



# Time-resolved neural reinstatement and pattern separation during memory decisions in human hippocampus

Lynn J. Lohnas<sup>a</sup>, Katherine Duncan<sup>b</sup>, Werner K. Doyle<sup>c</sup>, Thomas Thesen<sup>d,e</sup>, Orrin Devinsky<sup>e</sup>, and Lila Davachi<sup>f,g,1</sup>

<sup>a</sup>Department of Psychology, New York University, New York, NY 10003; <sup>b</sup>Department of Psychology, University of Toronto, Toronto, ON M5S 3G3, Canada; <sup>c</sup>Department of Neurosurgery, New York University School of Medicine, New York, NY 10016; <sup>d</sup>Department of Physiology, Neuroscience, and Behavioral Sciences, St. George's University School of Medicine, St. George, Grenada, West Indies; <sup>e</sup>Department of Neurology, Epilepsy Division, New York University School of Medicine, New York, NY 10016; <sup>f</sup>Department of Psychology, Columbia University, New York, NY 10027; and <sup>g</sup>Department of Biomedical Imaging & Neuromodulation, Nathan Kline Institute, Orangeburg, NY 10962

Edited by Larry R. Squire, University of California, San Diego, La Jolla, CA, and approved June 20, 2018 (received for review October 3, 2017)

**Mnemonic decision-making has long been hypothesized to rely on hippocampal dynamics that bias memory processing toward the formation of new memories or the retrieval of old ones. Successful memory encoding may be best optimized by pattern separation, whereby two highly similar experiences can be represented by underlying neural populations in an orthogonal manner. By contrast, successful memory retrieval is thought to be supported by a recovery of the same neural pattern laid down during encoding. Here we examined how hippocampal pattern completion and separation emerge over time during memory decisions. We measured electrocorticography activity in the human hippocampus and posterior occipitotemporal cortex (OTC) while participants performed continuous recognition of items that were new, repeated (old), or highly similar to a prior item (similar). During retrieval decisions of old items, both regions exhibited significant reinstatement of multivariate high-frequency activity (HFA) associated with encoding. Further, the extent of reinstatement of encoding patterns during retrieval was correlated with the strength (HFA power) of hippocampal encoding. Evidence for encoding pattern reinstatement was also seen in OTC on trials requiring fine-grained discrimination of similar items. By contrast, hippocampal activity showed evidence for pattern separation during these trials. Together, these results underscore the critical role of the hippocampus in supporting both reinstatement of overlapping information and separation of similar events.**

similar environments, the place cell location remained the same, but the firing rates differed between chambers (14). In a similar way, Knierim and coworkers (15, 16) examined place cell activity between two environments with global and local cues rotated and found that subsets of place cell locations were consistent with either the rotated local or global cue rotations, thus keeping track of the original environment location as well as the rotated cue of the new environment. However, place cell studies cannot answer whether, on a cognitive level, the rodent is successful in recognizing the similarities and differences across two environments. In addition, place cell studies typically record activity over more extended periods of time and multiple visits to the same location and thus are unable to answer how pattern completion and separation emerge upon the second encounter with a similar or identical location. One study began to address this latter question using a context-dependent associative reward-learning task (17). They found that different subsets of hippocampal cells exhibited firing rates that distinguished between the context (i.e., the spatial environment) and an item's identity, position, or valence. Nonetheless, these hippocampal responses were recorded following an initial learning phase and thus do not capture hippocampal dynamics during the initial phase of distinguishing between similar memories.

hippocampus | electrocorticography | memory reinstatement | pattern separation | mnemonic discrimination

Perhaps one of the most challenging functions of the episodic memory system is distinguishing between two experiences that contain highly overlapping content. Pattern separation refers to the process of representing highly similar events in a distinct way, thus allowing them to coexist with minimal interference (1–6). However, if two distinct experiences share overlapping features, then the second experience may also promote “pattern completion” of the first experience (2, 3, 7). These two processes—pattern completion and separation—reflect opposing if not contradictory functions, both of which have been attributed to the hippocampus. To reconcile how the hippocampus can accomplish both processes, it has been proposed that novelty may bias the hippocampal system toward pattern separation, while familiarity may promote memory retrieval, or pattern completion (8–11). How and when the hippocampus can support representations of both overlapping and distinctive features of events has been an active area of research. Nonetheless, the temporal dynamics examining how these processes emerge over the course of a single memory decision remains relatively unexplored.

Theoretical and rodent work provide evidence that the hippocampus exhibits sensitivity to differences in highly similar events while also representing the events' strong overlap (2, 3, 12, 13). For instance, when comparing place cell firing across two

## Significance

**One of the biggest computational challenges the memory system faces is to disambiguate highly similar experiences while at the same time preserving and reinstating prior memories. Remarkably, hippocampal processes have been implicated in both of these functions. However, how this is accomplished is unknown. Leveraging the spatiotemporal resolution of electrocorticography, we found evidence for memory reinstatement in both the hippocampus and occipitotemporal cortex. Interestingly, when a current experience was very similar but not identical to a prior one, occipitotemporal cortical activity still showed reinstatement of the prior memory, but hippocampal activity differentiated or disambiguated these two similar experiences.**

Author contributions: K.D. and L.D. designed research; K.D., W.K.D., T.T. and O.D. performed research; L.J.L. analyzed data; and L.J.L., K.D. and L.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: The behavioral and electrophysiological data reported here are deposited with the National Institute of Mental Health Data Archive at <https://data-archive.nimh.nih.gov>.

<sup>1</sup>To whom correspondence should be addressed. Email: [ld24@columbia.edu](mailto:ld24@columbia.edu).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1717088115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1717088115/-DCSupplemental).

Published online July 13, 2018.

In humans, multivariate approaches have been used to examine mnemonic reinstatement effects for similar and identical stimuli, both with functional magnetic resonance imaging (fMRI) (18) and electrocorticography (ECoG). In the hippocampus, there is evidence for reinstatement of encoding patterns during successful later retrieval both for similar stimuli from the same category (19) and for identical stimuli (9, 20–22). However, such effects are not unique to the hippocampus. Mnemonic reinstatement in several cortical regions has been shown with ECoG (23–27) and has been noted in visual cortex (28, 29), medial temporal lobe cortex (9, 20, 30, 31), and prefrontal cortex (28) using fMRI.

At a mechanistic level, cortical reinstatement during retrieval may result from or interact with hippocampal pattern completion (32, 33). Supporting this idea, cortical reinstatement has been shown to correlate with hippocampal univariate activity at encoding (29, 34) and at retrieval (9, 20, 22, 28, 31, 35–37). Thus, hippocampal computations may recover the memory representations formed during encoding, and this may, in turn, support cortical reinstatement. However, the fMRI response, on the time scale of seconds, is not well suited to address how reinstatement emerges over time and across regions, and to our knowledge, no ECoG study has contrasted reinstatement with separation in the hippocampus.

While there is strong evidence that both cortical regions and the hippocampus contribute to memory reinstatement, multivariate fMRI studies have uniquely implicated the hippocampus in supporting pattern separation of highly similar experiences. In particular, there is evidence that hippocampal pattern separation is greater for very similar item pairs compared with unrelated item pairs (4–6, 30). In addition, studies of univariate fMRI activity provide evidence that hippocampal activation is sensitive to whether a presented item is either identical or just highly similar to a previously presented item (30, 38–45). In one seminal example (38), unlike other hippocampal subregions or medial temporal lobe cortical regions, hippocampal subregion DG/CA3 did not show repetition suppression for highly similar lure items; rather, activity was not significantly different between lures and new items. These results have been interpreted as evidence that the DG/CA3 subregion of the hippocampus plays a unique and critical role in distinguishing highly overlapping memory representations (38), consistent with prior theoretical work (2, 13) and rodent work (46).

Taken together, there is accumulating evidence that hippocampal processes contribute both to memory reinstatement, or pattern completion, and to pattern separation. However, it is not understood how these distinct operations are orchestrated in time over the course of a memory decision. In particular, it is not known whether reinstatement and separation occur on similar time scales in the hippocampus, nor whether reinstatement occurs on similar time scales in cortical regions as in the hippocampus. Furthermore, it remains unclear how attentional focus on the overlapping or distinctive features of similar stimuli modulates neural reinstatement or separation (47). To address these questions, we took advantage of the spatiotemporal resolution of ECoG activity, comparing activity contributing to separation and reinstatement within individual trials and across different regions. Specifically, we recorded depth and surface cortical ECoG activity as participants performed a continuous recognition paradigm previously used to examine fine-grained mnemonic discrimination (3, 8, 10, 48). In each participant we examined dynamics of high-frequency activity (HFA; 45–115 Hz), an established correlate of firing rates of individual neurons (49–51) and of fMRI activity (52, 53). To directly address questions about pattern separation and completion, in addition to univariate measurements, we adopted a multivariate pattern similarity approach to measure the overlap between neural representations of presented stimuli.

