

# Episodic Reinstatement in the Medial Temporal Lobe

Bernhard P. Staresina, Richard N. A. Henson, Nikolaus Kriegeskorte, and Arjen Alink

MRC Cognition and Brain Sciences Unit, Cambridge CB2 7EF, United Kingdom

The essence of episodic memory is our ability to reexperience past events in great detail, even in the absence of external stimulus cues. Does the phenomenological reinstatement of past experiences go along with reinstating unique neural representations in the brain? And if so, how is this accomplished by the medial temporal lobe (MTL), a brain region intimately linked to episodic memory? Computational models suggest that such reinstatement (also termed “pattern completion”) in cortical regions is mediated by the hippocampus, a key region of the MTL. Although recent functional magnetic resonance imaging studies demonstrated reinstatement of coarse item properties like stimulus category or task context across different brain regions, it has not yet been shown whether reinstatement can be observed at the level of individual, discrete events—arguably the defining feature of episodic memory—nor whether MTL structures like the hippocampus support this “true episodic” reinstatement. Here we show that neural activity patterns for unique word-scene combinations encountered during encoding are reinstated in human parahippocampal cortex (PhC) during retrieval. Critically, this reinstatement occurs when word-scene combinations are successfully recollected (even though the original scene is not visually presented) and does not encompass other stimulus domains (such as word-color associations). Finally, the degree of PhC reinstatement across retrieval events correlated with hippocampal activity, consistent with a role of the hippocampus in coordinating pattern completion in cortical regions.

## Introduction

Episodic memory enables us to mentally travel back in time and relive past experiences in great detail (James, 1890; Tulving, 1985); an ability lacking in individuals who have amnesia following damage to their medial temporal lobe (MTL) (Scoville and Milner, 1957; Squire et al., 2004). A natural assumption is that this mental time travel is accomplished by reinstating neural representations of past events (Tulving and Thomson, 1973), and computational models suggest that such reinstatement (also termed “pattern completion”) is mediated by the hippocampus, a key region of the MTL (Marr, 1971; Norman and O’Reilly, 2003).

Previous functional magnetic resonance imaging (fMRI) studies have shown that retrieving a specific stimulus category (e.g., words vs pictures or sounds vs pictures) differentially reactivates category-specific brain regions involved in the initial stimulus processing (Wheeler and Buckner, 2004; Woodruff et al., 2005), and more recent fMRI studies using multivoxel pattern analysis (MVPA; Norman et al., 2006) have shown whole-brain reinstatement of coarse properties like stimulus category (Polyn et al., 2005) or task context (Johnson et al., 2009). While these findings suggest that overall category- or context-specific neural representations are reinstated during successful retrieval, it has

not yet been shown whether reinstatement can be observed at the level of individual, discrete events—arguably the defining feature of episodic memory—nor whether MTL structures like the hippocampus support this “true episodic” reinstatement.

To assess the mechanisms of episodic reinstatement in the human brain, we applied fMRI while participants encoded and later retrieved trial-unique words paired with one of four scenes (home, office, city, nature). During retrieval trials, participants indicated whether they thought a cue word was previously seen and if so, whether they also remembered the associated scene (which was not visually present at retrieval, Fig. 1*a*). Based on participants’ responses, events of interest were categorized as Word Only (WO) retrievals (word recognition without memory for the associated scene) or Word & Scene (WS) retrievals (word recognition plus recollection of the associated scene). Episodic reinstatement was quantified by measuring the similarity (Pearson correlation coefficient) of activity patterns between each encoding trial and the corresponding retrieval trial (Fig. 1*b*), henceforth referred to as encoding/retrieval similarity (ERS). We performed this analysis for individually, anatomically defined MTL regions of interest: the hippocampus (HIPPO) and—to probe nonhippocampal, neocortical reinstatement—the parahippocampal cortex (PhC), a region known to be involved in scene processing (Epstein and Kanwisher, 1998) (Fig. 2*a*).

## Materials and Methods

**Participants.** Twenty (14 female) right-handed native English speakers with normal or corrected-to-normal vision participated in the experiment (mean age: 28 years, range: 20–35). Informed consent was obtained in a manner approved by a local Psychological Research Committee and participants were paid for their participation.

**Experimental design.** The stimuli consisted of 360 English concrete nouns and 8 different associated source details: the scenes home, office, city and nature (Fig. 1), and the colors blue, green, red and yellow. Each

Received Aug. 30, 2012; revised Oct. 11, 2012; accepted Oct. 26, 2012.

