voxels; TE = 15 ms; flip angle = 82°; field of view = 240 x 192 mm; slice gap = 0%) with slices oriented parallel to the AC-PC axis. Slices were positioned ventrally to provide full coverage of the anterior temporal lobes and prefrontal cortex; this resulted in omission of parts of the superior parietal cortex and, occasionally, parts of motor

cortex. A high-resolution T1-weighted anatomical scan (magnetization-prepared rapid-acquisition gradient echo sequence, 1 x 1 x 1 mm) was also obtained for each subject following the final block of the localizer task.

Preprocessing of fMRI Data

Images were preprocessed using SPM8 software (Wellcome Trust Centre for Neuroimaging, London UK). Functional images were realigned to the within-run mean to correct for head motion (one run from one participant was discarded due to head motion > 1 voxel). Realigned images were corrected for slice acquisition time and were then coregistered to the anatomical image to correct for between-run motion. For group-level analyses, the coregistered images were first spatially normalized to an EPI template in Montreal Neurological Institute space, resliced to 2 x 2 x 2 mm voxels and finally smoothed with a 6 mm FWHM isotropic Gaussian kernel. Low frequencies (< 1.5 cycles/run) were removed from the functional data in both the subject-specific and group analyses.

ROI Definition

Anatomical hippocampal ROIs were drawn manually on each participant's T1-weighted anatomical scan using an in-house drawing tool written in Matlab (Mathworks, Sherborn MA) and according to standard anatomical convention (Insausti et al., 1998). For the analysis comparing pattern differentiation in anterior and posterior hippocampus, we split the hippocampus into thirds based on coronal slice position, taking the anterior and posterior thirds as the ROIs. We also defined functional ROIs based on fMRI activation during the retrieval task. We first created an across-participant random effects general linear model (GLM) that included eight conditions

of interest (source correct/item-only for each of the *overnight*/*same day*/*single session* lists; and correct rejections/false alarms) that modeled activation for each trial as a 2-TR boxcar convolved with a canonical hemodynamic response function (HRF). Trials without a response and forgotten trials were each separately modeled as confounds, as were estimates of across-run participant motion. The resulting beta estimates were then used to define MPFC; specifically, a contrast of *single session associative* and *item-only* trials was conducted against baseline at a voxel-wise threshold of $p < 0.001$ to functionally identify the region [MNI coordinates: (0, 56, 18)]. Based on a comparison of our region to available MNI coordinates of prior work, the center of our MPFC region is more anterior and/or dorsal to some ROIs that have been reported previously (e.g. Bonnici et al., 2012; Gais et al., 2007; Takashima et al., 2007), but is partially overlapping with others (e.g. Tompary & Davachi, 2017). Importantly, we used only *single session* trials to define the MPFC region in order to avoid biasing our main pattern similarity analyses, which were limited to comparisons between the *overnight* and *same day* conditions.

We also defined, individually for each subject, a mask of the 5000 voxels showing the strongest responses (positive or negative) in a contrast of Task > Baseline. We defined these brain-wide task active masks as a proxy for voxels that might be expected to show reinstatement of contextual information during retrieval (Howard et al., 2015).

Pattern Analysis of fMRI Data

Analysis of multivariate patterns evoked during memory retrieval was conducted on functional data from the retrieval runs. In brief, estimates of activation on each trial were computed at every voxel in the brain. These single-trial estimates were then extracted as spatially-distributed patterns across relevant ROIs. To compute single-trial estimates of activation, each of

the 24 trials in a run was modeled as a separate condition in a general linear model (GLM). Trials were modeled as two-TR boxcars beginning at trial onset convolved with a canonical hemodynamic response function (HRF). Mean intensity across the run, linear drift and estimates of subject motion were modeled as nuisance regressors. The procedure was repeated within each of the 12 runs of the retrieval task, resulting in a GLM estimated for each run and a parameter estimate for each trial in each run in every voxel in the brain.

Patterns of parameter estimates were extracted from all voxels within an ROI and individual trial patterns were separated according to study condition (*overnight*/*same day*/*single session*) and memory (*associative*/*item-only*). We then measured pattern differentiation as follows: we computed the Pearson correlation between the pattern for each trial and all other trials in the same condition [i.e. a correlation value was computed between each *overnight-associative* trial and all other *overnight-associative* trials, (LaRocque et al., 2013)]. We then applied Fisher's *r*-to-*z* transformation to the between-trial similarities, converted the measures to differentiation by taking 1-*z* and averaged within participant to generate a global measure of differentiation across trials of the same condition. We then used ANOVAs and *t*-tests across participants to measure effects of learning condition and memory performance on neural differentiation.

Beta-Series Connectivity

To measure connectivity between hippocampus and MPFC we used the beta series method (Rissman, Gazzaley, & D'Esposito, 2004; Vilberg & Davachi, 2013). Within each subject, we used the same output of the single-trial GLM model that we used to compute pattern differentiation (see *Pattern Analysis of fMRI Data* above). These single-trial beta estimates were sorted by condition (*overnight* and *same day*) and averaged across voxels within the hippocampal and

MPFC ROIs. This yielded a single beta estimate for each region for each retrieval trial of the experiment. We then computed the Pearson correlation between the beta estimates across retrieval trials (the beta-series) from hippocampus and MPFC, and did so separately within the *overnight* and *same day* conditions. To assess the relationship between connectivity and pattern differentiation, we then correlated the difference in beta-series connectivity between the *overnight* and *same day* conditions with the difference in pattern similarity for the two conditions.

Results

Memory Performance

Participants performed well above chance on the Day 7 retrieval test. Collapsing across study condition (*overnight*/*same day*/*single session*) and memory status (associative/item only) status, the difference between the hit and false alarm rates for judging words as *Old*/*New* was significantly greater than zero [M = 0.60 ± 0.04; $t(18) = 14.8$; $p < 0.001$]. Participants also showed high levels of memory for the word-object associations. We compared the proportion of *associative* trials (collapsing across study list) to an expected mean of 0.5 (i.e. an equal number of *associative* and *item-only* trials), and found significantly better *associative* memory than expected by chance [*associative* M = 0.74 ± 0.03, *item-only* M = 0.22 ± 0.03; $t(18) = 9.0$; $p < 0.001$; Figure 2A].

To determine whether *associative* memory differed by study list, we conducted a one-way ANOVA with study condition (*overnight*/*same day*/*single session*) as a within-participant factor, and observed a significant main effect [$F(2,36) = 55.0$; $p < 0.001$]. In addition to our primary conditions of interest (*overnight*/*same day*), participants were also trained to criterion on a novel set of word-object associations just prior to the Day 7 retrieval test (*single session* condition; see *Experimental Procedures*). Planned comparisons revealed that the proportion of *associative* memory in the *single session* condition (M = 0.93 ± 0.02) was significantly greater than in both the *overnight* [M = 0.72 ± 0.05; $t(18) = 5.35$, $p < 0.001$] and *same day* conditions [M = 0.57 ± 0.03; $t(18) = 10.9$; $p < 0.001$]. This effect was expected given that participants had just learned the *single session* pairs to criterion prior to the retrieval test, while they had last seen the *overnight* and *same day* pairs the week before. Importantly, however, *associative* memory was also significantly higher in the *overnight* condition than in the *same day* condition [$t(18) = 4.9$; $p <$

0.001], demonstrating that distributing learning across two days, relative to study within a single day, benefitted associative memory retrieval even after one week. Thus, we defined the *overnight learning benefit* as the difference in *associative* memory between the *overnight* and *same day* conditions, and use this measure in later analyses of the fMRI data. Finally, we also examined whether the different study conditions resulted in different levels of forgetting (i.e. *Old* trials given a *New* response). A one-way ANOVA with study list as a within-participant factor showed a significant main effect [$F(2,36) = 10.9$; $p < 0.001$] that was driven by significantly more forgetting of the *same day* list [M = 0.07 ± 0.02] compared to the *overnight* [M = 0.02 ± 0.01; $t(18) = 3.58$, $p = 0.002$] and *single session* [M = 0.02 ± 0.01; $t(18) = 3.33$, $p = 0.004$] lists.