We examined both univariate and multivariate HFA measures in the hippocampus, posterior occipitotemporal cortex (OTC), and dorsolateral prefrontal cortex (DLPFC). We specifically considered regions both upstream (OTC) and downstream (DLPFC)

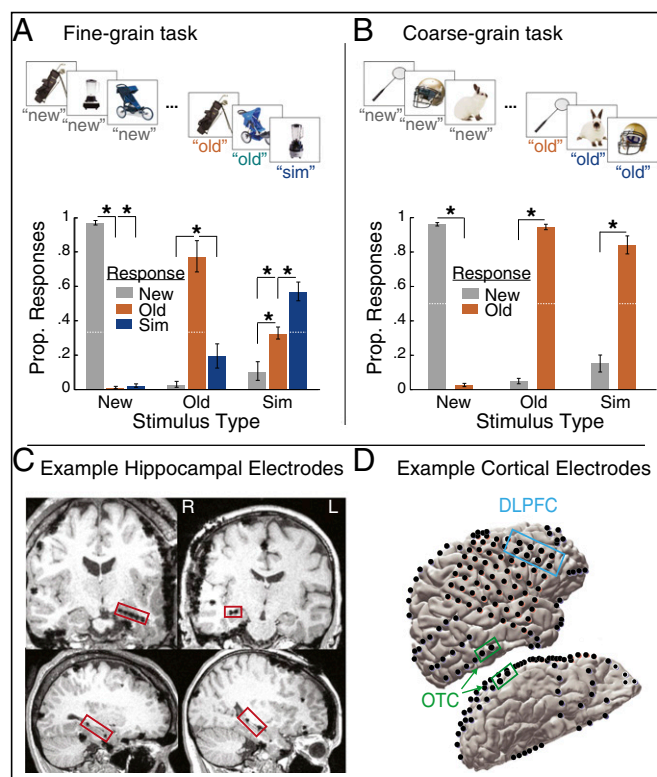
from the hippocampus to characterize the timing of the flow of mnemonic activity patterns across regions. We chose OTC because visual object processing regions may be sensitive to the perceptual similarity between a visually presented item and its similar prior presentation. We chose DLPFC because, much like the hippocampus, this region is necessary for maintenance of contextual aspects of episodic memory (54, 55) and studies of associative memory have consistently noted that DLPFC and hippocampal regions exhibit activity modulated by successful encoding (56–61) and successful retrieval (23, 42, 62, 63). Our experimental set-up additionally allowed us to address the extent to which reinstatement and separation are modulated by task demands. Across two blocks, participants viewed a series of objects, which could be new, repeated, or highly similar but not identical to a previously presented object. In the fine-grain task block (Fig. 1A), participants classified objects as new, old, or similar; distinguishing between the latter two categories required a fine-grained mnemonic discrimination based on the visual features of the stimuli. By contrast, in the coarse-grain task block participants were instructed to classify old and similar items as “old” (Fig. 1B).

## Results

**Behavior.** Fig. 1A and B shows the proportion of responses as a function of stimulus type. In both tasks, memory performance was above chance (dashed white lines) for all stimulus types [fine-grain task: chance = 0.33; new:  $P < 0.0001$ ,  $t(4) = 54.5$ , old:  $P = 0.0057$ ,  $t(4) = 5.39$ , similar:  $P = 0.0086$ ,  $t(4) = 4.80$ ; coarse-grain task: chance = 0.50; new:  $P < 0.0001$ ,  $t(4) = 53.3$ , old:  $P < 0.0001$ ,  $t(4) = 31.6$ , similar:  $P = 0.0019$ ,  $t(4) = 7.29$ ]. Further, in the coarse-grain task, participants were more likely to classify new items as new than as old [ $P < 0.0001$ ,  $t(4) = 54.7$ ] and old items as old than new [ $P < 0.0001$ ,  $t(4) = 32.1$ ]. Similarly, in the fine-grain task, participants were more likely to classify new items correctly than to classify them as old [ $P < 0.0001$ ,  $t(4) = 62.5$ ] or similar [ $P < 0.0001$ ,  $t(4) = 49.7$ ]. Participants were also accurate in their responses to old items in the fine-grain task, classifying them significantly more often as old items than as new [ $P = 0.0015$ ,  $t(4) = 7.74$ ] or as similar [ $P = 0.0158$ ,  $t(4) = 4.02$ ].

As expected, in the fine-grain task, accuracy was significantly lower for similar items in comparison with new items [ $P = 0.0016$ ,  $t(4) = 7.58$ ] and with old items [ $P = 0.0327$ ,  $t(4) = 3.21$ ]. However, in the coarse-grain task, the correct classification of similar items (as old) was not lower than correct classification of new items [ $P = 0.0638$ ,  $t(4) = 2.54$ ] or old items [ $P = 0.0623$ ,  $t(4) = 2.56$ ]. This reflects the fact that in the coarse-grain task, participants did not have to discriminate between old and similar items. Indeed, if a similar item was misclassified in the fine-grain task, it was more likely to be misclassified as old than new [ $P = 0.0253$ ,  $t(4) = 3.48$ ], further suggesting that errors for similar items in the fine-grain task primarily arose from an inability to discriminate whether the similar item was the same or slightly different from its corresponding original presentation, rather than an inability to recognize that a version of the stimulus was presented previously. Further supporting the notion that additional mnemonic discrimination was needed and deployed in the fine-grain task, response times were significantly faster in the coarse-grain task compared with the fine-grain task [ $P = 0.0036$ ,  $t(4) = 6.13$ ; fine-grain mean  $\pm$  SEM = 1.60 s  $\pm$  0.14; coarse-grain mean  $\pm$  SEM = 1.37 s  $\pm$  0.11; see also *SI Appendix, Fig. S1*].

**Memory-Related Differences in Univariate ECoG HFA.** Our first set of analyses focused on correct trials only, across conditions and regions. HFA in both hippocampus and OTC exhibited significant time-sensitive responses in the fine-grain task (Fig. 2). To examine the temporal dynamics of memory processing in each region, we examined differences in HFA divided into 500-ms time bins. First, we compared HFA for new and old presentations of a stimulus (64–67). In the hippocampus, HFA was significantly



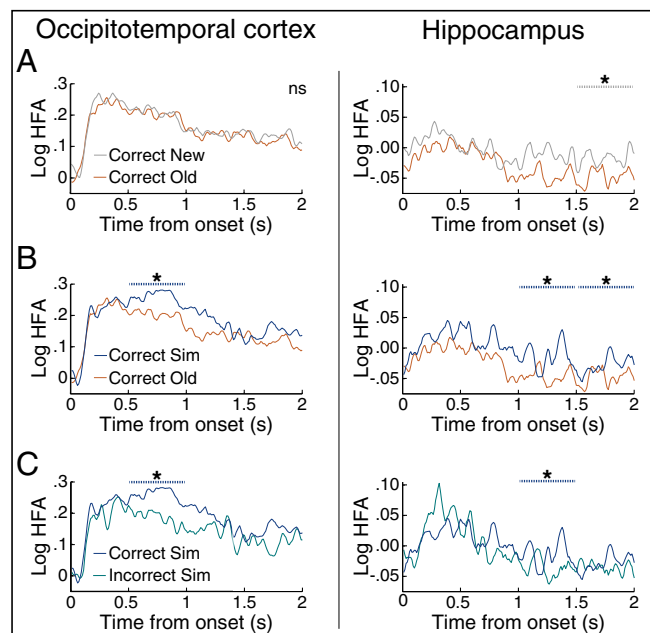
**Fig. 1.** Experiment design, behavioral results, and regions of interest. (A) In the fine-grain task block, participants viewed images one at a time and had to distinguish between new items (new), exact repeats (old), and items that were similar but not identical to a previous presentation (similar). (Top) The four stimulus types considered in the electrophysiological data are shown here: correct new items, correct old items, correct similar items, and similar items incorrectly classified as old. (Bottom) Proportion of responses (Prop. Responses) by stimulus type in the fine-grain task. (B) In the coarse-grain task block, participants viewed new, old, and similar items but classified both similar and old items as old, thus not requiring as fine-grained discrimination between these latter two stimulus types. (Top) The three stimulus types examined in the electrophysiological data are shown here: correct new items, correct old items, and correct similar items. (Bottom) Memory performance by stimulus and response type in the coarse-grain task. (C) Hippocampal electrode placements in the hippocampus for participant 5 (Left) and participant 1 (Right). Hippocampal electrodes were visualized using each participant's postoperative magnetic resonance imaging scan. (D) Electrode placements in posterior OTC and DLPCF for participant 1. Cortical surface electrode placements were visualized on each participant's rendered 3D brain (92). Error bars indicate mean  $\pm$  SEM across participants. \* $P < 0.05$ .  $n = 5$ .

greater for new items than old items during the later 1.5–2-s time window (Fig. 2A;  $P = 0.016$ , actual  $Z = 20.4$ , null mean  $Z = -0.283$ ). By contrast, HFA in OTC did not exhibit significant differences between old and new items.

We next asked whether HFA distinguishes between old items and similar items (correct sim), arguably one of the most challenging components of this task as similar items place more demands on mnemonic discrimination between the visual features of the current stimulus and one retrieved from memory. We found that HFA in both hippocampus and OTC was significantly greater for similar compared with old items (Fig. 2B). Interestingly, in OTC, this effect emerged during the early 0.5–1-s time bin ( $P < 0.001$ , actual  $Z = 15.8$ , null mean  $Z = 0.005$ ), whereas in hippocampus this effect occurred later and lasted longer, through 1–1.5 s and 1.5–2 s (1–1.5 s:  $P = 0.032$ , actual  $Z = 16.7$ , null mean  $Z = 0.289$ ; 1.5–2 s:  $P = 0.048$ , actual  $Z = 17.1$ , null mean  $Z = 0.136$ ). Moreover, hippocampal HFA was significantly greater for new items than old items during the same time

window as when HFA was greater for similar items than old items, with no significant difference between similar and new items ( $P > 0.3$ ). This pattern of activity, also seen in prior fMRI work (3, 38, 48, 68), is consistent with the notion that old items evoke stronger repetition suppression than highly similar lures. Elevated hippocampal HFA for similar relative to old items, occurring at a relatively late time point, may reflect attention to and/or encoding of the novel details of the similar items, thus allowing mnemonic resolution between old and similar items, consistent with the unique role of the hippocampus in separation. However, we see evidence that processing was necessary for similar items beyond novelty detection, as HFA for similar and new items is not identical in all time windows: HFA for similar items was significantly greater than new items during 0.5–1 s in both OTC ( $P < 0.001$ , actual  $Z = 15.1$ , null mean  $Z = -0.133$ ) and in hippocampus ( $P = 0.016$ , actual  $Z = 19.1$ , null mean  $Z = 0.417$ ).

