Decrements in Hippocampal Activity with Item Repetition during Continuous Recognition: An fMRI Study

Maki Suzuki[1,2], Jeffrey D. Johnson[1], and Michael D. Rugg[1]

Abstract

■ fMRI (1.5 mm isotropic voxels) was employed to investigate the relationship between hippocampal activity and memory strength in a continuous recognition task. While being scanned, subjects were presented with colored photographs that each appeared on four occasions. The requirements were to make one response when an item was presented for the first or the third time and to make a different response when an item appeared for the second or the fourth time. Consistent with prior findings, items presented for the first time elicited greater hippocampal and parahippocampal activity than repeated items. The activity elicited by repeated items declined linearly as a function of number of presentations ("graded" new > old effects). No medial-temporal lobe regions could be identified where activity elicited by repeated items exceeded that for new items or where activity elicited by repeated items increased with number of presentations. These findings are inconsistent with the proposal that retrieval-related hippocampal activity is positively correlated with memory strength. We also identified graded new > old effects in several cortical regions outside the medial-temporal lobe, including the left retrosplenial/posterior cingulate cortex and the right lateral occipito-temporal cortex. By contrast, graded old > new effects were evident in bilateral mid-intraparietal sulcus and precuneus. ■

INTRODUCTION

fMRI studies have demonstrated that neural responses in the hippocampus elicited by recognition memory test items differ according to the study status of the item and the nature of the associated behavioral judgment. Two patterns of differential activity have been observed. In several studies, correctly endorsed "old" test items were reported to elicit a smaller response than correctly classified "new" items (e.g., Vilberg & Rugg, 2009; Daselaar, Fleck, & Cabeza, 2006; Duzel et al., 2003; Rugg, Henson, & Robb, 2003; Stark & Okado, 2003; Rombouts, Barkhof, Witter, Machielsen, & Scheltens, 2001). This finding has frequently been interpreted as a hippocampal "novelty" response that reflects a bias to encode experimentally novel events into memory (e.g., Nyberg, 2005; Stark & Okado, 2003). A second, more frequently reported, pattern of retrieval-related hippocampal activity takes the form of relatively greater old item activity when recognition is accompanied, rather than unaccompanied, by retrieval of information about the study context, as indexed either by phenomenal report (e.g., Woodruff, Johnson, Uncapher, & Rugg, 2005; Yonelinas, Otten, Shaw, & Rugg, 2005; Wheeler & Buckner, 2004; Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000) or accurate source memory (e.g., Vilberg & Rugg, 2009; Weis et al., 2004; Dobbins, Rice, Wagner, & Schacter, 2003; Cansino, Maquet, Dolan, & Rugg, 2002).

These findings have been taken as support for the view that the hippocampus supports the recollection of episodic information but that it does not contribute to the undifferentiated sense of familiarity that also allows old and new recognition test items to be discriminated (e.g., Eichenbaum, Yonelinas, & Ranganath, 2007).

This interpretation of recollection-related enhancement of hippocampal activity has recently been challenged (Wais, Squire, & Wixted, 2010; Wais, 2008; Squire, Wixted, & Clark, 2007). These authors have argued that the relevant studies invariably confounded recollection with the "strength" of the memory signal supporting recognition, as operationalized by the confidence with which a test item is (or would have been) judged as old on a test of old/new recognition. According to Wais et al. (2010), if strength is equated between recollected and unrecollected items (in their case, by restricting analysis of the effects of source accuracy to studied items attracting confident "old" judgments on an immediately preceding recognition test), both classes of item elicit greater levels of hippocampal activity than do items of lesser strength, indicating that the hippocampus contributes to "strong" memories regardless of whether the memories contribute to recollection or familiarity.

We obtained findings relevant to this issue in a recent study that employed a continuous recognition procedure (Johnson, Muftuler, & Rugg, 2008). Test items (words and pictures) were presented a total of four times, and subjects were required to respond "new" to items seen for the first time and "old" to any repeated item. Accuracy

of old judgments—and, presumably, memory strength—increased markedly across the successive presentations of each test item. The activity elicited by new items was enhanced relative to the activity associated with correctly endorsed old items in several hippocampal regions, but no regions could be identified where activity was greater for old than for new items or where activity increased as a function of the number of old item presentations. Thus, there was no evidence for a positive association between hippocampal activity and memory strength, as would be expected from the perspective of Wais et al. (2010; see above). Johnson et al. (2008) speculated that their findings reflected the failure of their subjects to engage recollection when discriminating between old and new test items, the subjects relying instead on the less demanding strategy of evaluating the relative familiarity of the test items. Thus, item repetition increased familiarity strength but not the likelihood of recollection.

This conjecture receives support from the findings of a second study (Suzuki, Johnson, & Rugg, in press) that also employed a continuous recognition procedure. In this study, however, accurate task performance necessitated the engagement of recollection. Test items were each surrounded by a colored frame. The requirement was to respond "new" to items presented for the first time but to respond "old" to a repeated item only if its surrounding frame matched the frame color of the item when it was first presented, responding "new" otherwise. Thus, a correct response to an old item with a "mismatching" frame necessitated recollection of the frame color associated with its first presentation—reliance on familiarity alone would result in an incorrect response (Jacoby, 1991). Consistent with the view that retrieval-related enhancement of hippocampal activity is a signature of successful recollection, Suzuki et al. (in press) identified a right hippocampal region where recollection-related activity (operationalized as the difference in activity elicited by mismatching old items attracting correct vs. incorrect responses) positively correlated across subjects with probability of successful color recollection.

