Published in final edited form as:

Hippocampus. 2011 June ; 21(6): 575–583. doi:10.1002/hipo.20781.

Recollection-related Hippocampal Activity during Continuous Recognition: a High-Resolution fMRI Study

Maki Suzuki^{1,2,*}, Jeffrey D. Johnson¹, and Michael D. Rugg¹

¹ Center for the Neurobiology of Learning and Memory, and Department of Neurobiology and Behavior, University of California, Irvine, California, USA

² Japan Society for the Promotion of Science (JSPS), Tokyo, Japan

Abstract

We used high-resolution functional magnetic resonance imaging (fMRI) to investigate whether successful recollection during continuous recognition is associated with relative enhancement of hippocampal activity, consistent with prior findings from experiments employing separate study and test phases. While being scanned, subjects discriminated between new and repeated pictures. Each picture, which was repeated once after an interval of between 10 and 30 items, was surrounded by a frame that was colored grey, blue, or orange. When an item repeated, its frame color determined the correct response. Repeated items surrounded by a grey frame always required an 'old' judgment. A repeated item surrounded by a blue or an orange frame required a different response depending whether it was re-presented in the same (Target) or a different (Nontarget) color from the first presentation. Consistent with the results from previous continuous recognition experiments, robust new > old effects were found in bilateral hippocampus. Additionally, an across-subjects correlational analysis identified a cluster of voxels in right hippocampus where recollection-related activity (operationalized by the contrast between correctly vs. incorrectly judged Nontargets) was positively correlated with recollection performance. Thus, successful recollection during continuous recognition is associated with a relative enhancement of hippocampal activity.

Keywords

familiarity; novelty; hippocampus; medial temporal lobe; memory

INTRODUCTION

fMRI studies of hippocampal activity during tests of recognition memory have revealed two patterns of effects. Several studies have reported that test items eliciting recollection of episodic details (indexed by successful source memory judgments or by 'Remember' judgments) are associated with larger hippocampal BOLD responses than are items recognized through an acontextual sense of familiarity (e.g., Eldridge et al., 2000; Dobbins et al., 2003; Weis et al., 2004; Wheeler and Buckner, 2004; Woodruff et al., 2005; Yonelinas et al., 2005; Vilberg and Rugg, 2009). Additionally, in some of these fMRI experiments (e.g., Yonelinas et al., 2005; Vilberg and Rugg, 2009; see also Rombouts et al., 2001; Rugg et al., 2003; Daselaar et al., 2006), it was also reported that studied items elicited less hippocampal activity than unstudied items did.

*Corresponding author: Maki Suzuki, Center for the Neurobiology of Learning and Memory, and Department of Neurobiology and Behavior, University of California, Irvine, Irvine, CA 92697-3800, USA, Phone: +1-949-824-8861, Fax: 949-824-4807, msuzuki@uci.edu.

Findings that hippocampal activity is enhanced by test items recognized on the basis of recollection rather than familiarity converge with the results of neuropsychological studies to suggest that recollection is more dependent upon the hippocampus than familiarity (for review, see Eichenbaum et al., 2007). It has also been proposed, however, that these fMRI findings might reflect a role for the hippocampus in the support of ‘strong’ memories, regardless of whether the memories support recollection or familiarity (Squire et al., 2007; Wais, 2008; Wais et al., in press). By contrast, hippocampal ‘new > old’ effects have usually been interpreted not in terms of a signal capable of supporting memory judgments, but rather as evidence of encoding-related activity in the hippocampus that is preferentially elicited by situationally novel stimulus events (e.g., Nyberg, 2005; Duzel et al., 2003; see also Stark and Okado, 2003). On the basis of an analysis of single neuron firing patterns in the human hippocampus, however, it has been argued that decrements in the activity elicited by studied test items relative to unstudied items constitute a memory signal that can support both familiarity- and recollection-based memory judgments (Rutishauser et al., 2006, 2008). To the extent that these single neuron findings can be extrapolated to fMRI, they raise the possibility that fMRI ‘new > old’ effects might indeed constitute a signal that can support memory judgments.

The great majority of fMRI studies of memory retrieval, including all of the studies that have described recollection-related enhancement of hippocampal BOLD activity, have employed discrete study and test phases separated by intervals of at least several minutes. Four fMRI studies have been reported in which memory was tested with a continuous recognition procedure (Brozinsky et al., 2005; Huijbers et al., in press; Johnson et al., 2008; Yassa and Stark, 2008). In this procedure the task is to discriminate between test items presented in a single list for the first or a subsequent time. [A reviewer queried the scientific value of continuous recognition, asking whether ‘it better reflects our experience of daily life’ than the study-test procedure. We do not know whether one procedure should be regarded as better than the other, but note that there are numerous ‘real-life’ circumstances – for example, at a party - where one has to switch unpredictably and frequently between encoding novel stimulus events (the identity of someone you are meeting for the first time) and recognizing an old event (a person you have already met).] Yassa and Stark (2008) did not identify any hippocampal effects in their study. In the study of Huijbers et al. (in press) hippocampal activity elicited by repeated items increased as a function of inter-item lag. In both Brozinsky et al. (2005) and Johnson et al. (2008), retrieval-related hippocampal activity was greater for new than for old items (see Viskontas et al, 2006, for analogous findings in a study of hippocampal single neuron responses during continuous recognition).