We also analyzed response times to determine whether they varied according to memory and study condition. A comparison of response times for *associative* and *item-only* memories, collapsed across study condition, showed that participants were significantly faster to make *associative* responses than *item-only* responses [*associative* M = 2.60 ± 0.08 s, *item-only* M = 4.08 ± 0.23 s; $t(18) = 7.17$, $p < 0.001$, Figure 2B]. Because six participants did not have any *single session item-only* trials, we focused our analysis on comparing *associative* response times across study conditions using a one-way ANOVA. The main effect of study condition was significant [$F(2,36) = 59.1$, $p < 0.001$, Figure 2B] and planned comparisons showed that *single session associative* response times (M = 1.99 s ± 0.08 s) were significantly faster than both *overnight* [M = 2.63 s ± 0.10 s; $t(18) = 7.09$, $p < 0.001$] and *same day* [M = 3.18 s ± 0.12 s; $t(18) = 8.67$, $p < 0.001$] response times. Finally, *overnight* response times were also significantly faster than *same day* response times [$t(18) = 5.80$, $p < 0.001$]. Taken together, object-word pairs learned across two days, relative to those learned on a single day, were better and more quickly recalled after

one week, consistent with the notion that distributing learning across contexts benefited the access to and accuracy of those memory representations.

Finally, although we observed the previously described differences between the *overnight* and *same day* conditions on Day 7, these were not due to differences between conditions during training. As described in the Methods, both conditions were trained to 90% criterion. In addition, the conditions did not differ in the number of training rounds required to reach criterion (mean *overnight* = 3.3, mean *same day* = 3.2, $p = 0.60$ by Wilcoxon signed rank test) and in the number of exposures to each item during the training phases on Days 1 and 2 (mean *overnight* = 7.6, mean *same day* = 7.5, $p = 0.10$ by rank sum test). We also found that the difference in number of presentations did not predict the difference in Day 7 hit rates [$r(17) = -0.11$, $p = 0.66$, Spearman rank correlation] nor source correct memory [$r(17) = -0.07$, $p = 0.76$, Spearman rank correlation].

Pattern Differentiation in Hippocampus

In order to measure whether memories are represented in a distinct manner that benefits retrieval, we computed the differentiation in hippocampal multivariate BOLD activity patterns evoked during retrieval of each memory with patterns evoked during retrieval of all the other memories from that learning session. We found that, in left hippocampus, pattern differentiation was significantly higher for *associative* compared to *item-only* memories [$F(1,16) = 11.3$, $p = 0.004$], but did not vary by *overnight* vs. *same day* learning [$F(1,16) = 2.60$, $p = 0.12$], with no interaction between factors ($p = 0.58$, Figure 3A). Planned comparisons confirmed more within-condition hippocampal pattern differentiation for *associative* trials compared to *item-only* trials in both the *overnight* [*associative* $z = 0.90 \pm 0.01$, *item-only* $z = 0.86 \pm 0.03$; $t(16) = 2.17$, $p = 0.04$] and *same day* [*associative* $z = 0.88 \pm 0.02$, *item-only* $z = 0.85 \pm 0.02$; $t(16) = 2.36$, $p =$

386 0.03] conditions. We also conducted our pattern differentiation analyses separately for voxels in
387 the anterior and posterior thirds of the left hippocampus, defined by coronal slice position along
388 the anterior/posterior axis. Using a 3-way ANOVA (anterior/posterior X *associative*/*item-only* X
389 *overnight*/*same day*), we found that pattern differentiation in left hippocampus was greater for
390 anterior compared to posterior ROIs [main effect: $F(1,16) = 7.05$, $p = 0.02$, Figure 3B], greater
391 for *associative* compared to *item-only* trials [$F(1,16) = 7.27$, $p = 0.02$], and greater for *overnight*
392 compared to *same day* trials [$F(1,16) = 9.03$, $p = 0.008$]. There was a trend for a 3-way interac-
393 tion ($p = 0.10$), but all other interactions with anterior/posterior ROI were not significant ($p >$
394 0.41).

395 We also examined right hippocampus, which showed higher differentiation for associa-
396 tive compared to item memories [$F(1,16) = 14.4$, $p = 0.002$] and a trend for greater differentia-
397 tion with overnight study [$F(1,16) = 4.09$, $p = 0.06$] with a significant difference observed in
398 both the *overnight* [*associative* $z = 0.89 \pm 0.02$, *item-only* $z = 0.83 \pm 0.03$; $t(16) = 2.48$, $p = 0.02$]
399 and *same day* conditions [*associative* $z = 0.87 \pm 0.02$, *item-only* $z = 0.82 \pm 0.02$; $t(16) = 4.12$, $p <$
400 0.001]. There was no interaction between study list and memory ($p = 0.76$). These results pro-
401 vide evidence for hippocampal pattern separation during a one-week memory test that is related
402 to successful associative memory retrieval. When comparing anterior/posterior ROIs in right
403 hippocampus, we also found a main effect such that pattern differentiation was greater in anterior
404 compared to posterior right hippocampus [$F(1,16) = 9.81$, $p = 0.006$, Figure 3B] and greater for
405 *associative* compared to *item-only* trials [$F(1,16) = 13.0$, $p = 0.002$]. There were no other main
406 effects or interactions (all $p > 0.12$).

407 To determine whether the level of pattern differentiation in hippocampal memory repre-
408 sentations was related to memory performance, we correlated *overnight* > *same day associative*

memory differentiation with the *overnight learning benefit*. We also conducted the same analysis between *overnight > same day item-only* memory differentiation and the *overnight > same day item-only* behavioral memory difference. In both left and right hippocampus, neither the *associative* correlation [left: $r(15) = 0.31$, $p = 0.23$, Figure 3A; right: $r(15) = 0.30$, $p = 0.24$] nor the *item-only* correlation [left: $r(15) = -0.28$, $p = 0.27$; right: $r(15) = 0.32$, $p = 0.21$] was significant.

In sum, pattern separation in hippocampus was higher for retrieved memories that contained associated source information compared to memories containing only item information, irrespective of whether information was learned the week before across 24 hours or within the same session. We observed greater differentiation in anterior compared to posterior hippocampus, which highlights a dissociation along the long-axis of the hippocampus (also see, Tompary & Davachi, 2017), but is not fully consistent with models that predict more separated representations in posterior compared to anterior subregions (Poppenck et al., 2013). In contrast we do think the current findings are consistent with a role for anterior hippocampus in a distributed network that represents specific item details in episodic memory (Ranganath & Ritchey, 2012).

Pattern Differentiation in MPFC

As in left hippocampus, there was significantly more differentiation in MPFC for *associative* memories compared to *item-only* memories [$F(1,16) = 9.45$, $p = 0.007$]. Interestingly, however, unlike the results from hippocampus, there was a significant interaction between study list (*overnight*/*same day*) and memory [$F(1,16) = 7.39$, $p = 0.02$] that was driven by a difference in *associative*/*item-only* memory specific to the *overnight* list [*associative* $z = 0.92 \pm 0.02$, *item-only* $z = 0.84 \pm 0.03$; $t(16) = 3.18$, $p = 0.006$], with no difference in *associative*/*item-only* memory for the *same day* list ($p = 0.16$; Figure 4A). A direct comparison of *associative* trials in

the *overnight* and *same day* conditions showed that *overnight* differentiation was higher than *same day* differentiation [*same day* $z = 0.88 \pm 0.02$; $t(16) = 3.70$, $p = 0.002$].