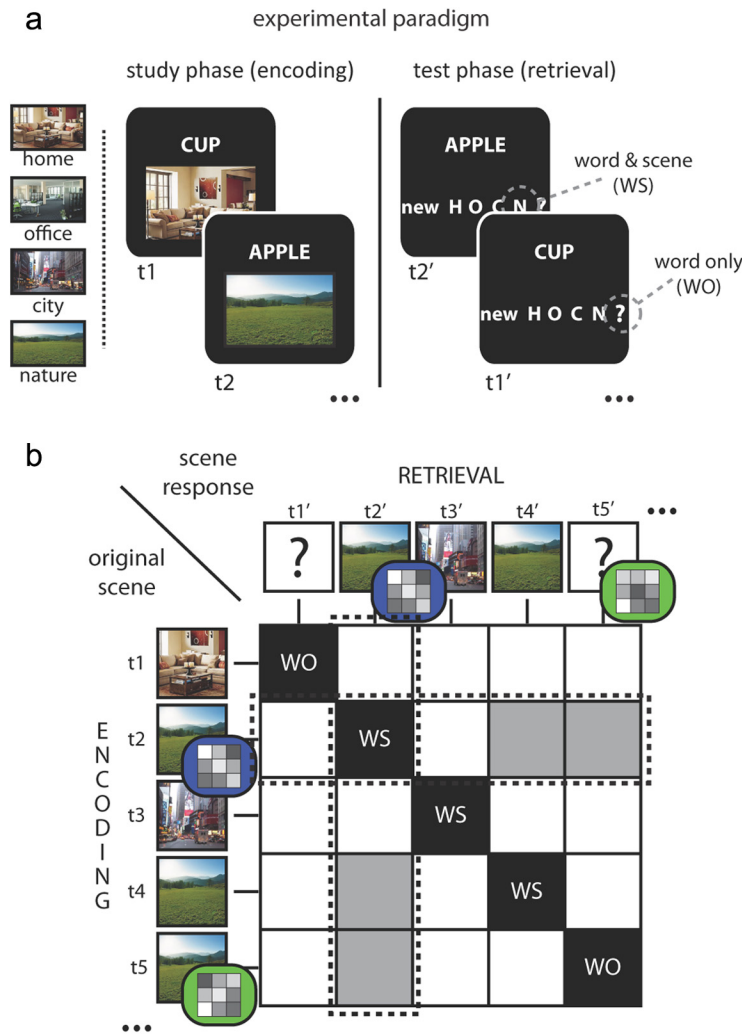
Author contributions: B.P.S. and R.N.A.H. designed research; B.P.S. performed research; N.K. and A.A. contributed unpublished reagents/analytic tools; B.P.S. and A.A. analyzed data; B.P.S., R.N.A.H., N.K., and A.A. wrote the paper.

This work was supported by a Sir Henry Wellcome Postdoctoral Fellowship to B.P.S., the UK Medical Research Council Programmes MC-A060-5PR10 (R.N.A.H.) and MC-A060-5PR20 (N.K.), a European Research Council Starting Grant to N.K., and a Rubicon Grant from the Netherlands Organisation for Scientific Research (NWO) 825.10.023 to A.A. We thank Katherine Duncan and Dean Mobbs for helpful discussion.

Correspondence should be addressed to Dr. Bernhard Staresina, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK. E-mail: bernhard.staresina@mrc-cbu.cam.ac.uk.

DOI:10.1523/JNEUROSCI.4156-12.2012

Copyright © 2012 the authors 0270-6474/12/3218150-07\$15.00/0



**Figure 1.** *a*, Schematic of the experimental design, showing all four scenes and two exemplary trials at encoding (t1, t2) and retrieval (t1', t2') with two different memory outcomes. *b*, Illustration of the ERS analysis. ERS is the correlation between the activation pattern over voxels within a given region for a given trial at encoding and the activation pattern for the corresponding trial at retrieval (same-trial ERS, black diagonal cells). In our initial analysis, the difference in same-trial ERS for WS vs WO retrievals was assessed. In a subsequent analysis, same-trial ERS values were compared with the ERS between words that shared the same scene exemplar (same-scene ERS; gray, off-diagonal cells for example trial 2).

Study (encoding) or Test (retrieval) block used only one category (scenes or colors), and scene and color blocks were alternated, with the assignment of the first block counterbalanced across participants. The assignment of words to the list of Study items or lures for the Test phase and to scene vs color blocks was randomized across participants. Because scene stimuli have been shown to strongly activate HIPP and PhC (Epstein and Kanwisher, 1998; Burgess et al., 2002), our current analysis focused on scene trials but we used color trials (for which all experimental parameters were identical) as a control to assess the category-specificity of the effects we obtained.

The paradigm is schematized in Figure 1. During each study block, participants saw 60 words together with one of four possible scenes. Words were presented in white uppercase letters and centered on a black background. The associated scene was presented in a 250 × 350 pixels frame positioned 150 pixels underneath the word. A given word-scene combination was presented on the screen for 4 s and participants indicated whether the given combination was plausible or implausible (Staresina and Davachi, 2006). Participants were encouraged to give their response as fast as possible. All Study blocks were presented in a row (with a short break in between), followed by a 5 min break during which the anatomical scan was acquired, and then all Test blocks were presented in a row (again with a short break in between). A Test block contained all 60 previously seen (studied) words as

well as 30 experimentally novel (unstudied) words (lures). The lures were pseudorandomly intermixed, holding the average delay between study and test constant across words. Upon being presented with a word, participants could give one of 6 possible answers: (1) new (word not seen during the Study phase), (2) old, seen with home, (3) old, seen with office, (4) old, seen with city, (5) old, seen with nature, or (vi) old, but cannot remember the scene (“don’t know” response). Thus participants indicated two memory decisions with one response: whether they thought the word was old or new and whether they also remembered the associated scene. Both Study and Test trials were presented at a fixed duration of 4 s. Trials were pseudorandomly intermixed with an active “arrows task” to maximize the efficiency of the rapid event-related design (see below).