In the study of Johnson et al. (2008), test items were presented a total of four times, affording the opportunity to independently contrast the activity elicited by items according to whether they were new or old, and the activity elicited by old items as a function of number of presentations. As just noted, new > old effects were evident throughout the hippocampus. Strikingly, whereas behavioral accuracy (and, presumably, ‘memory strength’) increased with number of presentations, no hippocampal voxels could be identified in which activity elicited by old items also increased as a function of presentation number. By contrast, several hippocampal clusters were identified where the activity elicited by old items declined linearly with repetition.

The findings of Johnson et al. (2008) are difficult to reconcile with the proposal that increases in the memory strength of test items are associated with increases in the level of hippocampal activity elicited by the items (Squire et al., 2007; Wais, 2008; Wais et al., in press). The findings are however easily accommodated by the idea that memory strength is signaled not only by enhancement of hippocampal responses, but also by response reductions (as proposed by Rutishauser et al., 2006, 2008). Johnson et al. (2008) rejected

this possibility in favor of the ‘standard’ account of hippocampal new > old effects, namely, that the effects reflect novelty-induced encoding activity and not an on-line memory signal. They argued that the absence of retrieval-related enhancement of hippocampal activity in their study reflected the use of item familiarity - which, they argued, does not depend upon the hippocampus - as the sole basis for recognition. Johnson et al. (2008) predicted that in a continuous recognition task that necessitated the items be recollected, recollection would be accompanied by the same neural signature as in studies employing separate study and test phases, namely, a relative enhancement of hippocampal activity.

The goal of the present study was to test this prediction. We combined a continuous recognition procedure with an ‘exclusion’ task (Jacoby, 1991) that can be performed accurately only if recollection is successful. In the exclusion tasks subjects must discriminate between test items according to which of two study contexts the items had originally been presented in, classifying items that were studied in one of the contexts (the ‘target’ context) as ‘old’, and classifying items studied in the alternate context (the ‘non-target’ context) as ‘new’. The rationale underlying the exclusion task is that a correct target judgment will be made not only when a test item elicits recollection of its study context, but also when recollection fails and the judgment is driven solely by familiarity. By contrast, a familiarity-based judgment on a non-target item will result in its misclassification as a target; accurate classification of a familiar non-target is possible only if its context is successfully recollected. Thus, the neural correlates of recollection can be operationalized in an exclusion task as the difference in the activity elicited by correctly and incorrectly categorized non-target test items.

MATERIALS AND METHODS

Subjects

Twenty-two subjects (10 males; 18–22 years of age; mean age 19.9) participated in the experiment. All subjects were right-handed, native English speakers with normal or corrected-to-normal vision, and with no history of neurological illness or other contraindications for MR imaging. In accordance with the requirements of the Institutional Review Board of UCI, which approved the study, informed consent was obtained from all subjects. Two subjects’ data were excluded because of inadequate memory performance (> 2 SDs below mean accuracy in at least one experimental condition). Thus, data from a total of 20 subjects (9 males) are reported here.

Stimulus materials

Stimuli were 288 color photographs of common objects. For each subject, these items were randomly sampled without replacement to create six item lists. Each of these lists was then used to form a continuous recognition list comprising 12 unrepeated (filler) items, and 36 critical items that were each repeated once (inter-item lag between 10 and 30 items; mean 18 items). The repeated critical items were distributed equally across three different experimental conditions - ‘Old’, ‘Target’ and ‘Nontarget’ (see below). A separate pool of 24 photographs was used to create a practice list organized in the same way.

The items in each continuous recognition list were surrounded by a frame that could appear in one of three colors: grey, blue or orange. When an item was repeated, the color of its frame determined its experimental condition and hence the correct response. Regardless of frame color, items presented for the first time (New items) required a ‘new’ response. Items presented for the second time surrounded by a grey frame (Old items) required an ‘old’ response, regardless of the frame color associated with the first presentation. Crucially, repeated items surrounded by an orange or a blue frame required a different response

depending on whether the current color matched or did not match the color associated with the first presentation. When an item was re-presented in the *same* color as on the first presentation (Targets), the correct response was 'old'. When the item was re-presented in a *different* color (Nontargets), however, the correct response was 'new'.

Each continuous recognition list was organized into a series of 7 sub-blocks (as in Johnson et al., 2008), and each sub-block comprised 12 items. Except for the initial sub-block, which included new items only (six of which were fillers), each consisted of half new and half repeated items (for the last sub-block, all of the new items were fillers). Repeated items were assigned in equal proportions to the three experimental conditions (12 items each for Old, Target, and Nontarget). All behavioral and fMRI results reported here are based on data obtained from these final six sub-blocks.

During each continuous recognition run, experimental items were projected onto a screen viewed by the subject via a mirror mounted on the scanner head coil. The visual angle of each item (including the surrounding frame) subtended $9.5 \text{ deg} \times 9.5 \text{ deg}$. Presentation duration was 800ms, with a stimulus onset asynchrony (SOA) of 3250ms, excluding null trials (see below). A white central fixation cross (+) was presented throughout the inter-item interval, which varied between 2450 and 12200 ms, depending on the number of intervening null trials. Five null trials were added to the end of each run. Each scanned run consisted of 84 item trials, randomly intermixed with 28 null trials.