We next investigated whether MPFC differentiation was related to the behavioral enhancement seen for *overnight associative* memory relative to *same day associative* memory. We found that *overnight* > *same day* pattern differentiation in MPFC for *associative* memory trials was significantly correlated with the *overnight learning benefit* [$r(15) = 0.53$, $p = 0.03$, Figure 4A]. Thus, greater differentiation in evoked MPFC neural patterns for the *overnight* condition relative to the *same day* condition was associated with *increases* in *overnight associative* memory, relative to *same day associative* memory (this correlation did not differ from the hippocampal correlation $p = 0.18$). There was no correlation between *overnight* > *same day* differentiation for *item-only* memories and *overnight* > *same day item-only* memory success [$r(15) = 0.38$, $p = 0.13$], but this did not differ from the *associative* correlation ($p = 0.30$). These data show that, relative to *same day* memories, greater pattern differentiation in MPFC patterns across *overnight associative* memories predicts a larger *overnight learning benefit*. More broadly, these data are consistent with the idea that differentiation in measured BOLD retrieval patterns in MPFC supports enhancements in associative memory derived from studying information across an overnight delay.

One explanation for the differentiation of retrieval patterns in MPFC is simply that *overnight* memories are tagged during encoding with information from two contexts, which is then reinstated at retrieval. If this were the case, then one prediction that would follow is that *overnight* memories should be more likely to cluster into two groups than *same day* memories. To test this we applied *k*-means clustering to segment the *overnight* and *same day* patterns for each subject into two clusters. We then assessed whether the between-cluster distance for the *over-*

night patterns was larger than for the *same day* patterns, as would be the case if there was greater evidence for two distinct clusters in the *overnight* condition. In MPFC we did not find greater evidence for two clusters in the *overnight* compared to *same day* conditions [mean within vs. between cluster correlation: *overnight* = 0.22, *same day* = 0.21, $t(16) = 1.21$, $p = 0.24$] (there was also no difference in the left hippocampus [*overnight* = 0.17, *same day* = 0.19, $t(16) = -1.02$, $p = 0.32$]).

To confirm that this null finding was not the result of lack of power to discriminate conditions of interest, we repeated the analysis and used it to cluster *associative* and *item-only* patterns into groups, as a way to classify the two trial types. We assessed classification for each subject using area under the curve (AUC) and then compared the AUC distribution across subjects to chance (0.50). We found that the mean of the group AUC distribution was significantly above chance (M = 0.55, $p < 0.001$), confirming that the analysis was sufficiently powered to separate *associative* and *item-only* MPFC patterns. Taken together, these results suggest that greater MPFC differentiation was not necessarily driven by the direct representation variable encoding of contextual information during learning.

Brain-Wide Pattern Differentiation

In our final set of analyses, we were interested in whether large-scale activity distributed across the brain exhibited similar patterns as observed in hippocampus and MPFC. For each subject we defined a mask of the 5000 most active voxels in a contrast of Task > Baseline and computed pattern differentiation across this distributed set of voxels (single-subject example, Figure 4B). Unlike the hippocampus and MPFC, brain-wide pattern differentiation did not differ based on memory retrieval success [associative vs. item-only main effect $F(1,16) = 0.11$, $p = 0.74$] but

was higher overall for pairs learned *overnight* compared to *same day* [$F(1,16) = 7.89$, $p = 0.01$, Figure 4B].

When comparing pattern differentiation across the three ROIs (MPFC/hippocampus/ brain-wide) we found that the interaction of region × study list (*overnight*/*same day*) × memory [$F(2,32) = 2.82$, $p = 0.07$] was driven by greater study list and memory-dependent pattern differentiation in MPFC than over the distributed task active brain mask ($p = 0.03$). In left hippocampus, study list and memory-dependent pattern differentiation did not differ from that observed in the task active brain mask ($p = 0.29$).

As in MPFC and hippocampus, we next examined whether the difference in pattern differentiation for *overnight* compared to *same day* pairs predicted the *overnight-same day* behavioral enhancement in in associative memory performance. Brain-wide pattern differentiation was not significantly correlated with the behavioral enhancement [$r(15) = 0.03$, $p = 0.88$, Figure 4B] and showed a trend for being lower than observed in MPFC (permutation test, $p = 0.08$).

MPFC—Hippocampal Connectivity Predicts Differentiation

To shed light on the mechanism by which MPFC representations of *overnight* memories were more differentiated than *same day* memories, our final analysis measured beta-series connectivity between hippocampus and MPFC during correct *associative* retrieval. We hypothesized that if context-based MPFC differentiation arises from reorganization in mnemonic networks (Siapas & Wilson, 1998; Takashima et al., 2009; van Kesteren, Fernández, Norris, & Hermans, 2010), then connectivity with hippocampus may be related to this mnemonic differentiation in MPFC. Indeed, we found that, across participants, the difference in hippocampal-MPFC connectivity for *overnight* memories relative to *same day* memories correlated with the increase in rep-

resentational differentiation for *overnight* compared to *same day* memories in MPFC [$r(15) = 0.60$, $p = 0.01$; Figure 5]. Thus, those participants who exhibited greater hippocampal-MPFC connectivity during retrieval of *overnight* memories compared to *same day* memories also showed more pattern separation for *overnight* compared to *same day* memories in MPFC. This suggests that study across an overnight delay leads to enhanced hippocampal-MPFC interactions that may support the development of differentiated memory representations over time.

Univariate Effects in Left Hippocampus and MPFC

Although we were focused on assessing pattern differentiation, we also examined univariate activity in our primary regions of interest. Specifically, we tested whether parameter estimates in left hippocampus and MPFC significantly differed for the *overnight*/*same day* and *associative*/*item-only* conditions. A 2x2 ANOVA applied to the data from the left hippocampus (Figure 6A) showed no significant main effects or interaction (all $p > 0.12$). Figure 6B shows that MPFC was more active for the overnight compared to the same day condition [$F(1,16) = 8.27$, $p = 0.01$] and also showed a study condition x memory interaction [$F(1,16) = 6.73$, $p = 0.02$]. A 3x2 ANOVA comparing the two ROIs showed a significant main effect of region [$F(1,16) = 138$, $p < 0.001$] but no interactions of region with study condition or memory.

Discussion

Understanding the mechanisms that promote the longevity of learning is a critical focus of cognitive neuroscience. We find that learning word-object pairs across the distinct temporal contexts of two days benefited long-term associative memory retrieval compared to word-object pairs learned on one day, even though both *overnight* and *same day* associations were trained to the same high learning criterion (90%) and were matched in the number of training sessions. At one-week retrieval, multivariate activity patterns in MPFC were more differentiated for associations learned across temporal contexts compared to associations learned within a single temporal context. Furthermore, the increase in pattern separation for *overnight* compared to *same day* memories in MPFC correlated with the behavioral memory benefit for the *overnight* memories, suggesting that differentiation of MPFC retrieval patterns contributes to the memory benefit for information studied across contexts. We also found a correlation between context-driven pattern differentiation of MPFC activity for individual memories and functional connectivity between hippocampus and MPFC, suggesting that persistent hippocampal-cortical connectivity one week after learning supports the observed MPFC differentiation. These data show that robust long-term memory is supported by differentiation in the neural activity patterns that represent memories and suggest that this differentiation arises as a consequence of context change during learning that promotes sustained MPFC—hippocampal interactions.

Given that participants encoded *overnight* pairs on two days and same day pairs on a single day, one explanation for greater MPFC pattern differentiation for the *overnight* condition is that participants encoded, and subsequently retrieved, associative details from two temporal contexts (Estes, 1955). One prediction that might follow from such a model is that *overnight* patterns should be more likely to cluster into two groups than *same day* memories, perhaps reflecting dif-

ferential levels of contextual associations across memories. However, we did not find evidence that *overnight* memories were more likely to cluster into two groups than *same day* memories, suggesting that MPFC pattern differentiation may arise as a consequence of repeated learning across days, but does not directly reflect variability in contextual associations across conditions. Repeated learning across contexts may facilitate the long-term retention of unique memory traces through mechanisms that enhance the fidelity of individual memory representations (Karlsson Wirebring et al., 2015; Thios & D'Agostino, 1976).