**MRI scanning and data analysis.** Scanning was performed on a 3-T Siemens Tim Trio MRI system using a 32-channel whole-head coil. Functional data were acquired using a gradient-echo, echo-planar pulse sequence (TR = 2000 ms, TE = 30 ms, 32 horizontal slices oriented parallel to the hippocampal axis, descending slice acquisition, 3 × 3 × 3 mm voxel size, 0.75 mm interslice gap, 165 and 245 volume acquisitions per Study and Test session, respectively). The first 7 volumes of each session were discarded to allow for magnetic field stabilization. High-resolution (1 × 1 × 1 mm) T1-weighted (magnetization-prepared rapid acquisition gradient echo) images were collected for anatomical visualization. Foam padding was used to minimize head motion. Visual stimuli were projected onto a screen that was viewed through a mirror, and responses were collected with magnet-compatible button boxes placed under the participant’s hands. Stimuli were created and presented using the Psychophysics Toolbox (Brainard, 1997) implemented in MATLAB (MathWorks).

Both the Study and the Test portion of the experiment were scanned. Trials were intermixed with an active, sensorimotor baseline task (arrows task; Stark and Squire, 2001), comprising a fourth of the total scanning time). Arrows that randomly pointed to the left or to the right for 1 s were repeatedly presented for the length of a baseline trial (2–10 s), and participants had to press the left hand index finger key if the arrow pointed to the left and the right hand index finger key if it pointed to the right. The sequence of “old” trials associated with each of the four possible scenes, “new” trials and the variable number of baseline trials, was pseudorandom and optimized for rapid event-related fMRI (using the “optseq” algorithm; Dale, 1999).

The experimental paradigm yielded two main conditions of interest: (1) studied words correctly identified as old, without remembering the correct scene from the study episode (WO recognition) and (2) studied words correctly identified as old, additionally remembering the correct scene (WS). For WO, to increase statistical power, we collapsed trials for which participants gave “don’t know” responses and trials for which the wrong scene was indicated. Correct rejections (unstudied words correctly classified as new), false alarms (unstudied words incorrectly classified as old) and misses (studied words incorrectly classified as new) were modeled but not included in the further analysis.

Data were analyzed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) and custom MATLAB code. During preprocessing, images were corrected for differences in slice acquisition timing, followed by motion

correction across all sessions. Regions of interest were created using hand-drawn, participant-specific masks, based on the individual structural image (which was coregistered to the functional images; see Fig. 1 for an example). Anatomical demarcation of the parahippocampal gyrus was done according to the method of Insausti et al. (1998) and Pruessner et al. (2002), and PhC was defined as the posterior third of the parahippocampal gyrus (Staresina et al., 2011).

ERS was quantified via Representational Similarity Analysis (Kriegeskorte et al., 2008; for a recent application on episodic encoding see Xue et al., 2010). Each individual trial of a scanning session was modeled as an individual impulse regressor convolved with a canonical hemodynamic response function in a mass-univariate general linear model (GLM; along with nuisance regressors for head motion, low-frequency scanner drift and run means) and the resulting  $\beta$  values were transformed into  $t$  values (Kriegeskorte et al., 2008). No smoothing or normalization was performed on the echoplanar imaging data that entered the GLM. ERS was quantified as the Pearson correlation coefficient of a trial's activation pattern during encoding with the corresponding activation pattern of that trial during retrieval.

As our ERS values were based on correlation coefficients for which the assumption of normal distribution is generally violated, values were Fisher-transformed before entering them into  $t$  tests and ANOVAs (note that the same statistical patterns were obtained when using nonparametric tests).

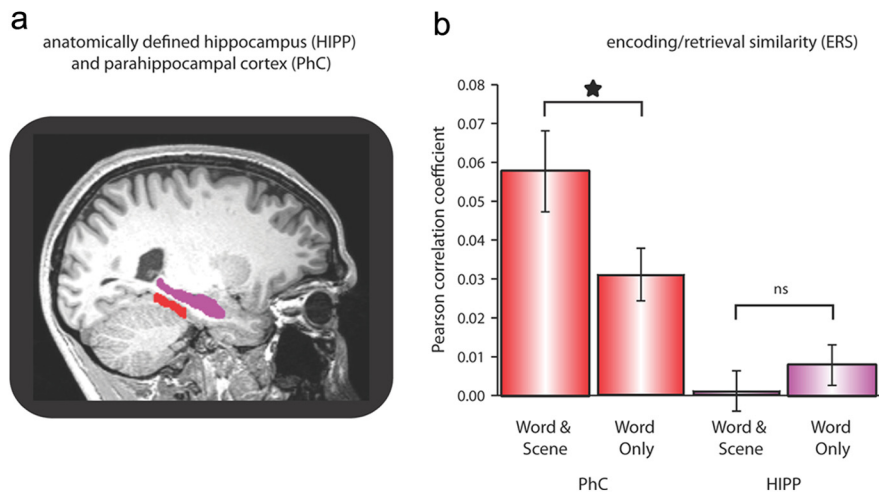
## Results

Behavioral results showed that of all studied words, 86% ( $\pm 2\%$ ) were correctly recognized ("hits"), of which 57% were WS and 43% were WO retrievals on average. Of unstudied new words, 83% ( $\pm 3\%$ ) were correctly identified as new ("correct rejection"). For the imaging analysis, the average number of trials was 59 (range 17–90) for WS retrievals and 42 (range 19–71) for WO retrievals.