Experimental procedure

Instructions and practice for the task were given prior to scanning outside the scanner. Subjects responded with the index ('old') and middle ('new') fingers of their right hand, discriminating between the different categories of experimental item as described above. They were instructed to respond as quickly as possible without sacrificing accuracy.

fMRI scanning

Scanning was performed on a 3T Philips Achieva scanner (Philips Medical Systems, Andover, MA) equipped with an 8-channel SENSE head coil. Functional images were acquired using a T2*-weighted, FE-EPI pulse sequence (TR, 2000 ms; TE, 25 ms; flip angle, 70° ; matrix size, 176×176 ; FOV, $240 \times 180 \text{ mm}$; slice thickness, 1.5 mm; interslice gap, 0.5 mm; resolution, 1.5 mm^3 ; SENSE factor, 2.5; half-scan factor, 0.75). Twenty-six oblique axial slices were acquired parallel to the long axis of hippocampus and covering the whole MTL. One-hundred and ninety five volumes were acquired during each run. The first four volumes of each run were discarded to allow for T1 equilibration. After the six functional runs, whole-brain anatomical images were acquired using a sagittal T1-weighted, 3D MP-RAGE (TR, 11 ms; TE, 4.6 ms; flip angle, 18° ; matrix size, 320×320 ; FOV, $240 \times 150 \text{ mm}$; slice thickness, 0.75 mm; resolution, 0.75 mm^3).

Data Analysis

For two subjects, the data from one of the six functional runs were not analyzed. In one case this was because of a technical problem; in the second case the subject reported becoming confused about the task requirements during that run. Functional data preprocessing and statistical analyses were performed with Statistical Parametric Mapping (SPM5, Wellcome Department of Cognitive Neurology, London; <http://www.fil.ion.ucl.ac.uk/spm>) implemented in MATLAB7.0.4 (The Mathworks, Natick, MA). Functional data were initially spatially realigned to the first image of the series and then realigned to the across-run mean image and, subsequently, resliced. Each subject's anatomical data were coregistered to the mean functional image, and then segmented to allow estimation of deformation parameters for different tissue compartments. The anatomical and resliced

functional images were spatially normalized to a sample template generated from all subjects' anatomical images using the 'DARTEL' algorithm (Ashburner, 2007; http://www.fil.ion.ucl.ac.uk/~john/misc/dartel_guide.pdf) and affine-transformed into Montreal Neurological Institute (MNI) stereotactic space (Cocosco et al., 1997). The functional images were then resampled into 1.5 mm³ voxels, and spatially smoothed with a 3 mm full-width half-maximum (FWHM) Gaussian kernel.

Statistical analyses were performed in two stages of a mixed effects model. In the first stage, neural activity was modeled by a delta function (impulse event) at stimulus onset. These functions were then convolved with a canonical hemodynamic response function (HRF) and its temporal and dispersion derivatives (Friston et al., 1998) to yield regressors in a general linear model that modeled the BOLD response to each event type. Results from the HRF derivatives did not add any theoretically meaningful information to those from the canonical HRF and are not reported (they are available from the corresponding author upon request).

Six event types of interest were defined: correctly categorized New items (New), correctly judged Old items (Old), correctly and incorrectly detected Targets (Target hit and Target miss), and Nontargets attracting correct (correct rejection) and incorrect (false alarm) responses (Nontarget CR and Nontarget FA). Six regressors modeling movement-related variance (three rigid-body translations and three rotations determined from the realignment stage) and session-specific constant terms modeling the mean over scans were also used in the design matrix. Parameter estimates for events of interest were estimated using a general linear model. Nonsphericity of the error covariance was accommodated by an AR(1) model in which the temporal autocorrelation was estimated by pooling over suprathreshold voxels (Friston et al., 2002). Effects of interest were tested using linear contrasts of the parameter estimates. These contrasts were carried forward to a second stage of analysis treating subjects as a random effect.

The fMRI results reported here were restricted to hippocampus, amygdala and surrounding MTL (based on standard anatomical landmarks; Insausti et al., 1998). An MTL mask was created by manually tracing the MTL of the across-subjects mean normalized 3D anatomical image and then smoothing it with a 3-mm FWHM Gaussian kernel image using MRICron (Rorden et al., 2007). The mask was used in combination with a small volume correction (SVC) procedure (Worsley et al., 1996) to identify significant effects. Clusters were accepted as significant if the number of contiguously activated voxels (thresholded at $P < 0.005$) exceeded a corrected cluster-level threshold of $P < 0.05$ (corresponding in this data set to a cluster extent threshold of approximately 24 voxels). All coordinates are reported in MNI space.

In addition, using the MarsBaR toolbox (Brett et al., 2002), subject-specific parameter estimates for event types of interest were extracted from the voxels comprising the hippocampal clusters that exceeded the statistical threshold mentioned above. The parameter estimates were averaged across voxels to give a mean value for each cluster.