Future work could investigate the degree to which the night of sleep itself (in the *overnight* condition) is important for promoting differentiation. We observed differentiation for *overnight* compared to *same-day* memories in spite of the fact that both types of memories were matched in several ways. First, both types were trained to the same high behavioral criterion. Second, both *overnight* and *same-day* memories had the opportunity for significant consolidation to occur during the one week between study and retrieval. These unique features of our design highlight the important role that reactivation following a 24-hour period played in the *overnight* condition. One possibility is that the overnight-delayed reactivation tagged *overnight* memories with enhanced behavioral relevance, leading these memories to undergo enhanced consolidation the following week (Lesburguères et al., 2011; Oudiette, Antony, Creery, & Paller, 2013; Singer & Frank, 2009; Wilhelm et al., 2011). Another possibility is that one night of sleep prior to restudy establishes a *schema* (in the sense of a pre-existing neural network), which then promotes additional learning following the night of sleep (Tse et al., 2007; Tse et al., 2011; van Kesteren et al., 2012). Each of these studies includes some period of rest before reactivation, which appears to be critical in enhancing memory retention (Litman & Davachi, 2008). Thus, future work that

563 can compare representations at encoding, rest and retrieval over extended timescales will be
564 needed to adjudicate between these and other possibilities.

565 The observation that distributing study episodes across time is beneficial for memory rel-
566 ative to massing study within a shorter time span is one of the most robust findings in the study
567 of memory (Cepeda et al., 2006; Dempster, 1989; Ebbinghaus, 1964; Litman & Davachi, 2008;
568 McGeoch & Irion, 1952; Ruch, 1928). Our behavioral results are consistent with this literature,
569 showing that object-word associations were better remembered one week later if initially learned
570 in two sessions distributed across two days than associations learned in two sessions within the
571 same day. Our design equated the number of learning sessions across the *overnight* and *same-*
572 *day* lists, as well as the level of immediate memory for the two lists. Thus, our data suggest that
573 the memory benefit for the *overnight* condition one week later is due to learning across two days.
574 Furthermore, the fMRI results highlight a potential mechanism for the efficacy of distributed
575 learning – namely, that it promotes efficient, differentiated neural representations of memories
576 that facilitates successful retrieval.

577 Prior fMRI work has suggested that the extent to which patterns overlap (pattern *similari-*
578 *ty*) across repeated learning trials of the same stimulus is related to successful recognition
579 memory. This work has shown that repeated individual stimuli that are later remembered were
580 associated with greater similarity in the evoked fMRI activity patterns across encoding repeti-
581 tions (Gordon et al., 2013; Ward et al., 2013; Watanabe et al., 2011; Xue et al., 2010, 2013).
582 Similarity across trial-unique same-category stimuli during encoding is predictive of memory
583 success (Kuhl, Rissman, & Wagner, 2012), as is similarity between patterns present during en-
584 coding and retrieval (Ritchey, Wing, Labar, & Cabeza, 2012; Zeithamova & Preston, 2010;
585 Tompary et al, 2016; Danker et al, 2016), again suggesting the important role that similarity

plays in processing the same events in memory. This general approach is distinct from the one we adopt here in which we measure how memories are represented compared to one-another: an inter-item similarity analysis. Prior work using a similar approach suggests that both similarity and differentiation during encoding support later memory, but are instantiated in distinct brain areas, the MTL cortex and hippocampus, that support slow extraction of commonalities across memories vs. rapid differentiation of overlapping experiences (LaRocque et al., 2013, Norman & O'Reilly, 2003, Tompary & Davachi 2017). It is possible that enhanced similarity across learning repetitions of the same item could promote differentiation *across items* at retrieval (i.e. our primary neural measure) by sharpening the representations of individual memories relative to one another, following repeated study/test attempts during learning. Testing-related neural differentiation is known to predict memory performance after a one-week delay (Karlsson Wirebring et al., 2015); therefore, one critical question for future research is to examine the extent to which repeated study supports later memory through engagement of similarity-based mechanisms, while repeated testing does so through differentiation-based mechanisms.

The present work extends our knowledge of how MPFC represents contextual information over time. Previous work showed that differentiation in MPFC BOLD activity patterns at encoding predicts temporal order memory (Jenkins & Ranganath, 2016). MPFC activity is also enhanced during retrieval of temporally remote memories compared to recent memories (Gais et al., 2007; Nieuwenhuis & Takashima, 2011; Sterpenich et al., 2009; Takashima et al., 2006). Furthermore, multivariate classification of MPFC BOLD patterns is better for remote than recent autobiographical memories, suggesting remote memory representations are more differentiated (Bonnici et al., 2012). Our work connects these data on MPFC encoding and retrieval activity by showing that MPFC pattern differentiation is related to the fidelity of individual associative

609 memories after one week, when those associations were learned across temporal contexts. Fur-
610 thermore, our univariate analysis showing greater MPFC activity in the *overnight* compared to
611 *same day* condition is consistent with a role for MPFC in retrieval of consolidated memories and
612 extends prior work by contrasting conditions that are matched in terms of remoteness.
613 Given that we scanned only at retrieval, an open question concerns the timecourse by
614 which differentiation across memory representations emerges following study (Bonnici &
615 Maguire, 2018). Understanding the timeline will help distinguish between the influence of pro-
616 cesses engaged during encoding, retrieval or in between (Hulbert & Norman, 2015). For exam-
617 ple, future work in which the encoding phase is scanned could determine whether differences in
618 neural representation emerge during study. This work will be important to determine the relation
619 between the context-based differentiation we observed and other mechanisms such as study-
620 phase retrieval (Braun & Rubin, 1998; Thios & D'Agostino, 1976). Neural markers of memory
621 stabilization in cortical networks have also been identified shortly after encoding of a to-be-
622 remembered experience (Ben-Yakov & Dudai, 2011; Tambini et al., 2010; 2014; Tompary et al.,
623 2015). Hippocampal-cortical connectivity following a one-day delay predicts subsequent re-
624 sistance to forgetting (Vilberg & Davachi, 2013) and persistent connectivity between hippocam-
625 pus and MPFC during the one-week delay is a potential mechanism that could lead to later pat-
626 tern differentiation as representations become established in cortical networks (Peyrache et al.,
627 2009; Richards et al., 2014). Our data showing hippocampal-MPFC connectivity predicting neu-
628 ral differentiation in MPFC are consistent with this hypothesis, but future work will be needed to
629 address how such connectivity emerges and evolves over time.
630 Prominent theories propose that MPFC contributes to long-term memory by supporting
631 the integration of new information into established memory networks (Tse et al., 2007; van

Kesteren, Ruiter, Fernández, & Henson, 2012), reducing forgetting of new information by providing a set of stable associations that can be used to access specific memories during later retrieval. A computational challenge that arises when integrating new information into established networks is avoiding interference between similar representations (McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & Rudy, 2001). Our findings suggest that one way to avoid such interference is to use contextual information to differentiate similar input representations, a notion that has figured prominently in theoretical and experimental work on the role of hippocampus in episodic memory encoding (Bakker et al., 2008; Hulbert & Norman, 2015; LaRocque et al., 2013; Leutgeb et al,, 2007; Leutgeb et al., 2004; Marr, 1971; Norman & O'Reilly, 2003).Our data show that distributing study context across two days leads to differentiation in activity patterns in MPFC. The extent of this differentiation predicts the increase in memory performance that results from distributed study and is related to functional connectivity between MPFC and hippocampus. These data implicate a role for MPFC in representing contextual information in long-term memory over long timescales and provide novel evidence that distributed study leads to greater memory retention via persistent hippocampal-MPFC interactions.