While the above result suggests that univariate HFA is sensitive to the increased mnemonic demands required to discriminate similar trials and old trials, to directly assess whether HFA is related to successful mnemonic discrimination, we next compared HFA during successful versus unsuccessful similar item discrimination. We hypothesized that HFA would be related to successful discrimination and thus would be significantly greater for correctly classified similar items compared with similar items incorrectly classified as old. We found that in both hippocampus and OTC, HFA was significantly greater for correct than incorrect similar items (Fig. 2C). Interestingly, whereas OTC exhibited a significant difference in the early 0.5–1-s window ( $P = 0.024$ , actual  $Z = 9.62$ , null mean  $Z = 0.168$ ), hippocampus exhibited a significant difference in a late 1–1.5-s window (one-tailed  $P = 0.032$ , actual  $Z = 14.5$ , null mean  $Z = 0.256$ ).



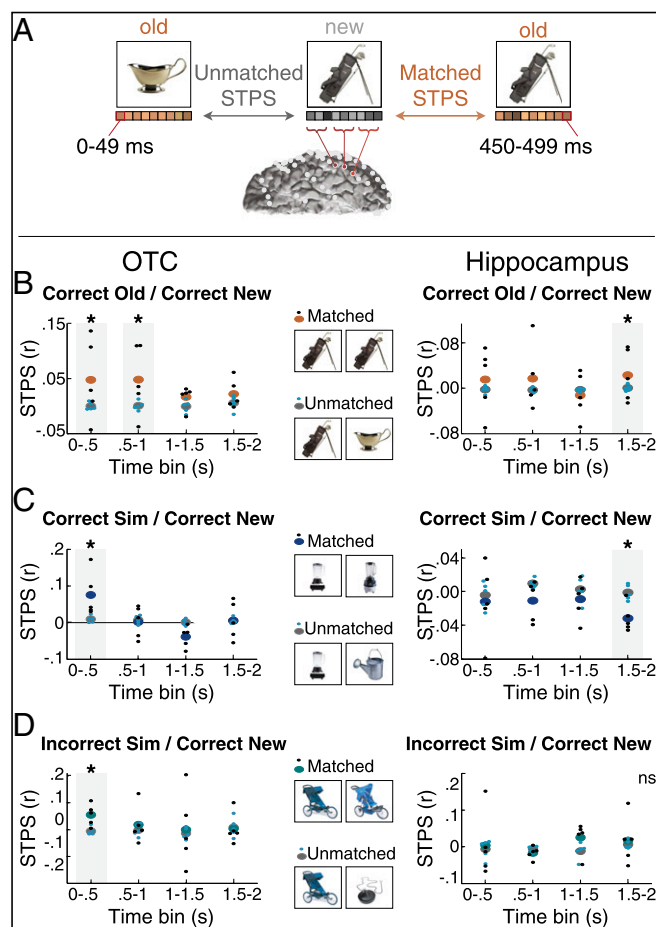
**Fig. 2.** In the fine-grain task, HFA in temporal lobe regions discriminated between stimulus and response type. Significance was assessed in the 2 s following poststimulus onset divided into four 500-ms time bins and is plotted as dashed lines above HFA. (A) HFA for correct old items and correct new items that were subsequently correctly classified as old. HFA in posterior OTC did not exhibit significant differences between old and new items, whereas hippocampal HFA was significantly reduced for old items during 1.5–2-s poststimulus onset. (B) HFA in both OTC and hippocampus was significantly greater for correct similar items than correct old items. (C) HFA in both OTC and hippocampus was significantly greater for correct similar items than similar items incorrectly classified as old. \* $P < 0.05$ . ns, not significant.  $n = 5$ .

In summary, analysis of mean HFA power in hippocampus and OTC suggest that both regions may contribute to successful mnemonic discrimination. First, both regions exhibited HFA that distinguished between correct similar and correct old items. More directly, HFA in both regions differentiated between trials where discrimination of similar items was successful versus not, with greater HFA for successful trials than those for which similar items were classified as old.

**Item-Level Reinstatement and Separation.** While the univariate data are suggestive of memory processes associated with pattern completion and separation, to more directly measure these processes, we next examined multivariate patterns in our data. Specifically, we calculated the similarity in HFA patterns evoked during each item's first presentation (as a new item) and its subsequent presentation either as an old or similar item. An HFA spatiotemporal pattern for each trial was defined as a vector of HFA values, where each vector element corresponded to HFA during a 50-ms time bin at a particular electrode. Thus, for each trial, a vector of HFA values was concatenated across all electrodes within a region of interest (ROI) and all 50-ms time bins within a 500-ms time bin (see *Materials and Methods*). We then calculated the similarity between HFA patterns for an item's first and second presentations (Fig. 3A). Critically, to eliminate more general contributions of univariate activity to pattern similarity, we compared the similarity scores of the actual matched trial pairs (i.e., between each item's first and second presentations) to a region-specific null distribution with shuffled trial labels (see *Materials and Methods*). In addition, for each time bin with a significant difference in univariate HFA between conditions, we performed follow-up analyses verifying that the difference in univariate HFA did not create spurious relationships in spatiotemporal pattern similarity (STPS; *SI Appendix, Table S1*).

We first hypothesized that correct old trials should be associated with significant reinstatement of the original encoding pattern, as measured by a greater correlation between a trial's retrieval pattern and its encoding pattern, consistent with past findings from fMRI in both hippocampus (9, 19–22) and cortex (9, 20, 24–31). We thus computed HFA SPS between matched correct old/correct new pairs, focusing our analyses on the same time windows as when univariate HFA was significantly different for old items in these regions (Fig. 3B). We found evidence for reinstatement in both OTC and hippocampus—namely, that SPS in both regions was significantly greater for matched pairs compared with the null distribution (hippocampus, 1.5–2 s: one-tailed  $P = 0.03$ , actual mean = 0.0228, null mean = 0.0004; OTC, 0.5–1 s:  $P = 0.015$ , actual mean = 0.0480, null mean = 0.0007). In OTC, this significant pattern similarity occurred in an even earlier window as well: 0–0.5 s ( $P = 0.02$ , actual mean = 0.0476, null mean = 0.00004).

Memory reinstatement is likely to be most robust when the presented item is the same as one that was initially presented. Furthermore, strong memory reinstatement may work against mnemonic discrimination when the current item is similar but not identical to the original presentation. For such items, the evoked pattern should be more distinct, or pattern separated, from its original presentation to support successful mnemonic discrimination. We thus asked whether hippocampal SPS between new items and their similar presentations would be significantly reduced, or more distinct, expecting such an effect to occur during one of the later time windows where we saw that HFA discriminated between old and similar items. Critically, during the 1.5–2-s time window, the hippocampus exhibited significantly reduced SPS compared with the unmatched (null) distribution, providing evidence for separation in this region (Fig. 3C; one-tailed  $P = 0.045$ , actual mean =  $-0.0319$ , null mean =  $-0.0013$ ). Indeed, during this time, SPS was significantly greater for old–new item pairs than similar–new item pairs (Fig. 4A;  $P = 0.015$ , actual mean = 0.0548, null mean = 0.0017). These results provide evi-

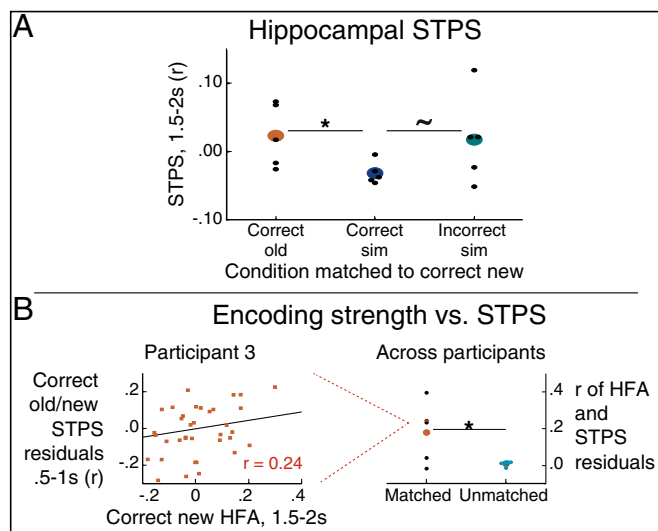


**Fig. 3.** STPS in the fine-grain task. (A) STPS was calculated across frequency and time and concatenated across all electrodes in each region of interest. We considered SPS between item's matched first and second presentations and assessed significance against a baseline of unmatched first and second presentations (see *Materials and Methods* for details). (B) STPS of old items in the fine-grain task. Both OTC and hippocampus exhibited positive STPS, indicative of reinstatement. (C) STPS of similar items in the fine-grain task. OTC exhibited positive STPS indicative of reinstatement, whereas hippocampus exhibited significantly negative STPS, indicative of differentiation. (D) STPS of similar items classified as old in the fine-grain task. OTC exhibited reinstatement of these items, whereas hippocampus did not exhibit significant differences. \* $P < 0.05$ . ns, not significant.  $n = 5$ .

dence for pattern separation in hippocampal HFA activity patterns during mnemonic discrimination of highly similar items (1, 3, 69).