RESULTS

Behavioral performance

Table 1 shows the mean proportions of correct responses to each item-type for all subjects, and, additionally, for the High- and Low-performing sub-groups (see below). Probability of successful recollection (pR) was estimated as the difference between the probability of correctly endorsing a Target minus the probability of incorrectly endorsing a Nontarget (pTarget hit – pNontarget FA). As outlined in the Introduction, the rationale for this metric (first proposed by Jacoby, 1991) is that Targets can be correctly endorsed on the basis of

either recollection or familiarity, whereas a correct response to a Nontarget requires that familiarity is successfully ‘opposed’ by recollection. Thus, when recollection fails, familiar Nontargets are liable to be incorrectly endorsed as Targets. The difference between Target hit and Nontarget FA rates therefore indexes the probability of successful recollection. In the present case, mean pR was 0.44 ($SD = 0.21$), which was significantly greater than zero ($t(19) = 9.49, p < 0.001$).

Table 2 shows the mean response times (RTs) for each response category. A repeated-measures one-way ANOVA of the RTs revealed a significant effect of response category ($F(2.6, 49.6) = 19.77, p < 0.001$). Pairwise tests indicated that New responses were significantly faster than all other response categories (all $ps < 0.001$). Additionally, Old responses were significantly faster than the responses to all other classes of repeated item (all $ps < 0.01$). Finally, RTs for Target hits were faster than those for Target misses ($p < 0.05$) whereas Nontarget CRs and Nontarget FAs did not significantly differ ($p > 0.1$).

fMRI findings

We first identified where within the MTL activity differed between New and Old items (cf. Johnson et al., 2008). We then tested the prediction (see Introduction) that items associated with successful recollection would elicit greater hippocampal activity than items for which recollection failed.

The across-subject means (and ranges) of the trial numbers associated with New and Old items, Target hits and misses, and Nontarget CRs and FAs were 171.1 (148–179), 59.3 (41–71), 54.2 (41–71), 13.5 (1–23), 43.5 (21–63) and 23.8 (9–49) respectively. The subject (who fell into the ‘Low-performing sub-group’ as defined below) contributing only a single trial to the Target miss category was omitted from all analyses involving these items.

New-old effects—The contrast between New and Old items revealed robust new > old effects in the hippocampus bilaterally and the left parahippocampal gyrus (Fig. 1A; Table 3). Paired t-tests on the mean parameter estimates of each of the three hippocampal clusters (Fig. 1B) demonstrated that activity elicited by New items not only exceeded the activity elicited by old items, but also the activity elicited by Targets and Nontargets (all $ps < 0.01$, though note that these latter contrasts are not fully independent of the contrast employed to identify the clusters).

No MTL region could be identified where activity elicited by Old items exceeded that for New items.

Recollection effects—We assume that Nontargets were correctly endorsed as ‘new’ when recollection successfully opposed their familiarity, and were incorrectly endorsed as ‘old’ when recollection failed (Jacoby, 1991; see Introduction). Thus, we searched for regions associated with successful recollection by identifying where activity elicited by Nontarget CRs was greater than that elicited by Nontarget FAs.

The Nontarget CRs > Nontarget FAs contrast failed to identify any above-threshold MTL clusters. As is evident from Table 1, however, subjects varied widely in their ability to successfully recollect Nontargets (pR ranged from 0.13 to 0.81). We took advantage of this variability to conduct a second analysis in which we asked whether there were any hippocampal regions where recollection-related activity (still operationalized by the contrast between Nontarget CRs and Nontarget FAs) co-varied across subjects with the probability of successful recollection. This analysis identified a single cluster in the right hippocampus (Fig. 2A; Table 3). Within the cluster, correlations varied across voxels from a minimum of 0.56 ($t(18) = 2.87, p < 0.005$, 1-tailed) to a maximum of 0.72 ($t(18) = 4.40, p < 0.001$, 1-

tailed). As is evident from Fig. 2A, these correlations reflect a relationship between recollection performance and neural activity that, at the cluster level, is continuous across the range of both variables. (Note that while providing assurance that the correlation with performance of the recollection-related activity in this hippocampal cluster is not driven by outlying data points, Fig 2A does not constitute evidence of a correlation beyond that provided by the original SPM analysis; Kriegeskorte et al., 2009).

To further elucidate the recollection-related right hippocampal effect identified by the above correlation analysis, we separated the subjects into two sub-groups (High- and Low-performers) based on a median split of their recollection performance (giving mean pRs of 0.61 and 0.27, respectively). Our aims in performing these analyses were two-fold. First, we wanted to contrast the parameter estimates associated with high and low levels of recollection; while the foregoing correlational analysis indicates that the difference in activity between Nontarget CRs and Nontarget FAs increases as recollection improves, this finding provides no information about whether this increase reflects performance-related variation in the activity elicited by correction rejections, false alarms, or both response categories. Second, to the extent that the findings for Nontargets do indeed reflect a hippocampal correlate of recollection, one would predict similar results in the analogous contrast for Targets, that is, for Target hits (at least some of which would have been based on successful recollection) versus Target misses (when recollection was always unsuccessful). Importantly, this contrast employs data that are entirely independent of the data employed to identify the putative recollection-sensitive hippocampal cluster illustrated in Fig. 2A.