Author Contributions

Conceptualization, Y.E., M.I., and L.D.; Methodology, Y.E., M.I., and L.D.; Software, Y.E. and M.I.; Formal Analysis, Y.E. and M.I.; Investigation, Y.E. and M.I.; Resources, L.D.; Writing - Original Draft, Y.E. and L.D.; Writing – Review & Editing, Y.E., M.I. and L.D.; Visualization, Y.E.; Supervision, L.D., Project Administration, L.D.; Funding Acquisition, L.D.

Acknowledgements

655 This work was supported by DART Neuroscience and NIH R01 MH074692 to L.D.

656

657 **References**

658 Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern separation in the human
659 hippocampal CA3 and dentate gyrus. *Science*, *319*(5870), 1640–1642.
660 http://doi.org/10.1126/science.1152882
661 Ben-Yakov, A., & Dudai, Y. (2011). Constructing realistic engrams: poststimulus activity of
662 hippocampus and dorsal striatum predicts subsequent episodic memory. *Journal of*
663 *Neuroscience*, *31*(24), 9032–9042.
664 Bonnici, H., Chadwick, M., Lutti, A., Hassabis, D., Weiskopf, N., & Maguire, E. (2012).
665 Detecting representatiosn of recent and remote autobiographical memories in vMPFC and
666 hippocampus. *Journal of Neuroscience*, *32*(47), 16982–16991.
667 Bonnici, H. M., & Maguire, E. A. (2017). Two years later–Revisiting autobiographical memory
668 representations in vmPFC and hippocampus. *Neuropsychologia*.
669 Bruan K., & Rubin, D. C. (1998). The spacing effect depends on an encoding deficit, retrieval,
670 and time in working memory: evidence from once-prsented words. *Memory*, *6*, 37-65.
671 Cepeda, N., Pashler, H., Vul, E., Wixted, J., & Rohrer, D. (2006). Distributed practice in verbal
672 recall tasks: a review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380.
673 Danker J.F., Tompary A., & Davachi L. (2016). Trial-by-trial hippocampal encoding activation
674 predicts the fidelity of cortical reinstatement during subsequent retrieval. *Cerebral Cortex*,
675 doi: 10.1093/cercor/bhw146.
676 Dempster, F. (1989). Spacing effects and their implications for theory and practice. *Educational*
677 *Psychology Review*, *1*(4), 309–330.
678 DuBrow, S., & Davachi, L. (2014). Temporal memory is shaped by encoding stability and inter-

679 vening item reactivation. *Journal of Neuroscience*, *34*(42), 13998-14005.
680 Ebbinghaus, H. (1964). *Memory: A Contribution to Experimental Psychology*. (H. A. Ruger & C.
681 E. Bussenius, Eds.). New York: Dover.
682 Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological*
683 *Review*, *62*(3), 145.
684 Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: pattern similarity within and
685 across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*,
686 *81*(5), 1179-1189.
687 Gais, S., Albouy, G., Boly, M., Dang-Vu, T. T., Darsaud, A., Desseilles, M., … Peigneux, P.
688 (2007). Sleep transforms the cerebral trace of declarative memories. *Proceedings of the*
689 *National Academy of Sciences*, *104*(47), 18778–18783.
690 Gordon, A., Rissman, J., Kiani, R., & Wagner, A. (2013). Cortical reinstatement mediates the
691 relationship between content-specific encoding activity and subsequent recollection
692 decisions. *Cerebral Cortex*, *24*(12), 3350–3364. http://doi.org/doi: 10.1093/cercor/bht194
693 Guise, KG., & Shapiro, ML. (2017) Medial prefrontal cortex reduces memory
694 interference by modifying hippocampal encoding. *Neuron* 94(1) (2017): 183-192.
695 https://doi.org/10.1016/j.neuron.2017.03.011
696 Heusser, A. C., Poeppel, D., Ezzyat, Y., & Davachi, L. (2016). Episodic sequence memory is
697 supported by a theta-gamma phase code. *Nature Neuroscience*, 19, 1374-1380.
698 Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context.
699 *Journal of Mathematical Psychology*, *46*(3), 269–299.
700 Howard, M. W., Shankar, K. H., Aue, W. R., & Criss, A. H. (2015). A distributed representation
701 of internal time. *Psychological review*, *122*(1), 24.

Hsieh, L. T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron*, *81*(5), 1165-1178.

Hulbert, J. C., & Norman, K. A. (2015). Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cerebral Cortex*, *25*(10), 3994–4008. http://doi.org/10.1093/cercor/bhu284

Hyman, J. M., Ma, L., Balaguer-Ballester, E., Durstewitz, D., & Seamans, J. K. (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proceedings of the National Academy of Sciences*, *109*(13), 5086-5091.

Insausti, R., Juottonen, K., Soininen, H., Insausti, A., Partanen, K., Vainio, P., … Pitkanen, A. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *American Journal of Neuroradiology*, *19*, 659–671.

Jenkins, L. J., & Ranganath, C. (2016). Distinct neural mechanisms for remembering when an event occurred. *Hippocampus, 26, 554-559*.

Karlsson Wirebring, L., Wiklund-Hornqvist, C., Eriksson, J., Andersson, M., Jonsson, B., & Nyberg, L. (2015). Lesser neural pattern similarity across repeated tests is associated with better long-term memory retention. *Journal of Neuroscience*, *35*(26), 9595–9602. http://doi.org/10.1523/JNEUROSCI.3550-14.2015

Kitamura, T., Ogawa, S. K., Roy, D. S., Okuyama, T., Morrissey, M. D., Smith, L. M., ... & Tonegawa, S. (2017). Engrams and circuits crucial for systems consolidation of a memory. *Science*, *356*(6333), 73-78. http://doi.org/10.1126/science.aam6808

Kuhl, B. A., Rissman, J., & Wagner, A. D. (2012). Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia*, *50*(4), 458–69. http://doi.org/10.1016/j.neuropsychologia.2011.09.002

LaRocque, K. F., Smith, M. E., Carr, V. A., Witthoft, N., Grill-Spector, K., & Wagner, A. D. (2013). Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. *Journal of Neuroscience*, *33*(13), 5466–74. http://doi.org/10.1523/JNEUROSCI.4293-12.2013

Lesburguères, E., Gobbo, O. L., Alaux-Cantin, S., Hambucken, A., Trifilieff, P., & Bontempi, B. (2011). Early tagging of cortical networks is required for the formation of enduring associative memory. *Science*, *331*(6019), 924–8. http://doi.org/10.1126/science.1196164

Leutgeb, J. K., Leutgeb, S., Moser, M.-B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science*, *315*(5814), 961–966. http://doi.org/10.1126/science.1135801

Leutgeb, S., Leutgeb, J. K., Treves, A., Moser, M. B., & Moser, E. I. (2004). Distinct Ensemble Codes in Hippocampal Areas CA3 and CA1. *Science*, *305*(5688), 1295–1298. http://doi.org/10.1126/science.1100265

Litman, L., & Davachi, L. (2008). Distributed learning enhanced relational memory consolidation. *Learning & Memory*, *15*, 711–716.

Long, N. M., Sperling, M. R., Worrell, G. A., Davis, K. A., Gross, R. E., Lega, B. C., ... & Kahana, M. J. (2017). Contextually Mediated Spontaneous Retrieval Is Specific to the Hippocampus. *Current Biology*, *27*(7), 1074-1079.

Mankin, E. A., Sparks, F. T., Slayyeh, B., Sutherland, R. J., Leutgeb, S., & Leutgeb, J. K. (2012). Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, *109*(47), 19462–19467.

Manns, J. R., Howard, M. W., & Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, *56*(3), 530-540.