To assess neural reinstatement for WS and WO retrievals in our regions of interest, we first applied a repeated-measures ANOVA with the factors region (HIPP, PhC), hemisphere (left, right) and memory (WS, WO) on the ERS values. We observed a significant region  $\times$  memory interaction ( $F_{(1,19)} = 29.23$ ,  $p < 0.0001$ ), which was due to significantly greater ERS during WS vs WO retrievals in PhC ( $t_{(19)} = 3.88$ ,  $p = 0.001$ ) but not in HIPP ( $t_{(19)} = 0.96$ ,  $p = 0.347$ ) (Fig. 2b). Given that there was no interaction with the factor hemisphere, subsequent analyses were collapsed across left and right PhC and HIPP (values were obtained separately for left and right hemispheres and then averaged before entering statistical assessment, to avoid any bias introduced by different numbers of voxels in left and right ROIs).

### ERS and overall activation in PhC

As a first control, we wanted to ensure that the ERS in PhC during WS retrievals did not merely reflect greater blood oxygen level-dependent (BOLD) levels during those trials and hence greater power to detect correlations between encoding and the corresponding retrieval trials. We therefore implemented, within each participant, a GLM in which the dependent variable was the ERS values across trials and the regressors were: (1) each trial's mnemonic status (WS or WO), (2) that trial's BOLD activation during encoding and (3) that trial's BOLD activation during retrieval. The resulting  $\beta$  values for the first regressor (mnemonic



**Figure 2.** *a*, Structural scan of the first participant, showing the individual PhC (red) and HIPP (violet) masks. *b*, ERS for WS and WO retrievals in PhC and HIPP. Star,  $p < 0.05$ ; ns, not significant ( $p > 0.1$ ).

status) for each participant were then tested against zero, with the critical question being whether the "memory" regressor still accounted for a significant amount of variability in ERS when controlling for overall BOLD activation at encoding and at retrieval. Indeed, this  $\beta$  was significantly greater than zero across participants ( $t_{(19)} = 2.92$ ,  $p = 0.009$ ), suggesting that a trial's mnemonic status impacted the ERS above and beyond overall activation levels in PhC.

### Overall scene reinstatement in PhC

While the previous analyses established that the ERS in PhC was greater for WS than for WO retrievals and suggest that ERS was not driven by overall levels of activation, the similarity values for WS retrievals could still reflect solely reinstatement of the encoded scene, rather than retrieval of a unique encoding trial (i.e., the specific word-scene combination). We therefore asked whether the ERS between retrieval and encoding patterns in PhC for a given trial ("same-trial ERS;" highlighted black cells for example trial 2 in Fig. 1b) was greater than the average ERS between (1) the retrieval pattern for that trial and the encoding patterns for all other words that were paired with the same scene and (2) the encoding pattern for that trial and the retrieval patterns for all other words that were paired with the same scene ("same-scene ERS;" gray cells for example trial 2 in Fig. 1b). Note that (1) and (2) were averaged, but the same results were obtained when using one or the other values. In a repeated-measures ANOVA with the factors ERS type (same-trial, same-scene) and memory (WS, WO), a significant interaction ( $F_{(1,19)} = 8.99$ ,  $p = 0.007$ ) reflected greater same-trial ERS than same-scene ERS for WS retrievals ( $t_{(19)} = 2.96$ ;  $p = 0.008$ ) but not for WO retrievals ( $t_{(19)} = 0.58$ ,  $p = 0.569$ ). This shows that remembering a particular word-scene combination increases the ERS beyond mere reinstatement of the target scene (otherwise no difference would be expected in the correlation between a given retrieval trial and its actual encoding counterpart [same-trial ERS] and the correlation between a given retrieval trial and any other encoding trial that contained the same scene [same-scene ERS]). In fact, the same-trial ERS during WS retrievals even exceeded the similarity with other trials for which the same scene was also later recollected ( $t_{(19)} = 2.84$ ;  $p = 0.010$ ). Note that by comparing each WS trial the similarity to all other WS trials for which the same scene was also recollected, both the scene representation and the mean

level of overall memory strength are matched. In sum, these results suggest that the reinstatement of encoding representations in PhC goes beyond mere reinstatement of the target scene (which would be identical for all retrievals sharing a particular scene, especially those for which that scene was recollected), and is bolstered by the trial-unique combination of a scene and a particular word.

### Word cue reinstatement in PhC

To what extent does the visual display of the cue word per se (which was also seen during encoding), rather than the particular word-scene combination, drive ERS? To assess this, we first derived a “baseline” ERS. This baseline ERS was defined as the average similarity between encoding trials and previously unseen (experimentally novel) words during retrieval. Note that there should be no mnemonic reinstatement between encoding trials and novel retrieval trials, so that the resulting correlation likely reflects similarity due to basic stimulus processing, physiological noise etc. This average baseline ERS was 0.023 in PhC. Critically, WO retrievals did not differ statistically from the baseline ERS ( $t_{(19)} = 0.142, p = 0.171$ ). However, ERS for WS retrievals showed a strong increase from this baseline ERS ( $t_{(19)} = 4.71, p < 0.001$ ), constituting ~250% enhancement. This suggests that recognizing the cue word itself, without remembering the associated scene, did not drive ERS in our paradigm.