Mean parameter estimates for both Nontargets and Targets, separated according to subgroup, are shown in Fig. 2B. ANOVA of the Nontarget parameter estimates (factors of group and response accuracy) revealed a significant interaction between the two factors ($F(1, 18) = 17.13, p < 0.001$). Follow-up contrasts revealed a significant difference between CRs and FAs in the High- but not the Low-performing sub-group ($ps < 0.005$ and > 0.1 respectively, 1-tailed). In addition, between-group contrasts revealed significant differences for FAs ($t(18) = 3.71, p < 0.005$, 2-tailed) but not for CRs ($t(18) = 0.63, p > 0.1$, 2-tailed). Analogous to the Nontarget data, ANOVA of the parameter estimates associated with Targets also gave rise to a group $\times$ response accuracy interaction ($F(1, 17) = 4.98, p < 0.05$). Pairwise contrasts again revealed a significant difference between hits and misses in the High-performing group ($p < 0.05$, 1-tailed), but not in the Low-performing group ($p > 0.1$, 1-tailed). The magnitudes of the Target miss responses also differed significantly between the two groups ($t(17) = 2.51, p < 0.05$, 2-tailed), but the magnitudes of the Target hit responses did not ($t(17) = 0.76, p > 0.1$, 2-tailed). Mean parameter estimates for New items also did not differ between the two groups ($t(18) = 0.06, p > 0.1$, 2-tailed). Paired t-tests demonstrated that New items elicited greater activity in this cluster than any other response categories in both groups (all $ps < 0.05$, 1-tailed), with the exception of incorrect responses to Targets and Nontargets in the Low-performing group. (Note that these contrasts with New items are independent of the procedure employed to identify the cluster.)

DISCUSSION

The purpose of this study was to test the prediction that successful recollection during continuous recognition is associated with greater hippocampal activity than unsuccessful recollection (Johnson et al., 2008). Whereas a categorical contrast between estimates of the activity elicited by recollected and unrecollected items failed to support this prediction, it was supported by the findings of an across-subjects correlational analysis. This analysis identified a right hippocampal region where the difference in activity elicited by recollected versus unrecollected items covaried positively with recollection performance. In addition,

the present findings replicated prior reports of extensive hippocampal ‘new > old’ effects during continuous recognition (Brozinsky et al., 2005; Johnson et al., 2008).

To the best of our knowledge, the present experiment provides the first direct evidence for recollection-related enhancement of hippocampal activity in a continuous recognition task. The experiment brings the findings for continuous recognition into alignment with those obtained in studies employing separate study and test phases (see Introduction). Thus, regardless of the experimental procedure employed to assess memory performance, successful recollection – at least among subjects whose recollection performance is relatively good – is indexed by an enhancement of hippocampal BOLD activity.

Before considering the fMRI findings in more detail, one aspect of the recognition exclusion task employed here deserves mention. As was outlined in the Introduction, a key assumption underlying this procedure is that the ability to correctly exclude a Nontarget item depends on accurate recollection of the item’s study context, allowing its familiarity to be successfully ‘opposed’ (Jacoby, 1991). This assumption does not hold in the case of Nontarget items with low familiarity, however, since such items, even if not recollected, would receive a correct response by virtue of being classified (mistakenly) as new. An estimate of how frequently Nontargets were correctly endorsed through this ‘forgetting’ mechanism is given by the proportion of Target misses, that is, items for which recollection failed and familiarity was too low to support a ‘Target’ response. This proportion was approximately 20%, suggesting that the great majority of Nontarget CRs were indeed recollection-based.

Possible insights into the nature of the hippocampal BOLD ‘recollection signal’ can be gleaned from the patterns of activity in our ‘High-performing’ and ‘Low-performing’ subject groups in the recollection-sensitive hippocampal region identified by the correlational analysis (Fig. 2B). Three findings are worthy of comment. First, the difference in the activity of this region in the High-performers between correctly and incorrectly categorized Nontargets was also evident for the analogous categories of Targets. Whereas the finding for the Nontargets is arguably unsurprising in light of how this region was identified (Kriegeskorte et al., 2009), Target item activity in this region is unconstrained by this prior analysis. Thus the finding that, in High-performing subjects, the Target items most likely to have been recollected (hits) elicited the larger hippocampal responses is consistent with the proposal that activity in this hippocampal region indexes successful recollection (unlike in the contrast between Nontarget CRs and FAs, however, we cannot rule out a possible confound with RT, since Target hits were faster than misses). Additionally, it is worth noting that because the response contingencies for Targets and Nontargets were reversed, this putative hippocampal correlate of successful recollection is not confounded with downstream decisional or response selection processes.

The second point arising from the patterns of activity illustrated in Fig. 2B concerns the relative magnitudes of the hippocampal responses to the different categories of test item. As already noted, correct responses for both Nontargets and Targets were associated with greater hippocampal activity than incorrect responses in High- but not Low-performing subjects, consistent with the proposal that successful recollection is associated with relatively greater hippocampal activity. [A reviewer noted that one might have expected hippocampal activity to be higher for correctly endorsed Nontargets than Targets, since the former responses were more dependent on recollection. As noted above, however, around 20% of the Nontargets were likely correctly classified because of a combination of recollection failure and low familiarity. These items would act to dilute the recognition signal elicited by the remainder of the correctly endorsed Nontargets.] Strikingly the relatively poor recollection performance of the Low-performing group was associated not with a diminution (compared to the High-performers) in the magnitude of Target hit and

Nontarget CR responses, but by enhancement of the responses elicited by Target misses and Nontarget FAs, items supposedly associated with recollection failure. We tentatively account for these seemingly paradoxical findings by assuming that they reflect not so much a failure to recollect on the part of the Low-performers, as a failure to recollect *selectively*. By this argument (for related ideas, see Anderson and Bjork, 1994; Herron and Rugg, 2003), the poor discrimination between Target and Nontarget items in the Low-performing subjects came about because of an inability to focus retrieval cue processing so as to effectively target color information. Whereas repeated test items elicited retrieval of episodic information in these subjects (and, concomitantly, elicited relatively large hippocampal responses), this information was ‘non-diagnostic’ (Yonelinas and Jacoby, 1996; Mulligan and Hirshman, 1997) for color.