In OTC, we also queried whether there would be significant STPS effects for similar items in the same time bins that we saw reinstatement in OTC. By contrast to hippocampus, no significant separation was seen in OTC. Instead, SPS between new and similar items was significantly greater than expected by chance, consistent with reinstatement in this region, in the 0–0.5-s time bin ( $P = 0.020$ , actual mean = 0.0755, null mean = 0.0094).

Thus far, we have examined pattern similarity for correct items only. A critical step, however, is to query more directly how reinstatement and separation are related to behavioral success in mnemonic discrimination. We next examined SPS for similar items incorrectly classified as old, limiting this analysis to trials whose first presentation was correctly classified as new. Given that in OTC we saw significant reinstatement of correct similar items 0–0.5 s after stimulus presentation, we anticipated that, if there were significant reinstatement of incorrect similar items, it would be during this time bin as well, and indeed this is what we



**Fig. 4.** Hippocampal STPS in the fine-grain task across conditions and correlations with univariate HFA. (A) In the hippocampus 1.5–2-s poststimulus onset, there was significantly more separation of similar items in comparison with correct old items and a trend toward more separation in comparison with incorrect similar items (classified as old). (B, Left) Representative example of calculating each participant's correlation between hippocampal encoding HFA and hippocampal STPS. (B, Right) Across participants, hippocampal HFA during encoding of items subsequently correctly classified as old was significantly correlated with the extent of their hippocampal STPS.  $\sim P < 0.1$ ;  $*P < 0.05$ .  $n = 5$ .

found: In OTC during the 0–0.5-s time bin, STPS was significantly greater for incorrect similar trials (incorrect sim/correct new) in comparison with a null distribution (Fig. 3D;  $P = 0.020$ , actual mean = 0.0548, null mean =  $-0.0059$ ). Thus, OTC exhibited reinstatement irrespective of whether the similar items were correctly classified or not, and pattern similarity between these conditions was not significant ( $P > 0.4$ ).

By contrast, hippocampal STPS did not exhibit any evidence for significant reinstatement or separation between matched and unmatched pairs for incorrect similar trials. Given that the hippocampus exhibited significant pattern reinstatement for old items and pattern separation for similar items in the same 1.5–2-s time bin, we asked whether, in this time bin, we would also see less separation for incorrect similar items in comparison with correct similar items. Thus, we compared STPS between similar items incorrectly labeled as old (incorrect sim) and correct similar items (correct sim) and found a trend toward more separation for correct similar items (Fig. 4A; one-tailed  $P = 0.065$ , actual mean = 0.049, null mean = 0.0076). Thus, we only see significant pattern separation in hippocampus for trials correctly identified as similar, suggesting that hippocampal reinstatement, as measured by STPS, is sensitive to the participant's mnemonic decision.

These pattern analyses reveal that, like univariate HFA power, both OTC and hippocampus distinguish between correct similar and old items. However, critically, OTC showed evidence only for reinstatement, whereas hippocampal patterns showed evidence for both memory reinstatement during exact repeats of old items and for pattern separation of highly similar items.

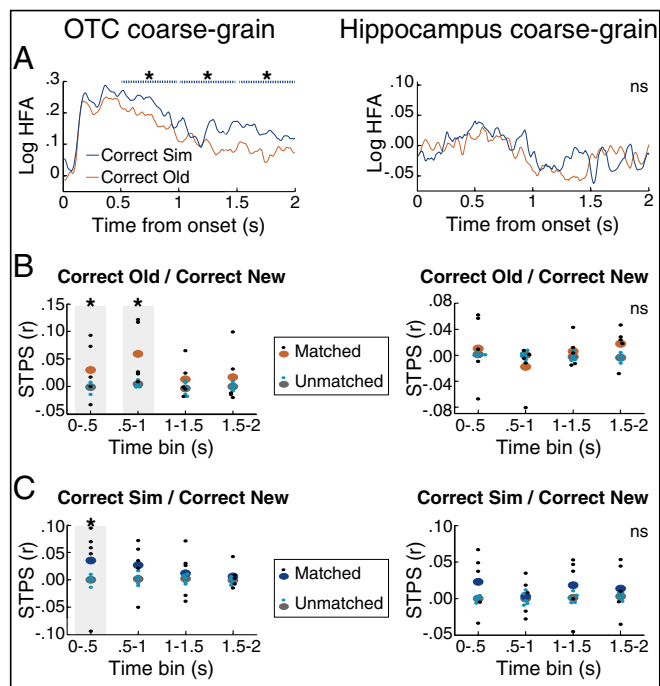
**Univariate Hippocampal HFA Correlates with Hippocampal Pattern Reinstatement.** Theoretically, the extent of memory reinstatement during retrieval for a given item should be related to how well that item was initially encoded—that is, its memory strength. Here we tested this assumption by asking whether hippocampal reinstatement was related to encoding strength. To this end, we quantified the encoding strength of each item as hippocampal HFA during encoding of new items (29, 34, 35) during 1.5–2 s, as this was

the time window with significant differences between old versus new items. Then, for each participant, we calculated the correlation between hippocampal encoding strength and hippocampal HFA pattern similarity on a trial-by-trial basis for matched correct old/correct new pairs for each of the four 0.5-s time bins from 0 to 2 s. Indeed, we found evidence that hippocampal HFA during encoding (in the 1.5–2-s bin) was correlated with pattern reinstatement in the 0.5–1-s time window ( $P = 0.020$ , actual mean = 0.1986, null mean =  $-0.0113$ ). To ensure that this relationship does not reflect contributions of univariate HFA to reinstatement, we calculated a linear regression with STPS from 0.5 to 1 s as the dependent measure, and regressors as univariate HFA during 0.5–1 s of old and new item presentations, as well as an interaction term between these two HFA measures. Using the residuals from this model, which better control for univariate contributions of HFA to STPS, we still expected this STPS measure to correlate with pattern reinstatement. As expected, there was a significant correlation between the STPS residuals during 0.5–1 s and univariate HFA during encoding from 1.5 to 2 s (Fig. 4B;  $P = 0.020$ , actual mean = 0.1776, null mean = 0.0074). Together, these results reduce concern that the correlation between these measures only reflects the increase in encoding HFA and support the interpretation that there is a meaningful relationship between hippocampal HFA during encoding with pattern reinstatement.

**Modulation of Hippocampal HFA by Task Demands.** So far, we have shown that univariate HFA in the hippocampus and OTC exhibited significant differences between correct old, similar, and new items in the fine-grain task, where task demands required discrimination between all three stimulus types. We also asked whether these effects are sensitive to task demands or are more automatic in nature. To this end, we examined HFA in these same regions and participants while they performed a task nearly identical to the one described thus far, except they did not have to distinguish between similar and old items; although participants viewed the same three stimulus types, they responded old to both similar and old items (Fig. 1B).

Interestingly, we found that in this coarse-grain task, hippocampal univariate HFA did not exhibit any significant differences between stimulus types (Fig. 5), in contrast to the fine-grain task. Although most of our critical comparisons were between stimulus types during retrieval (i.e., old or similar items), we verified that these retrieval differences were not an artifact of differences during encoding. Specifically, we compared hippocampal HFA of new items that were correctly recognized between the fine-grain task and the coarse-grain task and found that HFA was not significantly different between tasks in any time bin (all  $P$ s  $> 0.5$ ; SI Appendix, Fig. S2B).

In light of these striking differences from the fine-grain task, we performed several follow-up analyses to ensure that this reflected a difference between tasks. First, we examined how differences across tasks and conditions interact with time. In a three-way repeated measures ANOVA (rmANOVA) of hippocampal HFA with factors as task (fine-grain, coarse-grain), condition (correct old, correct similar), and time bin (0–0.5 s, 0.5–1 s, 1–1.5 s, 1.5–2 s), there was a main effect of time bin,  $F(3, 12) = 11.8$ ,  $P < 0.001$ , and an interaction of time bin and task,  $F(3, 12) = 3.47$ ,  $P = 0.051$ . We were particularly interested in the presence of a significant difference between old and similar items in the fine-grain task yet the absence of this effect in the coarse-grain task. Thus, we performed a follow-up analysis to examine this interaction. The significant interaction of time bin and condition from the rmANOVA motivated our focus on the time bins when the difference between old and similar items was significant in the fine-grain task. Specifically, for each participant, we calculated the mean Z value of the significance between old and similar items in hippocampal HFA during 1–1.5 s and 1.5–2 s in the fine-grain task. We contrasted these Z values to Z values from



**Fig. 5.** HFA and STPS of temporal lobe regions in the coarse-grain task. (A) HFA in posterior OTC distinguished between correct old and correct similar items, despite not being relevant to task demands. Significance was assessed in the 2 s following poststimulus onset divided into four 500-ms time bins and is plotted as dashed lines above HFA. By contrast, hippocampal HFA did not distinguish between stimulus types. (B) STPS of correct old items. As in the fine-grain task, OTC exhibited significant reinstatement for these items. Unlike the fine-grain task, hippocampal STPS did not exhibit significant differences. (C) STPS of correct similar items. As in the fine-grain task, OTC exhibited significant reinstatement for these items. By contrast, hippocampal STPS did not exhibit significant differences. \* $P < 0.05$ . ns, not significant.  $n = 5$ .

the same time bins and conditions in the coarse-grain task, hypothesizing that the Z values would be greater in the fine-grain task than the coarse-grain task. Consistent with our hypothesis, we found a significant interaction whereby the Z values exhibited a significantly greater difference in the fine-grain task than the coarse-grain task [one-tailed  $P = 0.041$ ,  $t(4) = 2.30$ ].