748 Marr, D. (1971). Simple memory: a theory for archicortex. *Philosophical Transactions of the*
749 *Royal Society of London. Series B, Biological Sciences*, *262*(841), 23–81.
750 http://doi.org/10.1098/rstb.1971.0078
751 McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary
752 learning systems in the hippocampus and neocortex: Insights from the successes and
753 failures of connessionist models of learning and memory. *Psychological Rewiew*, *102*(3),
754 419–457.
755 McGeoch, J., & Irion, A. (1952). The distribution of practice and reminiscence. In *The*
756 *Psychology of Human Learning* (2nd ed., pp. 138–193). New York: David McKay.
757 Nieuwenhuis, I. L. C., & Takashima, A. (2011). The role of the ventromedial prefrontal cortex in
758 memory consolidation. *Behavioural Brain Research*, *218*(2), 325–334.
759 http://doi.org/10.1016/j.bbr.2010.12.009
760 Norman, K., & O'Reilly, R. (2003). Modeling hippocampal and neocortical contributions to
761 recognition memory: a complementary-learning-systems approach. *Psychological Review*,
762 *110*(4), 611–646.
763 O'Reilly, R. C., & Rudy, J. W. (2001). Conjunctive representations in learning and memory:
764 principles of cortical and hippocampal function. *Psychological Review*, *108*(2), 311–345.
765 Oudiette, D., Antony, J. W., Creery, J. D., & Paller, K. A. (2013). The Role of Memory
766 Reactivation during Wakefulness and Sleep in Determining Which Memories Endure.
767 *Journal of Neuroscience*, *33*(15), 6672–6678. http://doi.org/10.1523/JNEUROSCI.5497-
768 12.2013
769 Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S. I., & Battaglia, F. P. (2009). Replay of
770 rule-learning related neural patterns in the prefrontal cortex during sleep. *Nature*

Neuroscience, *12*(7), 919–926. http://doi.org/10.1038/nn.2337

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009a). Task context and organization in free recall. *Neuropsychologia*, *47*(11), 2158-2163.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009b). A context maintenance and retrieval model of organizational processes in free recall. *Psychological review*, *116*(1), 129-156.

Poppenk, J., Evensmoen, H.R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. Trends in Cognitive Sciences, *17*, 230-240.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, *13*(10), 713.

Richards, B. A, Xia, F., Santoro, A., Husse, J., Woodin, M. A, Josselyn, S. A, & Frankland, P. W. (2014). Patterns across multiple memories are identified over time. *Nature Neuroscience*, *17*(7), 981–6. http://doi.org/10.1038/nn.3736

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*, *23*(2), 752–763. http://doi.org/10.1016/j.neuroimage.2004.06.035

Ritchey, M., Wing, E. A., Labar, K. S., & Cabeza, R. (2012). Neural Similarity Between Encoding and Retrieval is Related to Memory Via Hippocampal Interactions. *Cerebral Cortex*, *23*(12), 2818–2828.

Ritchey, M., Montchal, M. E., Yonelinas, A. P., & Ranganath, C. (2015). Delay-dependent contributions of medial temporal lobe regions to episodic memory retrieval. *Elife*, *4*, e05025.

Ruch, T. (1928). Factors influencing the relative economy of massed and distributed practice in learning. *Psychological Review*, *35*(1), 19–45.

794 Schlichting, M. L., & Preston, A. R. (2016). Hippocampal–medial prefrontal circuit supports
795 memory updating during learning and post-encoding rest. *Neurobiology of learning and*
796 *memory*, *134*, 91-106.
797 Siapas, A. G., Lubenov, E. V., & Wilson, M. A. (2005). Prefrontal phase locking to hippocampal
798 theta oscillations. *Neuron*, *46*(1), 141-151.
799 Siapas, A. G., & Wilson, M. A. (1998). Coordinated interactions between hippocampal ripples
800 and cortical spindles during slow-wave sleep. *Neuron*, *21*(5), 1123–1128.
801 http://doi.org/10.1016/S0896-6273(00)80629-7
802 Singer, A. C., & Frank, L. M. (2009). Rewarded Outcomes Enhance Reactivation of Experience
803 in the Hippocampus. *Neuron*, *64*(6), 910–921. http://doi.org/10.1016/j.neuron.2009.11.016
804 Sterpenich, V., Albouy, G., Darsaud, A., Schmidt, C., Vandewalle, G., Dang Vu, T. T., …
805 Maquet, P. (2009). Sleep promotes the neural reorganization of remote emotional memory.
806 *Journal of Neuroscience*, *29*(16), 5143–5152.
807 Takashima, A., Nieuwenhuis, I., Jensen, O., Talamini, L., Rijpkema, M., & Fernandez, G.
808 (2009). Shift from hippocampal to neocrotical centered retrieval network with
809 consolidation. *Journal of Neuroscience*, *29*(32), 10087–10093.
810 Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., …
811 Fernández, G. (2006). Declarative memory consolidation in humans: a prospective
812 functional magnetic resonance imaging study. *Proceedings of the National Academy of*
813 *Sciences*, *103*(3), 756–761. http://doi.org/10.1073/pnas.0507774103
814 Tambini, A., Ketz, N., & Davachi, L. (2010). Enhanced brain correlations during rest are related
815 to memory for recent experiences. *Neuron*, *65*(2), 280–290.
816 Thios S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase

retrieval. *Journal of Verbal Learning and Verbal Behavior*, *15*, 529-536.

Tompary, A., Duncan, K., & Davachi, L. (2015). Consolidation of Associative and Item Memory Is Related to Post-Encoding Functional Connectivity between the Ventral Tegmental Area and Different Medial Temporal Lobe Subregions during an Unrelated Task. *Journal of Neuroscience*, *35*(19), 7326–31.

Tompary, A,. Duncan, K., & Davachi, L. (2016). High-resolution investigation of memory-specific reinstatement in the hippocampus and perirhinal cortex. *Hippocampus*, 26(8), 995-1007.

Tompary, A., & Davachi, L. (2017). Consolidation Promotes the Emergence of Representational Overlap in the Hippocampus and Medial Prefrontal Cortex. *Neuron*, *96*(1), 228-241.

Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., … Morris, R. (2007). Schemas and memory consolidation. *Science*, *316*(5821), 76–82.

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., … Morris, R. G. M. (2011). Schema-dependent gene activation and memory encoding in neocortex. *Science*. http://doi.org/10.1126/science.1205274

van Kesteren, M. T. R., Fernández, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences*, *107*(16), 7550–5.

van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*. http://doi.org/10.1016/j.tins.2012.02.001

Vilberg, K., & Davachi, L. (2013). Perirhinal-hippocampal connectivity during reactivation is a

marker for object-based memory consolidation. *Neuron*, *79*(6), 1232–1242.

Ward, E. J., Chun, M. M., & Kuhl, B. A. (2013). Repetition suppression and multi-voxel pattern similarity differentially track implicit and explicit visual memory. *Journal of Neuroscience*, *33*(37), 14749–57. http://doi.org/10.1523/JNEUROSCI.4889-12.2013

Watanabe, T., Hirose, S., Wada, H., Katsura, M., Chikazoe, J., Jimura, K., … Konishi, S. (2011). Prediction of subsequent recognition performance using brain activity in the medial temporal lobe. *NeuroImage*, *54*(4), 3085–3092.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Mölle, M., & Born, J. (2011). Sleep selectively enhances memory expected to be of future relevance. *The Journal of Neuroscience*, *31*(5), 1563–1569. http://doi.org/10.1523/JNEUROSCI.3575-10.2011

Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., & Poldrack, R. A. (2010). Greater neural pattern similarity across repetitions is associated with better memory. *Science*, *330*(6000), 97–101. http://doi.org/10.1126/science.1193125

Xue, G., Dong, Q., Chen, C., Lu, Z.-L., Mumford, J. A., & Poldrack, R. A. (2013). Complementary role of frontoparietal activity and cortical pattern similarity in successful episodic memory encoding. *Cerebral Cortex*, *23*(7), 1562–1571. http://doi.org/10.1093/cercor/bhs143

Zeithamova, D., & Preston, A. R. (2010). Flexible memories: differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *Journal of Neuroscience*, *30*(44), 14676–84. http://doi.org/10.1523/JNEUROSCI.3250-10.2010