The final point to come out of Fig. 2B concerns the responses to New items. As was also true in three other hippocampal regions also (see below), New items elicited larger hippocampal responses in this ‘recollection-sensitive’ region than old items did (note that the results of these contrasts are unconstrained by the procedure employed to identify this region). This finding replicates prior observations of overlap within the hippocampus between new > old and recollection effects (e.g., Rugg and Yonelinas, 2003; Yonelinas et al., 2005; Vilberg and Rugg, 2009), and indicates that the overlap observed in earlier studies remains even when BOLD activity is acquired at a considerably higher spatial resolution than in those studies. This finding underscores the point made several times previously (e.g., Rugg et al., 1997; Martin, 1999; Stark and Okado, 2003; Woodruff et al., 2005) that ‘new’ items are a poor baseline against which to assess retrieval-related hippocampal activity, since these items elicit a hippocampal response in their own right (see below). That the same hippocampal region (at least at the spatial resolution afforded by fMRI) can be modulated both by the ‘novelty’ (or, equivalently, the ‘familiarity’) of an item, and the extent to which an item elicits recollection of a prior encoding episode, does nothing to simplify the interpretation of retrieval-related hippocampal activity.

As was noted in the Introduction, it has been argued that recollection-related enhancement of hippocampal activity is a consequence of the confounding effects of ‘memory strength’ (Squire et al., 2007; Wais, 2008; Wais et al., in press). By this argument, the methods that have been used to segregate recollected and unrecollected test items (for example, source memory and ‘Remember/Know’ procedures) actually segregate items according to the strength of an undifferentiated recognition signal. Hence, findings that hippocampal activity is selectively modulated by recollection can be reinterpreted as evidence that the hippocampus supports ‘strong’ memories, regardless of whether those memories contain qualitative episodic information (see Wais et al., in press for relevant findings). From this perspective, the present finding of a positive relationship between recollection performance and recollection-related hippocampal activity does not signify a selective role for the hippocampus in the retrieval of context information, but rather the greater memory strength of the test items for which such information can be retrieved. Thus, a similar relationship between strength and hippocampal activity should also be evident when only a simple recognition decision is required, rather than a decision that requires retrieval of source information. By themselves, the present data do not allow this prediction to be evaluated, since too few of the items requiring a simple recognition decision (i.e., the ‘Old’ items) were incorrectly endorsed as ‘new’ to permit contrasts between recognition hits and recognition misses. As was noted in the Introduction, however, the study of Johnson et al. (2008) was able to contrast old item activity across multiple presentations, but no hippocampal voxels could be identified where activity increased in concert with the additional memory strength that accrued with each presentation. Although more work is needed to address the relationship between hippocampal activity, ‘strength’, and recollection, these prior findings (see also Diana et al., in press), together with the present results, add weight to the notion

that retrieval-related enhancement of hippocampal activity is selectively associated with recollection of qualitative information.

In addition to the hippocampal regions that demonstrated recollection effects, robust new > old effects were also found in other hippocampal clusters where activity did not otherwise differ according to item type (see Fig. 1A, B). These findings replicate prior reports of hippocampal new > old effects during continuous recognition (Brozinsky et al., 2005; Johnson et al., 2008; see Viskontas et al., 2006 for analogous single neuron results). As was discussed in detail by Johnson et al. (2008), although these findings are consistent with the proposal that hippocampal new > old effects represent a memory signal that supports behavioral responding (as suggested by Rutishauser et al., 2006, 2008), there are grounds for doubting that this is generally the case. Most notably, the limited human single neuron data that are available (Rutishauser et al., 2006, 2008; Viskontas et al., 2006) suggest that hippocampal new > old effects take several hundreds of milliseconds to onset, in contrast to the considerably faster onset latency of seemingly analogous effects in primate perirhinal cortex (for review, see Brown and Xiang, 1998).

In conclusion, the present findings indicate that, as in experiments employing separate study and test phases, successful recollection of a prior study episode is associated with a relative enhancement of the hippocampal fMRI BOLD signal. Together with the results of a prior study (Johnson et al., 2008), these findings support the view that enhanced hippocampal activity is associated with successful retrieval of qualitative information (recollection), and does not signify retrieval of 'strong' memories regardless of whether the memories contain episodic content.

Acknowledgments

We thank the members of the University of California, Irvine Research Imaging Center for their assistance with fMRI data acquisition.

Grant sponsor: National Institute of Mental Health; **Grant number:** R01-MH072966

Grant sponsor: Japan Society for the Promotion of Science (to M.S.)