In addition, we examined the interaction of hippocampal HFA between old and new items across tasks. Using the same approach as the preceding analysis, we performed a three-way rmANOVA with factors as task, stimulus type (correct old, correct new), and time bin. This rmANOVA revealed a main effect of time bin,  $F(3, 12) = 7.96$ ,  $P = 0.003$ , and trended toward an interaction between time bin and task,  $F(3, 12) = 2.97$ ,  $P = 0.075$ . As a follow-up analysis, we next compared the Z values of the significance for hippocampal HFA between old and new items in the fine-grain task. We focused on the time bin when this effect was significant, 1.5–2 s, motivated by the significant and trending effects in the rmANOVA. We compared the Z values from the fine-grain task to the Z values from the same stimulus types and time bin in the coarse-grain task. Because the difference was significant in the fine-grain task but not the coarse-grain task, we hypothesized that the Z values would be greater in the fine-grain task, and the effect trended in this direction [one-tailed  $P = 0.094$ ,  $t(4) = 1.59$ ].

Consistent with the hippocampal HFA results, the hippocampus did not exhibit significant STPS differences for matched vs. unmatched pairs in the coarse-grain task. To query whether these effects were significantly different from the fine-grain task, we again performed an interaction analysis. Given that hippocampal HFA only exhibited significant differences between old

and similar items during 1–1.5 s and 1.5–2 s, we examined the across-task differences in STPS during these time bins. First, we calculated the mean STPS values across 1–1.5 s and 1.5–2 s across the four conditions of fine-grain old, fine-grain similar, coarse-grain old, and coarse-grain similar items. We also used the shuffled STPS baselines from these four conditions to generate the null distribution of what we would expect by chance for the difference across tasks and conditions. We then calculated where the actual mean STPS value of these four conditions fell on the null distribution. We expected that the difference would be significantly greater in the fine-grain than the coarse-grain task, and this prediction trended toward significance (one-tailed  $P = 0.070$ , actual mean = 0.0297, null mean = 0.0031). Thus, based on both the univariate and multivariate measures, these results suggest that the hippocampal contribution to memory is sensitive to task demands, such that its activity discriminates between similar and old items more strongly when participants are attending to and are required to respond to these differences.

By contrast, OTC still showed greater HFA for similar compared with old items during the 0.5–1-s time window ( $P = 0.012$ , actual mean  $Z = 11.4$ , null mean  $Z = -0.020$ ). We queried contributions of timing to the OTC HFA effects by performing a three-way rmANOVA with task, stimulus type (correct old, correct similar), and time bin as factors. This rmANOVA revealed a main effect of stimulus type,  $F(1, 4) = 11.9$ ,  $P = 0.026$ , a main effect of time bin,  $F(3, 12) = 4.46$ ,  $P = 0.025$ , and trended toward an interaction of time bin and task,  $F(3, 12) = 2.96$ ,  $P = 0.075$ . Based on these effects, we focused on time bins exhibiting a significant effect in at least one of the two tasks. First, we asked whether the difference between old and similar items was significantly greater in the fine-grain task than the coarse-grain task during 0.5–1 s, when this effect was significant in both tasks, but there was no significant difference between tasks [paired  $t$  test,  $t(4) = 0.90$ ,  $P = 0.42$ ]. Unlike the fine-grain task, this significant effect continued throughout the 2-s interval (1–1.5 s:  $P = 0.008$ , actual mean  $Z = 10.6$ , null mean  $Z = -0.021$ ; 1.5–2 s:  $P = 0.02$ , actual mean  $Z = 10.4$ , null mean  $Z = -0.121$ ). Nonetheless, this later difference between old and similar items was not significantly greater than in the fine-grain task [paired  $t$  test,  $t(4) = 1.24$ ,  $P = 0.28$ ]. Paralleling the encoding analysis in hippocampus, we examined whether HFA in OTC would differ for correct new items between the coarse-grain and the fine-grain task, but there were no significant differences between tasks ( $P_s > 0.5$ ; *SI Appendix, Fig. S24*).

In addition, we asked whether STPS in OTC would continue to exhibit significant reinstatement in the same time bins as in the fine-grain task. We found this to be the case: Reinstatement for old items was significantly greater for matched versus unmatched items during 0–0.5 s (one-tailed  $P = 0.04$ , actual mean = 0.0299, null mean = -0.0013) and 0.5–1 s ( $P < 0.001$ , actual mean = 0.0592, null mean = 0.0041) as well as for similar items during 0–0.5 s ( $P = 0.035$ , actual mean = 0.0352, null mean = 0.0001). We next asked whether the extent of reinstatement in OTC differed between tasks. Unlike the hippocampal results, the time bins in OTC with significant STPS effects were not contained within the time bins with significant HFA. We thus queried contributions of timing to the OTC STPS effects by performing two-way rmANOVAs with time bin and task as factors. Such rmANOVAs revealed no significant effects or interactions, whether the dependent measure was STPS of old items (all  $P_s > 0.2$ ) or STPS of similar items (all  $P_s > 0.05$ ). Thus, we examined the potential interactions across tasks in OTC STPS in all four time bins. None of these pairwise comparisons yielded significant differences (old items: all  $P_s > 0.5$ ; similar items: all  $P_s > 0.3$ ), thus suggesting that reinstatement in OTC did not vary by task demands.

The significant increase in STPS in OTC, occurring relatively early and irrespective of stimulus type and task type, is perhaps more suggestive of a stimulus-evoked perceptual response than of a mnemonic operation. If STPS only reflected a perceptual

similarity between conditions, we would not expect STPS to differ based on an item's classification during retrieval (i.e., its second presentation). Thus, to further query the contribution of memory processes in OTC effects, we examined STPS based on whether participants classified the second presentation of an item as being old or similar (i.e., an awareness that it was a second presentation) versus classified the item as new, including items from both tasks. We found that reinstatement was significant when second presentations were classified as old or similar, and this reinstatement was significantly greater than for second items classified as new (*SI Appendix*). These results suggest that OTC reinstatement is modulated by retrieval of the item's prior presentation. Taken together, our results show that HFA in hippocampus, but not in OTC, was sensitive to task demands.

**HFA in DLPFC.** Thus far, our results have shown that OTC appears to be sensitive to stimulus type irrespective of task demands, and hippocampal HFA is very sensitive to task demands, only distinguishing between the memory status of items in the fine-grain task. To address whether the hippocampus' sensitivity to task demands might be a more general property of "higher level" mnemonic brain regions, we examined HFA in the DLPFC. Like hippocampus, DLPFC has been implicated in successful memory encoding (56–61) and successful memory retrieval (23, 42, 62, 63).

DLPFC HFA did discriminate between stimulus types, but these dissociations were upheld irrespective of task demands (Fig. 6). Most importantly, HFA was significantly greater for similar items than old items in the fine-grain task during the 1–1.5-s and 1.5–2-s time windows (1–1.5 s:  $P = 0.036$ , actual  $Z = 27.5$ , null mean  $Z = -0.542$ ; 1.5–2 s:  $P = 0.020$ , actual  $Z = 31.0$ , null mean  $Z = -0.225$ ), and significant differences in these time windows remained when considering the coarse-grain task (1–1.5 s:  $P < 0.001$ , actual  $Z = 36.1$ , null mean  $Z = 0.074$ ; 1.5–2 s:  $P = 0.030$ , actual  $Z = 25.5$ , null mean  $Z = -0.056$ ). We next asked whether there was an interaction between tasks for the difference in HFA between old and similar items. Following the approach in hippocampus and OTC, we performed a three-way rmANOVA with task, stimulus type, and time bin as factors. This rmANOVA revealed a main effect of stimulus type,  $F(1, 4) = 7.91$ ,  $P = 0.048$ , and main effect of time bin,  $F(3, 12) = 5.80$ ,  $P = 0.011$ . From these results, we could not justify focusing our analyses only on the time bins exhibiting significant differences between old and similar items. We thus performed a paired  $t$  test for each of the four time bins, comparing the  $Z$  values between old and similar items in the fine-grain task to those in the coarse-grain task. There were no significant differences in any time bin (all  $P$ s  $> 0.5$ ), suggesting that DLPFC was not sensitive to task demands. In addition, there were no significant differences between tasks for HFA of new items that were correctly recognized in any time bin (all  $P$ s  $> 0.5$ ; *SI Appendix*, Fig. S2C). These cross-task similarities in DLPFC univariate HFA contrast with the task differences in hippocampal HFA yet are consistent with the role of DLPFC in successful

memory processing (23, 57, 70, 71). Lastly, we examined HFA pattern similarity effects in DLPFC (*SI Appendix*, Fig. S3), but there were no significant differences (fine-grain task: all  $P$ s  $> 0.08$ ; coarse-grain task: all  $P$ s  $> 0.06$ ).