### Category-specificity in PhC (scenes vs colors)

Finally, we wanted to test whether episodic reinstatement in PhC was specific to scenes (arguably that region’s preferred stimulus category; Epstein and Kanwisher, 1998), or whether this is a more global phenomenon that also transfers to other stimulus classes. To this end, we applied the same encoding/retrieval similarity analysis to color trials, where all experimental parameters were identical except that instead of pairing a word with one of four scenes, we used one of four colors (see Materials and Methods). Behaviorally, there were no differences in performance between word-scene and word-color trials. In terms of encoding strategy, we compared the plausibility ratings and found no difference in the proportion of “plausible” to “implausible” responses (scene trials: 46% plausible, color trials: 40% plausible,  $t_{(19)} = 1.60, p > 0.1$ ). Regarding memory performance, the proportion of correctly remembered word plus scene/color retrievals out of hits was also comparable: 57% for scene trials and 55% for color trials,  $t_{(19)} = 0.52, p > 0.1$ ). However, despite comparable behavioral performance for word-scene and word-color trials, there was no evidence in PhC for differential ERS during successful versus unsuccessful color recollection ( $t_{(19)} = 0.98, p = 0.339$ ). [Average numbers for color trials entering the analysis: 59 (range 17–99) for word-color retrievals and 46 (range 13–81) for word-only retrievals.] Indeed, there was a significant memory (successful vs unsuccessful scene/color memory)  $\times$  stimulus category (scenes, colors) interaction ( $F_{(1,19)} = 5.58, p = 0.045$ ), reflecting the specificity of PhC ERS to scene recollection. To ensure that we had sufficient power to detect ERS for word-color trials, we queried—based on previous data showing word-color binding in this region (Staresina and Davachi, 2006, 2010)—the perirhinal cortex (PrC) for word-color ERS (again using individually, anatomically defined ROIs). An initial ANOVA with the factors memory (word plus color, word only) and hemisphere (left, right) revealed a significant memory  $\times$  hemisphere interaction ( $F_{(1,19)} = 4.64, p = 0.045$ ), due to a significant difference in ERS for word plus color vs word only retrievals in the right ( $t_{(19)} = 2.66, p = 0.016$ ) but not the left PrC ( $t_{(19)} = 0.09, p = 0.927$ ). Importantly,

there was no difference in right PrC regarding ERS for scene recollection (word plus scene vs word only retrievals,  $t_{(19)} = 0.34, p = 0.737$ ). In fact, we observed a double dissociation between ERS for word only vs word plus scene/color retrievals in PhC and PrC [interaction of the factors memory (word plus scene/color, word only)  $\times$  stimulus category (scenes, colors)  $\times$  region (bilateral PhC, right PrC),  $F_{(1,19)} = 5.20, p = 0.034$ ]. While ERS for color trials in PrC did not survive the same-exemplar comparison described above (which could be due to poorer BOLD contrast in PrC and/or coarser reinstatement in PrC and/or less discriminability of colors compared with scenes), this double dissociation shows that we had sufficient power to detect ERS for color trials and thus emphasizes that episodic reinstatement in PhC was specific to word-scene information.

The same scene-specificity was also observed during encoding, where different scene exemplars were visually presented (Fig. 1a). This was assessed by comparing the within-exemplar similarity (e.g., the similarity of DOG-home with all other words paired with the “home” scene) to the between-exemplar similarity [e.g., the similarity of DOG-home with all other words paired with any of the other scenes (“office,” “city,” “nature”)]. The within-exemplar similarity was significantly greater than the between-exemplar similarity for scene exemplars ( $t_{(19)} = 3.89, p = 0.001$ ) but not for color exemplars ( $t_{(19)} = 1.21, p = 0.24$ ), again with a significant interaction of stimulus category (scene, color)  $\times$  similarity (within, between),  $F_{(1,19)} = 15.55, p = 0.001$ . These results are consistent with the idea that there is scene-specific information coding in PhC at the exemplar level (Bonnici et al., 2012; Liang et al., 2012).

### Hippocampus

At first sight, it might seem surprising that there was no evidence for reinstatement in the hippocampus, given this region’s consistent association with recollection in univariate fMRI studies (Davachi, 2006; Eichenbaum et al., 2007; Mayes et al., 2007). To ensure that we had sufficient statistical power in our hippocampal data, we first tested whether we would observe recollection effects in a more conventional univariate analysis. To this end, we took each participant’s individually, anatomically defined left and right hippocampus and extracted the estimated magnitude of fMRI response for each trial (using the same time series model as for the ERS analysis). In a repeated-measures ANOVA with the factors hemisphere (left, right) and memory (WS, WO), there was a significant main effect of memory ( $F_{(1,19)} = 7.40, p = 0.014$ ), due to WS retrievals showing stronger activation than WO retrievals. This result replicates a large body of work showing recollection-related activation in the hippocampus (for review, see Eichenbaum et al., 2007), and confirm that we had sufficient overall power in our hippocampal data. Although direct evidence for pattern completion in the human hippocampus is currently lacking, there are compelling theoretical/computational accounts for pattern completion in the hippocampus (Marr, 1971; Norman and O’Reilly, 2003). However, these accounts assign different roles in pattern completion vs pattern separation to different hippocampal subfields, with the dentate gyrus (DG) thought to support pattern separation, and the CA3 region thought to support pattern completion (Marr, 1971). As a result, our data might reflect a combination of hippocampal pattern separation and completion processes occurring concurrently in different subfields (that cannot be clearly separated even with high-resolution fMRI; Bakker et al., 2008; Duncan et al., 2012), which might in turn preclude detection of pattern completion signals (Leutgeb et al., 2007).