References

- Anderson, MC.; Bjork, RA. Mechanisms of inhibition in long-term memory: a new taxonomy. In: Dagenbach, D.; Carr, TH., editors. *Inhibitory Processes in Attention, Memory and Language*. San Diego: Academic Press; 1994. p. 265-326.
- Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage*. 2007; 38:95–113. [PubMed: 17761438]
- Brett, M.; Anton, JL.; Valabregue, R.; Poline, JB. Region of interest analysis using an SPM toolbox. The eighth international conference on functional mapping of the human brain; Sendai, Japan. 2002 June.
- Brown MW, Xiang JZ. Recognition memory: neuronal substrates of the judgment of prior occurrence. *Prog Neurobiol*. 1998; 55:149–189. [PubMed: 9618747]
- Brozinsky CJ, Yonelinas AP, Kroll NEA, Ranganath C. Lag-sensitive repetition suppression effects in anterior parahippocampal gyrus. *Hippocampus*. 2005; 15:557–561. [PubMed: 15889401]
- Cocosco C, Kollokian V, Kwan RS, Evans A. Brainweb: online interface to a 3D MRI simulated brain database. *Neuroimage*. 1997; 5:S425.
- Daselaar SM, Fleck MS, Cabeza R. Triple dissociation in the medial temporal lobes: recollection, familiarity, and novelty. *J Neurophysiol*. 2006; 96:1902–1911. [PubMed: 16738210]
- Diana RA, Yonelinas AP, Ranganath C. Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *J Cogn Neurosci*. in press.

- Dobbins IG, Rice HJ, Wagner AD, Schacter DL. Memory orientation and success: separable neurocognitive components underlying episodic recognition. *Neuropsychologia*. 2003; 41:318–333. [PubMed: 12457757]
- Duzel E, Habib R, Rotte M, Guderian S, Tulving E, Heinze HJ. Human hippocampal and parahippocampal activity during visual associative recognition memory for spatial and nonspatial stimulus configurations. *J Neurosci*. 2003; 23:9439–9444. [PubMed: 14561873]
- Eichenbaum H, Yonelinas AP, Ranganath C. The medial temporal lobe and recognition memory. *Annu Rev Neurosci*. 2007; 30:123–152. [PubMed: 17417939]
- Eldridge LL, Knowlton BJ, Furmanski CS, Bookheimer SY, Engel SA. Remembering episodes: a selective role for the hippocampus during retrieval. *Nat Neurosci*. 2000; 3:1149–1152. [PubMed: 11036273]
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. Event-related fMRI: characterizing differential responses. *Neuroimage*. 1998; 7:30–40. [PubMed: 9500830]
- Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J. Classical and bayesian inference in neuroimaging: applications. *Neuroimage*. 2002; 16:484–512. [PubMed: 12030833]
- Herron JE, Rugg MD. Strategic influences on recollection in the exclusion task: electrophysiological evidence. *Psychon Bull Rev*. 2003; 10:703–710. [PubMed: 14620367]
- Huijbers W, Pennartz CMA, Daselaar SM. Dissociating the “retrieval success” regions of the brain: Effects of retrieval delay. *Neuropsychologia*. in press.
- Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitkanen A. MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *Am J Neuroradiol*. 1998; 19:659–671. [PubMed: 9576651]
- Jacoby LL. A process dissociation framework: separating automatic from intentional uses of memory. *J Mem Lang*. 1991; 30:513–541.
- Johnson JD, Muftuler LT, Rugg MD. Multiple repetitions reveal functionally and anatomically distinct patterns of hippocampal activity during continuous recognition memory. *Hippocampus*. 2008; 18:975–980. [PubMed: 18548578]
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci*. 2009; 12:535–540. [PubMed: 19396166]
- Martin A. Automatic activation of the medial temporal lobe during encoding: lateralized influences of meaning and novelty. *Hippocampus*. 1999; 9:62–70. [PubMed: 10088901]
- Mulligan NW, Hirshman E. Measuring the bases of recognition memory: an investigation of the process-dissociation framework. *J Exp Psychol*. 1997; 23:280–304.
- Nyberg L. Any novelty in hippocampal formation and memory? *Curr Opin Neurol*. 2005; 18:424–428. [PubMed: 16003119]
- Rombouts SA, Barkhof F, Witter MP, Machielsen WC, Scheltens P. Anterior medial temporal lobe activation during attempted retrieval of encoded visuospatial scenes: an event-related fMRI study. *Neuroimage*. 2001; 14:67–76. [PubMed: 11525338]
- Rorden C, Karnath HO, Bonilha L. Improving lesion-symptom mapping. *J Cogn Neurosci*. 2007; 19:1081–1088. [PubMed: 17583985]
- Rugg MD, Fletcher PC, Frith CD, Frackowiak RS, Dolan RJ. Brain regions supporting intentional and incidental memory: a PET study. *Neuroreport*. 1997; 8:1283–1287. [PubMed: 9175130]
- Rugg MD, Henson RN, Robb WG. Neural correlates of retrieval processing in the prefrontal cortex during recognition and exclusion tasks. *Neuropsychologia*. 2003; 41:40–52. [PubMed: 12427564]
- Rugg MD, Yonelinas AP. Human recognition memory: a cognitive neuroscience perspective. *Trends Cogn Sci*. 2003; 7:313–319. [PubMed: 12860190]
- Rutishauser U, Mamelak AN, Schuman EM. Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. *Neuron*. 2006; 49:805–813. [PubMed: 16543129]
- Rutishauser U, Schuman EM, Mamelak AN. Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. *PNAS*. 2008; 105:329–334. [PubMed: 18162554]
- Squire LR, Zola-Morgan J, Clark RE. Recognition memory and the medial temporal lobe: a new perspective. *Nat Rev Neurosci*. 2007; 8:872–883. [PubMed: 17948032]