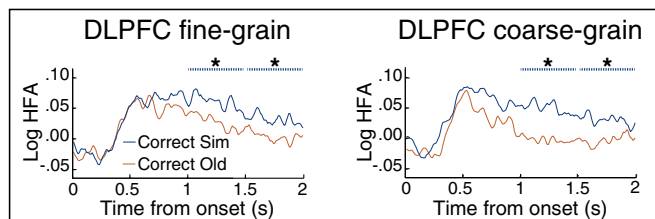
## Discussion

Distinguishing between two highly similar events poses a unique challenge to the episodic memory system. While there is an advantage to preserving access to the unique aspects of each memory, representing the overlapping aspects of two events promotes generalization and learning. Interestingly, hippocampal processes have been hypothesized to contribute to both functions—reinstating prior memory patterns and forming unique memories. Understanding how and when these processes unfold over time in the hippocampus is not known. To this end, we examined electrophysiological activity in the human hippocampus, OTC, and DLPFC for signatures of memory reinstatement and separation.

We found, first, that overall HFA in hippocampus, an established measure of local neuronal firing (49–51), was significantly greater for similar items compared with old items and similar items incorrectly classified as old. This is consistent with prior work showing increased BOLD activation for similar trials compared with old trials (38, 40, 43, 47). Importantly, we used multivariate pattern analyses of the HFA signal to gain leverage on memory-related pattern reinstatement and pattern separation. We found that hippocampal reinstatement and separation were both significant during the 1.5–2-s time window, the same time window where hippocampal univariate HFA was enhanced for similar items compared with old items. Although these effects may be considered relatively "late" with respect to the participant responses (72), relatively little is known about the timing of hippocampal responses. Many past hippocampal electrophysiological studies report activity only from sites responsive to some component of the task (17, 22, 73). By contrast, we include all available hippocampal electrodes in our analyses. In addition, in this small dataset we chose to look at more discrete time bins, and thus, it is always possible with more statistical power that we would detect earlier more punctate differences. However, the late hippocampal responses reported here, often occurring after participants' responses, raise an important question of how these processes are related to the memory decision. It is possible that these later differences in neural activity reflect prospective processes that update the item in memory. For instance, the dissimilarity in hippocampal multivariate HFA activity for similar items might reflect distinctive encoding of these novel items, promoting the separation between the representation of the similar item and its original presentation.

We also found that, on a trial-by-trial basis, hippocampal pattern reinstatement during retrieval of old items was correlated with the hippocampal HFA during encoding of those same items. These results suggest that the strength of reinstatement correlates with the strength of encoding (29, 34). Further, early memory reinstatement may be triggered by the cue strength, and perhaps at other timepoints, reinstatement is modulated by operations engaged during retrieval itself.

Perhaps surprisingly, all of the hippocampal effects were task-dependent and were not evident in a task that presented but did not require participants to discriminate between old and similar items. In the coarse-grain task, hippocampal HFA did not differentiate conditions, either in univariate or pattern similarity analyses. It is unlikely that the lack of significant effects in the coarse-grain task is driven by global factors, such as the task being easier, because activity patterns in other regions (OTC and DLPFC; see below) were similar across the fine- and coarse-grain task. Furthermore, it is unlikely that the null effects were due to lack of power because the hippocampal results were significant in the more fine-grain version of the task, and there was evidence of interactions of the effects across tasks. Thus, the differences in hippocampal responses across



**Fig. 6.** HFA in DLPFC. HFA in DLPFC was significantly greater for correct similar items than correct old items in both (*Left*) the fine-grain task and (*Right*) the coarse-grain task. Significance was assessed in the 2 s following poststimulus onset divided into four 500-ms time bins and is plotted as dashed lines above HFA. \* $P < 0.05$ .  $n = 5$ .

tasks suggest that active attentional and goal processes need to be considered and incorporated into existing models of hippocampal pattern separation and reinstatement.

The results from the coarse-grain task may seem particularly difficult to reconcile with past studies lacking an explicit memory task, which nonetheless find evidence of hippocampal separation and reinstatement. For instance, several fMRI studies report differences in hippocampal activity as participants passively view stimuli (38, 40, 74). Further, many of the seminal rodent studies establishing hippocampal pattern separation and reinstatement involve simple foraging tasks. However, whereas these studies are arguably neutral with respect to promoting separation versus reinstatement of highly similar stimuli, our coarse-grain task specifically discourages pattern separation because participants are supposed to classify similar stimuli as if they were exact repeats. As a result, participants may not attend to the overlapping details between similar stimuli.

Our hippocampal findings fit well and build upon prior ECoG work providing evidence for hippocampal HFA activity and pattern reinstatement during memory decisions that engendered the recovery of contextual details (22, 75). A recent study (73) examined hippocampal ECoG activity during mnemonic discrimination of old and highly similar pictures of celebrities. Using single-unit recordings, they found decreased neuronal firing in CA3/DG when viewing lures (similar items) in comparison with targets (old items) and that the extent of this decrease correlated with successful memory discrimination. Given that HFA has been shown to correlate with neuronal firing (49–51), it might seem surprising that, in a similar early time window in the fine-grain task, we found significantly increased HFA for similar than old items. However, our univariate HFA results reflect the activity of a multitude of single units and thus do not provide information about how the firing rates of single units are modulated by the task. By contrast, our pattern separation result is consistent with the notion that a single unit sensitive to a learned stimulus may be less responsive to a new stimulus, even if it is highly similar, akin to what is reported (73).

In the current study, we were limited by the electrode placements, which were determined based on clinical criteria. However, there is ample theoretical and empirical work suggesting that subregions CA1 versus CA3/DG may be related to the distinct processes of pattern separation and reinstatement (3, 15, 38, 40, 41, 43, 76–78). Although our postsurgical images of electrode placements are not at a resolution to determine the hippocampal subregions of the electrodes, across participants ~60% of the electrodes were in the posterior hippocampus, which generally has a larger proportion of CA3/DG (79, 80). Future work can aim to fully characterize how hippocampal subregions are modulated by task demands and stimulus types.

It is critical to note that hippocampus was not the only brain region where HFA differentiated between similar and old trials. Like hippocampus, HFA in regions of the visual cortex (OTC) was also greater for similar compared with both old items and similar items classified as old. Interestingly, the HFA difference between correct and incorrect similar items emerged in OTC during 0.5–1 s, yet in hippocampus, this effect emerged 1–1.5 s. The fact that activity in visual cortex differentiates highly similar stimuli is consistent with recent work showing that fMRI activity in inferior temporal regions can distinguish between highly similar stimuli (48, 81). Using pattern analyses, we also found evidence for mnemonic reinstatement during old item presentations. However, critically, unlike the hippocampus, OTC did not show any evidence for pattern separation. Rather, OTC patterns showed evidence for reinstatement, not separation, during similar item presentations. These results build on prior work implicating cortical reinstatement in memory success (32, 33, 82–84).

It is tempting to think that pattern separation and reinstatement may be relatively automatic processes that support

attention to overlapping and novel features of an environment during memory decisions. Indeed, the fact that the same pattern of neural effects was evident in OTC and DLPFC in both versions of the task (but not hippocampus, see above) suggests that the processes supporting memory discrimination in these regions may be relatively automatic. OTC may be sensitive to the perceptual details in repeated items irrespective of the task, such that increased correlations in OTC activity between matched pairs of items may reflect their shared perceptual processing. Although we cannot rule out the possibility that part of the multivariate patterns in OTC reflects this shared perceptual processing, it is unlikely that activity in the OTC region is purely perceptual, as both univariate and multivariate activity in OTC distinguished between items based on participants' memory responses. Indeed, it is intriguing to speculate that the early univariate effects in OTC may be related to a familiarity signal that has also extensively been shown to be rapid and relatively automatic (69, 72, 85, 86). Further work that specifically differentiates the subjective sense of familiarity from recollection, however, would be needed before this claim could be directly tested. In DLPFC, differences in univariate HFA paralleled those of OTC, with significantly greater HFA for similar than old items in both tasks. The univariate findings fit well with the reported role of DLPFC in both successful memory encoding (56–61) and retrieval (23, 42, 62, 63). However, perhaps somewhat surprisingly, the DLPFC results did not differ in the two tasks (63, 87–89). One possibility is that our results reflect the role of DLPFC in postretrieval monitoring, a process by which the retrieved information is assessed based on task demands (62, 87, 90, 91). Both tasks may involve postretrieval monitoring because discriminating between similar and old items may require retrieval of the item's original presentation as well as postretrieval monitoring to help determine, based on the task and stimulus type, whether a response of *sim* or *old* is appropriate.

In summary, our findings underscore the unique role of the hippocampus in mnemonic decisions. In the hippocampus, pattern analyses revealed significant pattern separation of items that were similar but not identical to an earlier item. By contrast, the hippocampus exhibited significant reinstatement of encoding activity during presentation of exact repeats. Reinstatement and separation in hippocampus occurred in the same late time window, suggesting that these processes may emerge over similar time scales. By contrast, an earlier visual region in OTC also exhibited reinstatement of old items but exhibited reinstatement for similar items as well. Taken together, these results provide support for the idea that the occipitotemporal and prefrontal cortical regions may be sensitive to the demands required to discriminate highly similar lures from exact repeats—an especially challenging mnemonic operation—but that only the hippocampus may promote distinctive representations for these highly similar lures.