Figure Captions

Figure 1. Study design.
(A) General Procedures for the experiment. *Day 1*: Participants were trained to criterion on a set of word-object associations on Day 1 (*overnight* condition). *Day 2*: Twenty-four hours later participants returned to the lab and were trained on a new set of word-object associations (*same day* condition) before being re-trained to criterion on both the *overnight* and *same day* lists intermixed. *Day 7*: One week after the *Day 1* session, participants were trained to criterion on a novel set of word-object associations (*single session*) before being placed in the fMRI scanner and performing a retrieval test that included all lists (*overnight*, *same day* and *single session*) and novel lures.
(B) *Left*, Each training session began with a Study Phase in which participants were exposed to the word-object associations. *Middle*, Training then began with a test in which word-object association memory was tested by having participants respond 'Natural' or 'Manmade' for the object that had been associated with each word (*Nat*, Natural; *Man*, Manmade; *Uns*, Unsure). After each response, feedback ('Correct!/Incorrect!') was given along with presentation of the associated object. A correct response was required in two consecutive training rounds in order for the word-object pair to be dropped from future training. This training protocol was performed for three blocks of 24 stimuli until all 72 pairs from the list were learned. Following the final block of 24, all 72 words were then re-presented for a final test in order to ensure that participants were above 90% across the entire list. Participants who did not meet the 90% criterion on this final test were re-exposed to all 72 pairs in a final Study Phase. *Right*: In the scanner participants performed the final Retrieval Test for all of the words from the *overnight*, *same day* and *single session* lists intermixed with novel lures.

Figure 2. Behavioral data for day 7 retrieval test.
(A) The proportions of *associative*, *item-only* and *forgotten* pairs are plotted for the *overnight*, *same day*, and *single session* lists. *Associative* memory was higher for *overnight* pairs compared to *same day* pairs ($p < 10^{-4}$), evidence of a distributed learning effect. *Associative* memory for *single session* pairs was higher than both *overnight* and *same day* pairs (both pairwise comparisons $p < 10^{-4}$).
(B) Response times for *associative* and *item-only* conditions. In addition to showing higher source memory, participants were also faster to respond for *overnight associative* memories compared to *same day associative* memories ($p < 10^{-5}$). Participants were overall fastest for *single session associative* memories (both pairwise comparisons with *overnight* and *same day* trials, $p < 10^{-6}$). Error bars denote SEM.

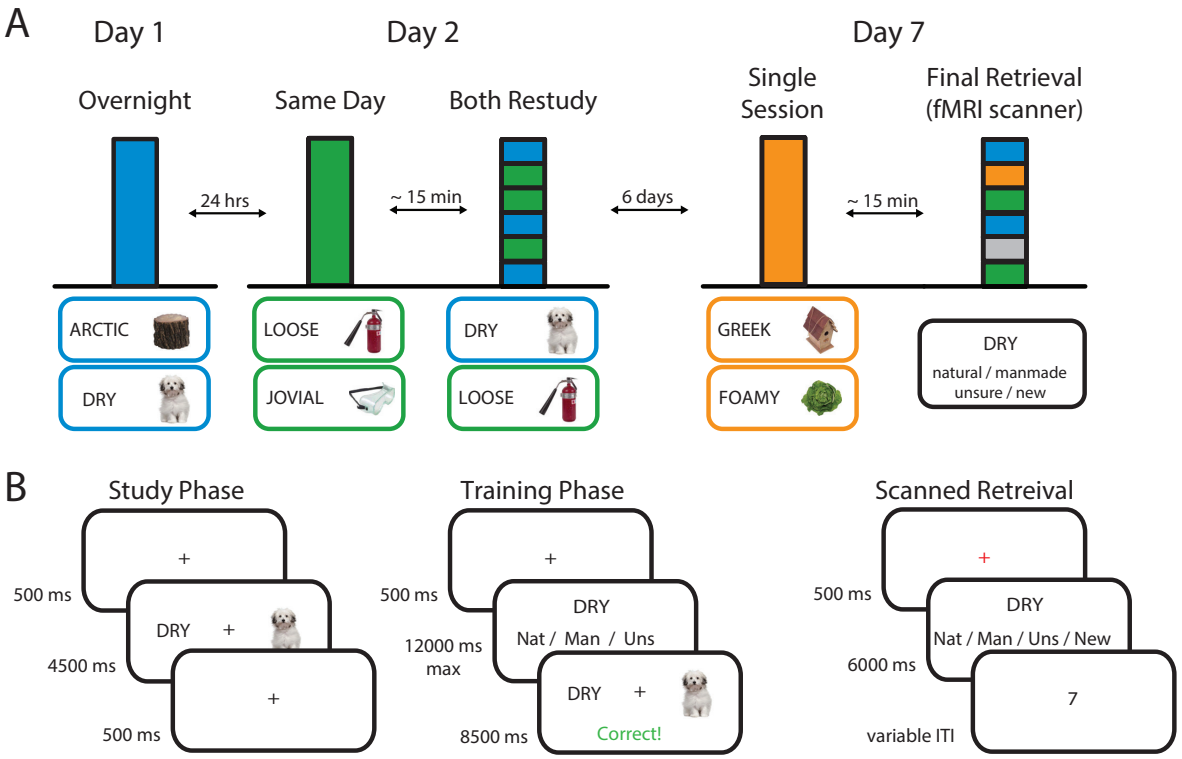
Figure 3. Hippocampal pattern differentiation.
(A) *Left*, example subject left hippocampus anatomical region of interest. *Middle*, in left hippocampus, differentiation did not vary between *overnight* and *same day* memory ($p = 0.12$) but was significantly higher for *associative* compared to *item-only* memory (main effect: $p = 0.004$), a difference that was significant within the *overnight* and *same day* conditions individually. *Right*, the difference in *overnight*/*same day* pattern differentiation in hippocampus did not correlate with the overnight memory benefit ($p = 0.23$).
(B) *Left*, in left hippocampus pattern differentiation was significantly greater in anterior compared to posterior regions ($p = 0.02$). *Right*, in right hippocampus, pattern differentiation was al-

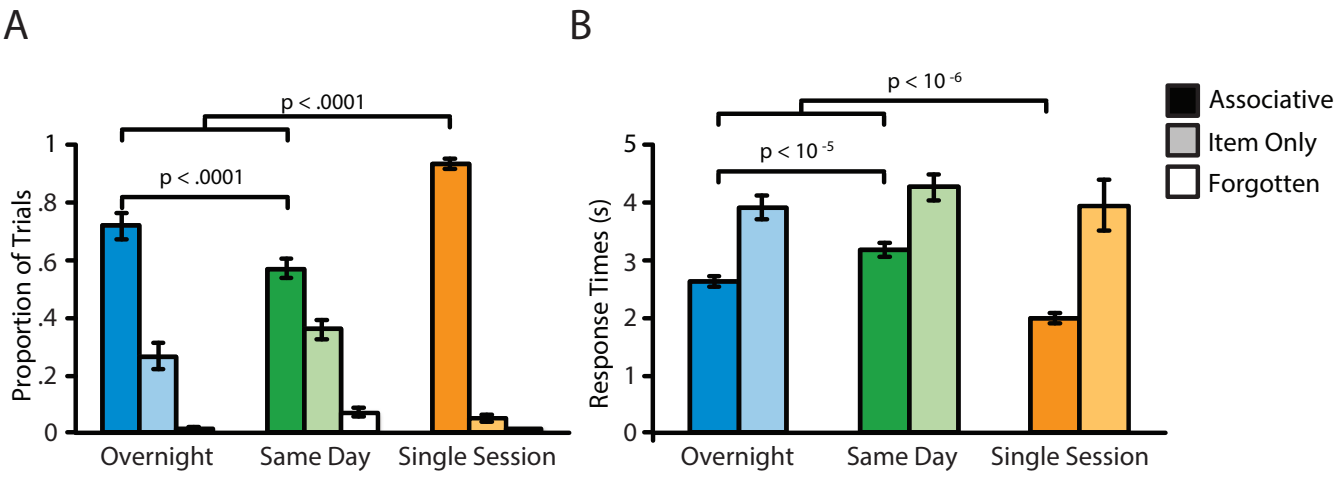
so significantly greater for anterior compared to posterior regions ($p = 0.006$).