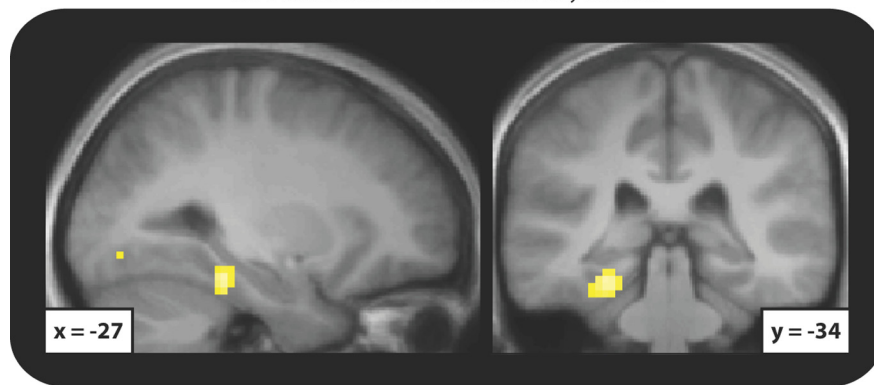
We then asked whether hippocampal activation might—if not reflecting reinstatement per se—correlate with the level of ERS in PhC, consistent with a role in driving cortical pattern completion in downstream regions (Marr, 1971; Norman and O'Reilly, 2003). To test this, we correlated (Spearman's rank correlation) the strength of each trial's ERS in PhC with the corresponding trial-specific response magnitude in the hippocampus. Correlation coefficients were significantly greater than zero across participants ( $t_{(19)} = 3.59$ ,  $p = 0.002$ ; average of 0.1 across participants), which held when calculating the correlation separately for WS and WO retrievals (both  $t_{(19)} > 2.59$ ,  $p < 0.019$ ). Interestingly, no such correlation was seen during encoding ( $t_{(19)} = 1.38$ ,  $p > 0.1$ ). In fact, the correlation between hippocampal activation and PhC ERS was significantly greater during retrieval than during encoding ( $t_{(19)} = 2.43$ ,  $p = 0.025$ ), in line with hippocampally mediated pattern completion occurring specifically during retrieval. As a final control, we correlated the trial-by-trial response magnitude of PhC with each trial's ERS in the hippocampus (i.e., reversing the correlation analysis described above). Indeed, that correlation did not differ from zero ( $t_{(19)} = 0.12$ ,  $p = 0.91$ ; average of 0.003 across participants) and was significantly smaller than the correlation between the trial-by-trial response magnitude of the hippocampus and ERS in PhC ( $t_{(19)} = 2.75$ ,  $p = 0.013$ ).

### Whole-brain analysis

Despite our a priori regions of interest in the MTL, it is important to know whether other brain regions show similar effects. To this end, we queried the whole brain for differential ERS for WS vs WO retrievals in an exploratory manner. Note that this analysis was not used to define the ROIs of the main hypothesis-driven analyses, so it does not render these analyses circular (Kriegeskorte et al., 2009). We applied a searchlight procedure (Kriegeskorte, 2011) in which a sphere (3-voxel-radius) was centered on every voxel (normalized to a standard brain template), and subjected the resulting set of voxels to the same ERS analysis as described above. This results in an ERS difference (WS minus WO) associated with each voxel (searchlight center), and the resulting ERS effect maps were smoothed with a 6 mm full-width at half-maximum Gaussian kernel and compared against zero across participants. The resulting  $t$  map was initially thresholded at  $p < 0.001$  (uncorrected, 10 contiguous voxels minimum). As shown in Figure 3, the effect we observed was—within the MTL—highly specific to the PhC [surviving familywise error (FWE) correction in an MTL mask that captured HIPP, PhC, and perirhinal cortex across all participants (1417 voxels) at  $P_{\text{corrected}} = 0.032$ ]. Outside the MTL, the only clusters surpassing the initial threshold were in early visual cortex ( $\sim$ BA 18) and posterior middle frontal gyrus ( $\sim$ BA 4), of which none survived whole-brain FWE correction.

It is worth noting that while the left PhC cluster that emerged from this whole-brain searchlight analysis also passed the test for trial-specific reinstatement during WS retrievals described above (comparing the same-trial ERS with that trial's similarity to all other trials for which the same scene was also recollected) ( $t_{(19)} = 2.21$ ,  $p = 0.040$ ), the clusters in early visual cortex and posterior middle frontal gyrus did not (both  $t_{(19)} < 1.39$ ,  $p > 0.184$ ), indicating that ERS in

whole-brain searchlight encoding/retrieval similarity (ERS)  
word & scene retrievals minus word only retrievals



**Figure 3.** Left parahippocampal cortex cluster surviving small volume correction, displayed on the normalized anatomical scan averaged across participants.

those other regions did not reach the level of trial-uniqueness as did ERS in PhC.