## Materials and Methods

**Participants.** Five participants (4 female; 19–42 y old) with intractable epilepsy were recruited via the Comprehensive Epilepsy Center of the New York University School of Medicine. Participants had elected to undergo intracranial monitoring for clinical purposes and provided informed consent to participate in this study under the approval of the local Institutional Review Board. Relevant clinical and demographic information for these participants is summarized in *SI Appendix, Table S2*.

**Task Design.** Participants performed two separate blocks: one block of the fine-grain task and one block of the coarse-grain task (Fig. 1 A and B). The order of the blocks was counterbalanced across participants. In each block, participants were presented with a series of images on a computer screen. Each image was either novel (*new*), an exact repetition of a prior new stimulus (*old*), or an image that was highly overlapping, but not identical to, a prior new stimulus ("*similar*"). Each image was presented for 2.5–5 s, with a blank 2.5 s interstimulus interval separating trials. Presentation of the stimulus terminated following a participant's response or 2.5 s, whichever came later. If no response was made after 5 s, item presentation ended.



In the fine-grain task block, participants were presented with 96 new images. Of these, half were presented again as old images, and the other half were presented as similar trials. Participants were instructed to indicate, on each trial, whether the presented image was new, old, or similar. The three response options appeared in black on the bottom of the stimulus screen, in the same order as the response keys. The number of intervening items between a new image and its subsequent old/similar trial ranged from 1 to 8 trials. There were no differences by lag in memory accuracy, condition-level HFA, or any of the critical STPS analyses (*SI Appendix, Fig. S4*).

The coarse-grain task block had the same design as the fine-grain block, except that participants were instructed to designate the similar items as old. One participant only completed the task for 64 new items (and thus 32 similar and old items each) for each of the blocks.

### Analysis.

**Conditions.** When comparing a first presentation item (new) to a second presentation item (e.g., old), we consider the same set of stimuli in both cases: that is, those items that were correctly classified as new for their first presentation and subsequently correctly classified as, for example, old items for the second presentation. In this way, the comparisons between first and second presentation items were matched for the number of observations and the types of stimuli that were tested. Furthermore, all comparisons between old and similar items were only conducted when their first presentations were correctly identified as new, thus providing some control for initial encoding. **HFA.** Given that there were no clear peaks in the power spectrum in higher frequencies, we defined an HFA band at 45–115 Hz, above the beta band but below the second line noise harmonic. We calculated spectral power by applying a Morlet wavelet transform (wavelet number, 6) during stimulus presentation (0–2,000 ms poststimulus onset) at 5-Hz intervals for each electrode and trial within an ROI (*SI Appendix, Table S3*). A 1,000-ms buffer was included on both sides of the data to minimize edge effects. Due to the broad distribution of power values, we took the (natural) log transform of the power values. Power values for each trial were normalized by subtracting the mean power at the same frequency during the corresponding baseline period 1,500–500 ms before stimulus onset (see *SI Appendix* for details of statistics on HFA univariate power). Although statistics on HFA univariate power are reported based on the distribution of values across participants, we also calculated significance on the individual participant level (*SI Appendix, Fig. S5 and Table S4*). For every reported significant difference in HFA, at least one participant showed this significant difference, aside from the comparison of hippocampal HFA for old versus new items. Nonetheless, even for this comparison, several participants trended toward

significance, and thus across participants, this effect was significant. For illustrative purposes only, in Figs. 2, 5, and 6, HFA is plotted as the mean across every 50 ms with a 10-ms sliding window.

**HFA STPS.** At each electrode for each trial, an HFA pattern vector was constructed for each 500-ms time bin and type/response condition. Specifically, HFA was calculated at each electrode as above, in nonoverlapping 50-ms time bins. In this way, 10 (50-ms) time bins  $\times$  the number of participant's electrodes were included in each HFA pattern to yield a single vector of HFA values per trial.

To ensure that condition differences in pattern vectors do not reflect condition differences in univariate HFA, mean HFA across all trials of a stimulus/response type (e.g., correct old in the fine-grain task) was subtracted from each time-frequency element in every vector, within participant (see also *SI Appendix, Table S1*). These vectors could then be compared with each other to determine their correlation or similarity. Matched pattern similarity was calculated as the Pearson's  $r$  correlation between the spatiotemporal pattern vector of an item's first presentation (correct new) and the pattern vector of the item's second presentation (as correct old, correct sim, incorrect sim). We used permutation tests to assess significance of pattern similarity values, as this allowed us to estimate a fair baseline of expected pattern similarity within and across participants. Specifically, for each participant, pair of conditions, and ROI, we permuted the trial labels of the first presentations and calculated the pattern similarity between first and second presentations based on these permuted trial labels at each time bin. The null distribution was defined as the mean of the pattern similarity values across 200 such permutations, and thus, the null distribution was unique to each ROI and pair of conditions. We then took the actual mean STPS value across participants and took the mean of the null distribution across participants. The point at which the actual matched STPS fell in the region-specific, condition-specific null distribution determined the  $P$  value. Unless noted otherwise in the text, reported  $P$  values are Bonferroni-corrected for the number of time windows. Although we report significance based on the mean across participants, we also calculated the significance for each participant (*SI Appendix, Fig. S6 and Table S4*). For all but two of the reported significant differences in STPS, at least one participant exhibited a significant effect. For those conditions where no one participant exhibited a significant effect, several participants trended toward significance.

**ACKNOWLEDGMENTS.** We thank Olga Felsovalyi and Amy Trongnetrpunya for their assistance with data collection as well as the patients who participated in this study. This work was supported by National Institutes of Mental Health Grants MH106266 (to L.J.L.) and MH074692 (to L.D.).

- O'Reilly RC, Norman KA (2002) Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. *Trends Cogn Sci* 6: 505–510.
- Treves A, Rolls ET (1994) Computational analysis of the role of the hippocampus in memory. *Hippocampus* 4:374–391.
- Yassa MA, Stark CEL (2011) Pattern separation in the hippocampus. *Trends Neurosci* 34:515–525.
- Hulbert JC, Norman KA (2015) Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cereb Cortex* 25: 3994–4008.
- Favila SE, Chanales AJH, Kuhl BA (2016) Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nat Commun* 7: 11066.
- Chanales AJH, Oza A, Favila SE, Kuhl BA (2017) Overlap among spatial memories triggers divergence of hippocampal representations. *Curr Biol* 27:1–47.
- Marr D (1971) Simple memory: A theory for archicortex. *Philos Trans R Soc Lond B Biol Sci* 262:23–81.
- Yassa MA, et al. (2011) Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus* 21:968–979.
- Tompary A, Duncan K, Davachi L (2016) High-resolution investigation of memory-specific reinstatement in the hippocampus and perirhinal cortex. *Hippocampus* 26: 995–1007.
- Duncan K, Sadanand A, Davachi L (2012) Memory's penumbra: Episodic memory decisions induce lingering mnemonic biases. *Science* 337:485–487.
- Duncan K, Tompary A, Davachi L (2014) Associative encoding and retrieval are predicted by functional connectivity in distinct hippocampal area CA1 pathways. *J Neurosci* 34:11188–11198.
- Jezek K, Henriksen EJ, Treves A, Moser EI, Moser M-B (2011) Theta-paced flickering between place-cell maps in the hippocampus. *Nature* 478:246–249.
- Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychol Rev* 110:611–646.
- Leutgeb S, et al. (2005) Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science* 309:619–623.
- Lee I, Yoganarasimha D, Rao G, Knierim JJ (2004) Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature* 430:456–459.
- Lee H, Wang C, Deshmukh SS, Knierim JJ (2015) Neural population evidence of functional heterogeneity along the CA3 transverse axis: Pattern completion versus pattern separation. *Neuron* 87:1093–1105.
- McKenzie S, et al. (2014) Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron* 83:202–215.
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity—Connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4.
- Backus AR, Bosch SE, Ekman M, Grabovetsky AV, Doeller CF (2016) Mnemonic convergence in the human hippocampus. *Nat Commun* 7:11991.
- Mack ML, Preston AR (2016) Decisions about the past are guided by reinstatement of specific memories in the hippocampus and perirhinal cortex. *Neuroimage* 127: 144–157.
- van den Honert RN, McCarthy G, Johnson MK (2016) Reactivation during encoding supports the later discrimination of similar episodic memories. *Hippocampus* 26: 1168–1178.
- Staresina BP, et al. (2016) Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *eLife* 5:1–18.
- Sederberg PB, et al. (2007) Gamma oscillations distinguish true from false memories. *Psychol Sci* 18:927–932.
- Manning JR, Polyn SM, Baltuch GH, Litt B, Kahana MJ (2011) Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proc Natl Acad Sci USA* 108:12893–12897.
- Manning JR, Sperling MR, Sharan A, Rosenberg EA, Kahana MJ (2012) Spontaneously reactivated patterns in frontal and temporal lobe predict semantic clustering during memory search. *J Neurosci* 32:8871–8878.
- Yaffe RB, et al. (2014) Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proc Natl Acad Sci USA* 111:18727–18732.
- Yaffe RB, Shaikhouni A, Arai J, Inati SK, Zaghoul KA (2017) Cued memory retrieval exhibits reinstatement of high gamma power on a faster timescale in the left temporal lobe and prefrontal cortex. *J Neurosci* 37:4472–4480.