Figure 4. MPFC and brain-wide pattern differentiation
(A) *Left*, MPFC region of interest rendered on the group-average brain. *Middle*, in MPFC, there was significantly more differentiation for *associative* versus *item-only* memories only within the *overnight* condition ($p = 0.006$) but not in the *same day* condition ($p > 0.16$; interaction $p = 0.02$). Differentiation for *overnight associative* trials was also significantly greater than for *same day associative* memories ($p = 0.002$). Error bars denote SEM. ⊗ indicates significant interaction ($p < 0.05$). *Right,* the difference in *overnight*/*same day* pattern differentiation on Day 7 correlated with the overnight memory benefit [$r(15) = 0.53$, $p = 0.03$].
(B) *Left*, example subject brain-wide task active mask. *Middle*, pattern differentiation was greater for overnight than same day retrieval ($p < 0.02$), but did not differ as a function of memory success. Error bars denote SEM. *Right,* the difference in *overnight*/*same day* pattern differentiation on Day 7 was not correlated with the overnight memory benefit [$r(15) = 0.03$, $p = 0.88$].

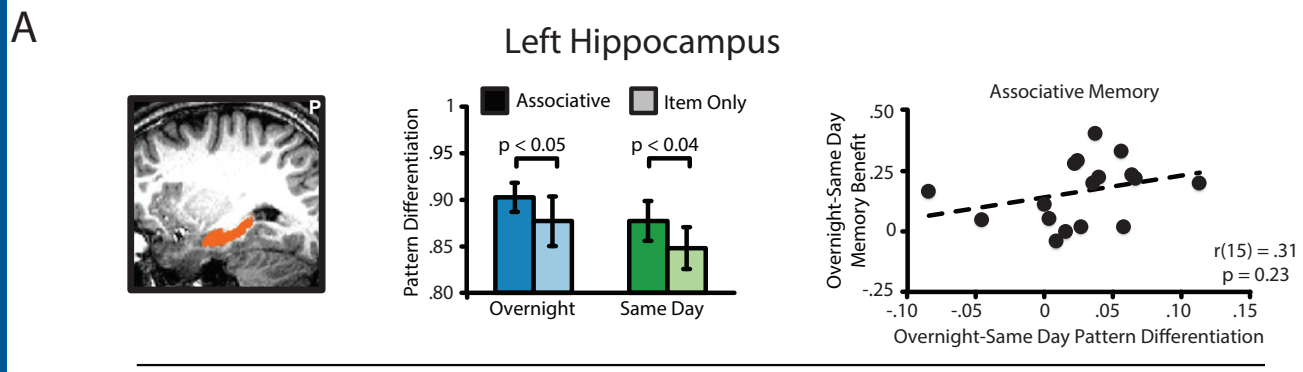
Figure 5. Hippocampal-MPFC connectivity predicts MPFC pattern differentiation.
The difference in hippocampal-MPFC connectivity between the *overnight-same day* conditions correlates with the level of pattern differentiation in MPFC [$r(15) = 0.60$, $p < 0.02$].

Figure 6. Univariate activity in hippocampus and MPFC.
A, we did not observe significant modulation of univariate activity in left hippocampus across conditions. (all $p > 0.12$). *B*, in MPFC, there was a main effect of overnight/same day condition ($p = 0.01$), as well as an interaction ($p = 0.02$). Errorbars denote SEM. ⊗ indicates significant interaction ($p < 0.05$).

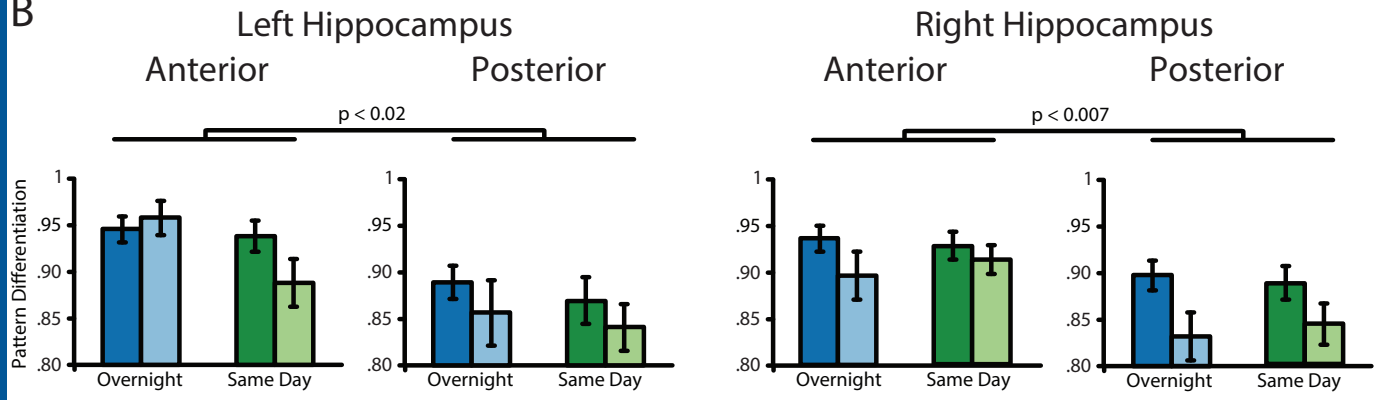




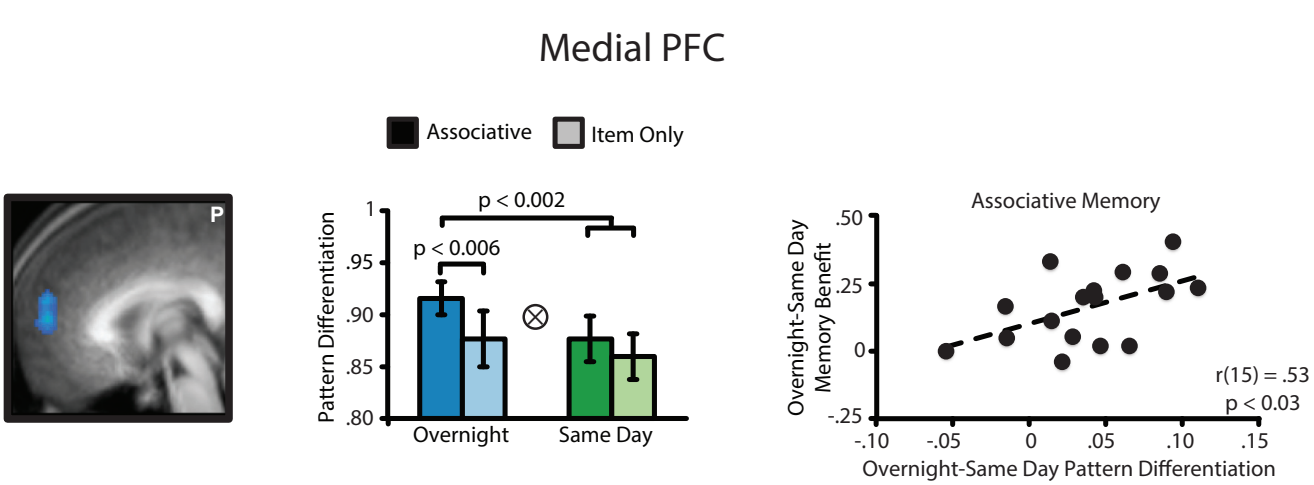
A



B



A



B