### Discussion

The current results provide evidence for reinstatement of neural representations that are unique to individual encoding trials. While previous findings suggested that overall category- or context-specific neural representations are reinstated during successful retrieval (Wheeler and Buckner, 2004; Polyn et al., 2005; Woodruff et al., 2005; Johnson et al., 2009), they did not determine whether reinstatement can be observed at the level of individual, discrete events. In this study, we further link this trial-unique reinstatement to the medial temporal lobe, in that reinstatement occurred in the PhC and was specific to successful recollection of word-scene associations. Finally, the level of cortical reinstatement correlated with retrieval activation in the hippocampus, pointing to a role of the hippocampus in orchestrating cortical pattern completion.

The PhC has consistently been reported in episodic memory tasks, especially (but not exclusively) when the relevant episodic details are scene-related (for review, see Diana et al., 2007), and has been shown to code for other stimulus classes (e.g., objects) if those stimuli have been associated with navigationally relevant information (Janzen and van Turennout, 2004) or generally strong contextual details (Aminoff et al., 2007). Here we provide evidence that PhC not only becomes more responsive during successful scene recollection, but that activation patterns elicited by word-scene combinations during encoding are reliably reinstated during recollection in this region even in the absence of an external scene stimulus. Moreover, our results indicate that the conjunction of a scene with a particular word sets this combination's reinstated pattern in PhC apart from the patterns associated with other trials that shared the same scene but were paired with a different word. This was evidenced by a significant increase of the same-trial ERS (the similarity between a retrieval trial and its actual encoding counterpart) relative to the global same-scene ERS. One possible scenario is that imagining a given scene together with a particular word during encoding renders the representation of that scene in PhC sufficiently unique to differ from trials in which the same scene is paired with another word. Consistent with this idea, data from an fMRI adaptation study show that PhC is just as sensitive to subtle alterations within the same scene (such as a viewpoint-change) as it is to complete scene changes (Epstein et al., 2003). By that token, it is plausible to

assume that a different word shifts the internal focus to a different aspect of the same scene (e.g., DOG-home might emphasize the internal representation of the couch, where the dog is imagined to sleep, whereas CUP-home might emphasize the internal representation of the coffee-table, where the cup is imagined to stand).

So could it be that the visually re-presented cue word (DOG in the example above) alone—regardless of whether that word was successfully bound to a particular scene—accounts for the reinstatement effects we observed? In our current paradigm, this seems unlikely on several grounds. First, ERS for WO retrievals, in which the re-presented cue word was recognized (but the associated scene was not recollected) did not differ from our baseline ERS (i.e., the similarity between encoding words and new retrieval words), where no psychological reinstatement should occur. Only when the cue word was recognized AND the associated scene was also recollected (WS retrievals) did we observe a significant increase in ERS. Could it then be that recollection of any associated information bolsters ERS, perhaps through increasing the signal-to-noise ratio via enhanced BOLD response? This also seems unlikely. First, we controlled for increased BOLD levels (both during encoding and retrieval) in our calculation of ERS and still found a significant increase of ERS for WS retrievals relative to WO retrievals. Perhaps more importantly, no ERS increase was observed for word-plus-color retrievals in PhC. If any of the effects we observed were driven by the cue word rather than by remembering the cue word along with the associated scene, the same pattern should be seen for word-color trials, which was not the case. In sum, the results that (1) a given word-scene combination elicited greater ERS than the same scene paired with other words and that (2) memory for the cue word alone did not drive ERS suggests that the reinstatement effects we observed in PhC were indeed sensitive to particular and unique word-scene conjunctions.

Finally, it should be mentioned that although we use the term “recollection” when describing WS retrievals, this is not to say that reinstatement does not occur during phenomenologically different memory states. Given that we used only four scene exemplars to probe associative memory, there might still be a certain contribution of familiarity to our observed ERS effects for WS retrievals, or recollection of other, non-scene episodic information in WO retrievals (Yonelinas and Jacoby, 1996). Indeed, a recent fMRI study using MVPA showed graded levels of contextual reinstatement with graded levels of memory strength (Johnson et al., 2009).

Regarding reinstatement beyond the MTL, although the early visual region emerging from our whole-brain searchlight analysis (Fig. 3) did not show clear trial-unique reinstatement (i.e., ERS for a given WS trial was not greater than ERS with other WS trials sharing the same scene), an interesting possibility is that episodic reinstatement mediated by the MTL draws on or elicits representations in early sensory regions. Future studies could assess the connectivity between MTL regions and different (e.g., visual or auditory) sensory regions during reinstatement of information initially presented in different modalities.

With regard to the hippocampus, the key finding of our study is that trial-by-trial fluctuations in hippocampal activation positively correlate with the level of reinstatement in PhC. Although interpretive caution is warranted with respect to causality, this result fits well with the pivotal role the hippocampus is thought to play in driving cortical pattern completion (Norman and O'Reilly, 2003). To the extent that pattern completion is predominantly exerted during retrieval, the finding that hippocampal activation correlated with PhC reinstatement only during re-

trieval (and significantly stronger so than during encoding) further strengthens this notion. Note that even for WO retrievals, where the correct scene was not successfully remembered and in which the ERS is on average lower than for WS retrievals, there are fluctuations in the amount of ERS. It is thus not surprising that hippocampal activation correlated with PhC ERS for both WS and WO retrievals.