- Stark CEL, Okado Y. Making memories without trying: medial temporal lobe activity associated with incidental memory formation during recognition. *J Neurosci.* 2003; 23:6748–6753. [PubMed: 12890767]
- Vilberg KL, Rugg MD. An investigation of the effects of relative probability of old and new test items on the neural correlates of successful and unsuccessful source memory. *Neuroimage.* 2009; 45:562–571. [PubMed: 19146963]
- Viskontas IV, Knowlton BJ, Steinmetz PN, Fried I. Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. *J Cogn Neurosci.* 2006; 18:1654–1662. [PubMed: 17014370]
- Wais PE. fMRI signals associated with memory strength in the medial temporal lobes: a meta-analysis. *Neuropsychologia.* 2008; 46:3185–3196. [PubMed: 18817791]
- Wais PE, Squire LR, Wixted JT. In search of recollection and familiarity signals in the hippocampus. *J Cogn Neurosci.* in press.
- Weis S, Specht K, Klaver P, Tendolkar I, Willmes K, Ruhlmann J, Elger CE, Fernández G. Process dissociation between contextual retrieval and item recognition. *Neuroreport.* 2004; 15:2729–2733. [PubMed: 15597043]
- Wheeler ME, Buckner RL. Functional-anatomic correlates of remembering and knowing. *Neuroimage.* 2004; 21:1337–1349. [PubMed: 15050559]
- Woodruff CC, Johnson JD, Uncapher MR, Rugg MD. Content-specificity of the neural correlates of recollection. *Neuropsychologia.* 2005; 43:1022–1032. [PubMed: 15769488]
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp.* 1996; 4:58–83. [PubMed: 20408186]
- Yassa MA, Stark CEL. Multiple signals of recognition memory in the medial temporal lobe. *Hippocampus.* 2008; 18:945–954. [PubMed: 18493928]
- Yonelinas AP, Jacoby LL. Noncriterial recollection: familiarity as automatic, irrelevant recollection. *Conscious Cogn.* 1996; 5:131–141.
- Yonelinas AP, Otten LJ, Shaw KN, Rugg MD. Separating the brain regions involved in recollection and familiarity in recognition memory. *J Neurosci.* 2005; 25:3002–3008. [PubMed: 15772360]

Figure 1.

Hippocampal regions associated with new > old effects. A: The bilateral hippocampus regions (a, b, and c) exhibiting differences between New > Old items. Effects ($p < 0.005$) are overlaid onto coronal sections (left, $y = -16.5$; right, $y = -10.5$) of across-subjects mean normalized anatomical images. L, left hemisphere. B: Graphs of the mean parameter estimates (with standard errors) for each response category derived from three hippocampal clusters.

Figure 2.

Hippocampal region associated with recollection. A: The right hippocampal region where recollection-related activity (operationalized by the contrast of Nontarget CRs > Nontarget FAs) co-varied across subjects ($p < 0.005$) with probability of successful recollection. The effect is overlaid on a coronal section ($y = -10.5$) of the across-subjects mean normalized anatomical image. The associated scatter plot graphs the relationship across-subjects between recollection performance and mean recollection effect (mean across-voxel parameter estimates) in the illustrated hippocampal cluster. L, left hemisphere. B: Graphs of the mean across-voxel parameter estimates for each response category (with standard errors) from the right hippocampal cluster of High-performing and Low-performing subject groups.

Table 1

Mean (SD) proportions of correct responses to each item-type for all subjects, and for High- and Low-performing sub-groups.

	all subjects	Sub-groups	
		High-performers	Low-performers
New	0.99 (0.01)	0.99 (0.01)	0.99 (0.01)
Old	0.88 (0.10)	0.92 (0.06)	0.84 (0.12)
Target hit	0.80 (0.09)	0.82 (0.07)	0.78 (0.10)
Nontarget CR	0.64 (0.20)	0.79 (0.09)	0.49 (0.16)

Table 2

Mean (SD) RTs for each response category for all subjects.

all subjects	
New	1081 (340)
Old	1201 (304)
Target hit	1310 (371)
Target miss	1445 (504)
Nontarget CR	1439 (404)
Nontarget FA	1461 (500)

Table 3

Hippocampal regions showing New > Old and Recollection effects.

Brain region	L/R	No. of voxels	coordinates			Z score	P (corrected)
			x	y	z		
<i>New > Old effects</i>							
Hippocampus	L	70	-22.5	-18	-16.5	4.96	0.001
Hippocampus	R	70	25.5	-16.5	-19.5	4.55	0.001
Hippocampus	R	62	21	-10.5	-16.5	3.83	0.001
Parahippocampal gyrus	L	24	-25.5	-34.5	-16.5	4.24	0.044
<i>Recollection effects (across-subjects correlation)</i>							
Hippocampus	R	34	25.5	-10.5	-21	3.53	0.016

Z values refer to the peak of the activated cluster. Ps are cluster-wise values corrected for multiple comparisons after small volume correction within a medial temporal lobe mask (see methods). L, left; R, right hemisphere.