28. Ritchey M, Wing EA, LaBar KS, Cabeza R (2013) Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cereb Cortex* 23: 2818–2828.
29. Wing EA, Ritchey M, Cabeza R (2015) Reinstatement of individual past events revealed by the similarity of distributed activation patterns during encoding and retrieval. *J Cogn Neurosci* 27:679–691.
30. LaRocque KF, et al. (2013) Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. *J Neurosci* 33:5466–5474.
31. Staresina BP, Henson RN, Kriegeskorte N, Alink A (2012) Episodic reinstatement in the medial temporal lobe. *J Neurosci* 32:18150–18156.
32. Davachi L, Danker JF (2013) Cognitive neuroscience of episodic memory. *The Oxford Handbook of Cognitive Neuroscience*, eds Ochsner KN, Kosslyn S (Oxford Univ Press, New York), pp 375–388.
33. Davachi L, Preston AR (2014) *The Medial Temporal Lobe and Memory*, eds Gazzaniga MS, Mangun GR (MIT Press, Cambridge, MA), 5th Ed.
34. Danker JF, Tompary A, Davachi L (2017) Trial-by-trial hippocampal encoding activation predicts the fidelity of cortical reinstatement during subsequent retrieval. *Cereb Cortex* 27:3515–3524.
35. Bosch SE, Jehes JFM, Fernández G, Doeller CF (2014) Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *J Neurosci* 34: 7493–7500.
36. Gordon AM, Rissman J, Kiani R, Wagner AD (2014) Cortical reinstatement mediates the relationship between content-specific encoding activity and subsequent recollection decisions. *Cereb Cortex* 24:3350–3364.
37. St-Laurent M, Abdi H, Buchsbaum BR (2015) Distributed patterns of reactivation predict vividness of recollection. *J Cogn Neurosci* 27:2000–2018.
38. Bakker A, Kirwan CB, Miller M, Stark CEL (2008) Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science* 319:1640–1642.
39. Kirwan CB, Stark CEL (2004) Medial temporal lobe activation during encoding and retrieval of novel face-name pairs. *Hippocampus* 14:919–930.
40. Lacy JW, Yassa MA, Stark SM, Muftuler LT, Stark CEL (2010) Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learn Mem* 18:15–18.
41. Duncan K, Ketz N, Inati SJ, Davachi L (2012) Evidence for area CA1 as a match/mismatch detector: A high-resolution fMRI study of the human hippocampus. *Hippocampus* 22:389–398.
42. Hannula DE, Ranganath C (2008) Medial temporal lobe activity predicts successful relational memory binding. *J Neurosci* 28:116–124.
43. Azab M, Stark SM, Stark CEL (2014) Contributions of human hippocampal subfields to spatial and temporal pattern separation. *Hippocampus* 24:293–302.
44. Copara MS, et al. (2014) Complementary roles of human hippocampal subregions during retrieval of spatiotemporal context. *J Neurosci* 34:6834–6842.
45. Pidgeon LM, Morcom AM (2016) Cortical pattern separation and item-specific memory encoding. *Neuropsychologia* 85:256–271.
46. Leutgeb S, Leutgeb JK, Treves A, Moser M-B, Moser EI (2004) Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* 305:1295–1298.
47. Hashimoto R, et al. (2012) Changing the criteria for old/new recognition judgments can modulate activity in the anterior hippocampus. *Hippocampus* 22:141–148.
48. Kirwan CB, Stark CEL (2007) Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learn Mem* 14:625–633.
49. Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560–1563.
50. Hirabayashi T, et al. (2014) Distinct neuronal interactions in anterior inferotemporal areas of macaque monkeys during retrieval of object association memory. *J Neurosci* 34:9377–9388.
51. Manning JR, Jacobs J, Fried I, Kahana MJ (2009) Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *J Neurosci* 29:13613–13620.
52. Crone NE, Korzeniewska A, Franaszczuk PJ (2011) Cortical  $\gamma$  responses: Searching high and low. *Int J Psychophysiol* 79:9–15.
53. Lachaux J-P, Axmacher N, Mormann F, Halgren E, Crone NE (2012) High-frequency neural activity and human cognition: Past, present and possible future of intracranial EEG research. *Prog Neurobiol* 98:279–301.
54. Janowsky JS, Shimamura AP, Squire LR (1989) Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia* 27:1043–1056.
55. McAndrews MP, Milner B (1991) The frontal cortex and memory for temporal order. *Neuropsychologia* 29:849–859.
56. Ranganath C, Johnson MK, D'Esposito M (2000) Left anterior prefrontal activation increases with demands to recall specific perceptual information. *J Neurosci* 20:RC108.
57. Jenkins LJ, Ranganath C (2010) Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *J Neurosci* 30:15558–15565.
58. Blumenfeld RS, Parks CM, Yonelinas AP, Ranganath C (2011) Putting the pieces together: The role of dorsolateral prefrontal cortex in relational memory encoding. *J Cogn Neurosci* 23:257–265.
59. Ranganath C, Cohen MX, Brozinsky CJ (2005) Working memory maintenance contributes to long-term memory formation: Neural and behavioral evidence. *J Cogn Neurosci* 17:994–1010.
60. Sederberg PB, et al. (2007) Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cereb Cortex* 17:1190–1196.
61. Ranganath C, Johnson MK, D'Esposito M (2003) Prefrontal activity associated with working memory and episodic long-term memory. *Neuropsychologia* 41:378–389.
62. Dobbins IG, Foley H, Schacter DL, Wagner AD (2002) Executive control during episodic retrieval: Multiple prefrontal processes subservise source memory. *Neuron* 35:989–996.
63. Ranganath C, Cohen MX, Dam C, D'Esposito M (2004) Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *J Neurosci* 24:3917–3925.
64. Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: Neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23.
65. Rodriguez Merzagora A, et al. (2014) Repeated stimuli elicit diminished high-gamma electrocorticographic responses. *Neuroimage* 85:844–852.
66. Zhang H, et al. (2015) Gamma power reductions accompany stimulus-specific representations of dynamic events. *Curr Biol* 25:635–640.
67. Gruber T, Müller MM (2005) Oscillatory brain activity dissociates between associative stimulus content in a repetition priming task in the human EEG. *Cereb Cortex* 15: 109–116.
68. Rugg MD, et al. (2012) Item memory, context memory and the hippocampus: fMRI evidence. *Neuropsychologia* 50:3070–3079.
69. Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 30:123–152.
70. Nolde SF, Johnson MK, Raye CL (1998) The role of prefrontal cortex during tests of episodic memory. *Trends Cogn Sci* 2:399–406.
71. Rugg MD, Yonelinas AP (2003) Human recognition memory: A cognitive neuroscience perspective. *Trends Cogn Sci* 7:313–319.
72. Brown MW, Aggleton JP (2001) Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nat Rev Neurosci* 2:51–61.
73. Suthana NA, et al. (2015) Specific responses of human hippocampal neurons are associated with better memory. *Proc Natl Acad Sci USA* 112:10503–10508.
74. Huffman DJ, Stark CEL (2014) Multivariate pattern analysis of the human medial temporal lobe revealed representational categorical cortex and representationally agnostic hippocampus. *Hippocampus* 24:1394–1403.
75. Merkow MB, Burke JF, Kahana MJ (2015) The human hippocampus contributes to both the recollection and familiarity components of recognition memory. *Proc Natl Acad Sci USA* 112:14378–14383.
76. Schapiro AC, Kustner LV, Turk-Browne NB (2012) Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Curr Biol* 22: 1622–1627.
77. Wills TJ, Lever C, Cacucci F, Burgess N, O'Keefe J (2005) Attractor dynamics in the hippocampal representation of the local environment. *Science* 308:873–876.
78. Leutgeb JK, Leutgeb S, Moser M-B, Moser EI (2007) Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315:961–966.
79. Malykhin NV, Lebel RM, Coupland NJ, Wilman AH, Carter R (2010) In vivo quantification of hippocampal subfields using 4.7 T fast spin echo imaging. *Neuroimage* 49: 1224–1230.
80. Poppenk J, Evensmoen HR, Moscovitch M, Nadel L (2013) Long-axis specialization of the human hippocampus. *Trends Cogn Sci* 17:230–240.
81. Reagh ZM, Yassa MA (2014) Object and spatial mnemonic interference differentially engage lateral and medial entorhinal cortex in humans. *Proc Natl Acad Sci USA* 11: E4264–E4273.
82. Wheeler ME, Petersen SE, Buckner RL (2000) Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proc Natl Acad Sci USA* 97:11125–11129.
83. Nyberg L, Habib R, McIntosh AR, Tulving E (2000) Reactivation of encoding-related brain activity during memory retrieval. *Proc Natl Acad Sci USA* 97:11120–11124.
84. Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes retrieval during memory search. *Science* 310:1963–1966.
85. Yonelinas AP (2002) The nature of recollection and familiarity: A review of 30 years of research. *J Mem Lang* 46:441–517.
86. Yonelinas AP, Jacoby LL (2012) The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Mem Cognit* 40:663–680.
87. MacDonald AW, 3rd, Cohen JD, Stenger VA, Carter CS (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288:1835–1838.
88. Braver TS, et al. (1997) A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* 5:49–62.
89. Rudolf S, Hare TA (2014) Interactions between dorsolateral and ventromedial prefrontal cortex underlie context-dependent stimulus valuation in goal-directed choice. *J Neurosci* 34:15988–15996.
90. D'Esposito M, et al. (1995) The neural basis of the central executive system of working memory. *Nature* 378:279–281.
91. Henson RN, Shallice T, Dolan RJ (1999) Right prefrontal cortex and episodic memory retrieval: A functional MRI test of the monitoring hypothesis. *Brain* 122:1367–1381.
92. Yang AI, et al. (2012) Localization of dense intracranial electrode arrays using magnetic resonance imaging. *Neuroimage* 63:157–165.