Together, the current findings suggest that our ability to mentally reexperience past events in great detail is accompanied by reinstatement of neural activation patterns in the medial temporal lobe and that this reinstatement is mediated by the hippocampus.

## References

- Aminoff E, Gronau N, Bar M (2007) The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb Cortex* 17:1493–1503. Medline
- Bakker A, Kirwan CB, Miller M, Stark CE (2008) Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science* 319:1640–1642. CrossRef Medline
- Bonnici HM, Kumaran D, Chadwick MJ, Weiskopf N, Hassabis D, Maguire EA (2012) Decoding representations of scenes in the medial temporal lobes. *Hippocampus* 22:1143–1153. Medline
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. CrossRef Medline
- Burgess N, Maguire EA, O'Keefe J (2002) The human hippocampus and spatial and episodic memory. *Neuron* 35:625–641. CrossRef Medline
- Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109–114. CrossRef Medline
- Davachi L (2006) Item, context and relational episodic encoding in humans. *Curr Opin Neurobiol* 16:693–700. CrossRef Medline
- Diana RA, Yonelinas AP, Ranganath C (2007) Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends Cogn Sci* 11:379–386. CrossRef Medline
- 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. CrossRef Medline
- Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. *Annu Rev Neurosci* 30:123–152. CrossRef Medline
- Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601. CrossRef Medline
- Epstein R, Graham KS, Downing PE (2003) Viewpoint-specific scene representations in human parahippocampal cortex. *Neuron* 37:865–876. CrossRef Medline
- Insausti R, Juottonen K, Soinen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitkänen A (1998) MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *AJNR Am J Neuroradiol* 19:659–671. Medline
- James W (1890) *The principles of psychology*. New York: Dover Publications.
- Janzen G, van Turenout M (2004) Selective neural representation of objects relevant for navigation. *Nat Neurosci* 7:673–677. CrossRef Medline
- Johnson JD, McDuff SG, Rugg MD, Norman KA (2009) Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. *Neuron* 63:697–708. CrossRef Medline
- Kriegeskorte N (2011) Pattern-information analysis: from stimulus decoding to computational-model testing. *Neuroimage* 56:411–421. CrossRef Medline
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis—connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4. CrossRef Medline
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12:535–540. CrossRef Medline
- Leutgeb JK, Leutgeb S, Moser MB, Moser EI (2007) Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315:961–966. CrossRef Medline
- Liang JC, Wagner AD, Preston AR (2012) Content representation in the human medial temporal lobe. *Cereb Cortex*. Advance online publication. Retrieved Aug. 15, 2012. doi:10.1093/cercor/bhr379. CrossRef

- Marr D (1971) Simple memory: a theory for archicortex. *Philos Trans R Soc Lond B Biol Sci* 262:23–81. [CrossRef Medline](#)
- Mayes A, Montaldi D, Migo E (2007) Associative memory and the medial temporal lobes. *Trends Cogn Sci* 11:126–135. [CrossRef Medline](#)
- Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol Rev* 110:611–646. [CrossRef Medline](#)
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430. [CrossRef Medline](#)
- Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes retrieval during memory search. *Science* 310:1963–1966. [CrossRef Medline](#)
- Pruessner JC, Köhler S, Crane J, Pruessner M, Lord C, Byrne A, Kabani N, Collins DL, Evans AC (2002) Volumetry of temporopolar, perirhinal, entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus. *Cereb Cortex* 12:1342–1353. [CrossRef Medline](#)
- Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 20:11–21. [CrossRef Medline](#)
- Squire LR, Stark CE, Clark RE (2004) The medial temporal lobe. *Annu Rev Neurosci* 27:279–306. [CrossRef Medline](#)
- Staresina BP, Davachi L (2006) Differential encoding mechanisms for subsequent associative recognition and free recall. *J Neurosci* 26:9162–9172. [CrossRef Medline](#)
- Staresina BP, Davachi L (2010) Object unitization and associative memory formation are supported by distinct brain regions. *J Neurosci* 30:9890–9897. [CrossRef Medline](#)
- Staresina BP, Duncan KD, Davachi L (2011) Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. *J Neurosci* 31:8739–8747. [CrossRef Medline](#)
- Stark CE, Squire LR (2001) When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proc Natl Acad Sci U S A* 98:12760–12766. [CrossRef Medline](#)
- Tulving E (1985) Memory and consciousness. *Can Psychol* 26:1. [CrossRef](#)
- Tulving E, Thomson DM (1973) Encoding specificity and retrieval processes in episodic memory. *Psychol Rev* 80:352. [CrossRef](#)
- Wheeler ME, Buckner RL (2004) Functional-anatomic correlates of remembering and knowing. *Neuroimage* 21:1337–1349. [CrossRef Medline](#)
- Woodruff CC, Johnson JD, Uncapher MR, Rugg MD (2005) Content-specificity of the neural correlates of recollection. *Neuropsychologia* 43:1022–1032. [CrossRef Medline](#)
- Xue G, Dong Q, Chen C, Lu Z, Mumford JA, Poldrack RA (2010) Greater neural pattern similarity across repetitions is associated with better memory. *Science* 330:97–101. [CrossRef Medline](#)
- Yonelinas AP, Jacoby LL (1996) Noncriterial recollection: familiarity as automatic, irrelevant recollection. *Conscious Cogn* 5:131–141. [CrossRef Medline](#)