Together, the findings of Suzuki et al. (in press) and Johnson et al. (2008) are consistent with the proposal that, during continuous recognition, old items elicit enhanced hippocampal activity when they are recollected and not merely when they elicit a "strong" memory. It can be argued however that comparison of the findings of the two studies is complicated by a task confound; whereas in the study of Johnson et al., the task was simply to discriminate between new and old items, in the study of Suzuki et al., it was further necessary to discriminate between different classes of old item. These differing task demands raise the possibility that the disparate findings of the two studies reflect differences not in the nature of the retrieved information (i.e., familiarity vs. recollection) but in how that information was used in service of the memory test. To address this issue, in the present study, we employed the same basic procedure as in Johnson et al., but with the requirement

to discriminate between test items on the basis of the number of times the items had been presented. Subjects were required to make one response when an item was presented for the first or the third time and to make a different response when an item appeared for the second or the fourth time. Thus, as was the case in Suzuki et al., it was necessary to discriminate between different classes of old item to meet task requirements. Importantly, although the present task involves a more fine-grained judgment than does new/old recognition, we assumed that the primary basis for what amounts to a discrimination based on frequency of occurrence would be the relative familiarity (the "familiarity strength") of the test items (cf. Chalmers, 2005; Shiffrin, 2003; Hintzman, 1988) rather than recollected content. If the failure to identify hippocampal regions where activity scaled positively with memory strength in the study of Johnson et al. merely reflected the use of an experimental task that did not require discrimination between the memory signals elicited by different classes of old test items, then strength-sensitive regions should be identified in the present study, given that the present task necessitates just such a discrimination. However, if enhancement of retrieval-related hippocampal activity in the study of Suzuki et al. is linked specifically with recollection, the findings from the present study should replicate those of its predecessor: abundant evidence of hippocampal "new > old effects," but no evidence that activity in hippocampal regions is positively associated with the memory strength of old items (as operationalized by number of prior item presentations).

Although the primary aim of the present study was to further elucidate the relationship between hippocampal activity and memory strength, we took advantage of the relatively extensive field of view of our high-resolution imaging protocol to also investigate retrieval-related activity in cortical regions outside the medial-temporal lobe (MTL)—including retrosplenial and posterior cingulate cortex, medial parietal cortex, and intraparietal sulcus (IPS)—that have been implicated in numerous previous studies of successful recognition memory (Vilberg & Rugg, 2009; Shannon & Buckner, 2004; for reviews, see Ciaramelli, Grady, & Moscovitch, 2008; Skinner & Fernandez, 2007; Wagner, Shannon, Kahn, & Buckner, 2005; Rugg & Henson, 2002). Thus, we were able to ask which, if any, of these extrahippocampal regions demonstrated response profiles that paralleled profiles evident within the hippocampus.

METHODS

Subjects

Eighteen subjects (8 men; age range = 18–31 years; mean age = 21.2 years) participated in the experiment. All subjects reported themselves to be right-handed, native English speakers with normal or corrected-to-normal vision, and 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 item category). Thus, data from a total of 16 subjects (8 men) are reported here.

Stimulus Materials

Stimuli were 360 color photographs of common objects. For each subject, these items were randomly sampled to create six item lists. Each of these lists was then used form a continuous recognition list comprising 39 filler items, and 21 critical items that were each presented between one and four times (interitem lag between 3 and 34 items; mean = 16 items)—hereafter referred to as the first, second, third, and fourth presentations. Allocation of items to the lists and experimental conditions was randomized on a subject-specific basis.

Each continuous recognition list was organized into a series of seven subblocks (as in Johnson et al., 2008). After the initial three subblocks, each subblock consisted of half new and half repeated items and included an equal number of critical items from every experimental condition (12 items each for the items presented for first, second, third, and fourth time). All behavioral and fMRI results reported here are based on data obtained from these final four subblocks.

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. Items were presented one by one, each surrounded by a gray frame. The visual angle of each item (including the surrounding frame) subtended 9.5° × 9.5°. Presentation duration was 500 msec, with an SOA of 2400 msec excluding null trials (see below). A white central fixation cross (+) was presented throughout the interitem interval, which varied between 1900 and 9100 msec, depending on number of intervening null trials. Five null trials were added to the end of each run. Each scanned run consisted of 105 item trials, randomly intermixed with 35 null trials.

Experimental Procedures

Instructions and practice for the task were given prior to the scanning, outside the scanner. Subjects responded with the index ("first" and "third" presentations) and middle ("second" and "fourth" presentations) 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 3-T Philips Achieva scanner (Philips Medical Systems, Andover, MA) equipped with an eight-channel SENSE head coil. Functional images were acquired using a T2*-weighted, FE-EPI pulse sequence (repetition time = 2000 msec, echo time = 25 msec, flip angle = 70°, matrix size = 176 × 176, field of view = 240 × 180 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 sixty-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 (repetition time = 11 msec, echo time = 4.6 msec, flip angle = 18°, matrix size = 320 × 320, field of view = 240 × 150 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 due to a technical problem; in the second case, the subject felt asleep 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 the Montreal Neurological Institute (MNI) stereotactic space (Cocosco, Kollokian, Kwan, & Evans, 1997). The functional images were then resampled into 1.5 mm^3 voxels and spatially smoothed with a 3-mm 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 judged items presented for the first time, correctly judged items presented for the second time, and both correctly and incorrectly judged items presented for third time and

fourth time. Null trials were modeled as an implicit baseline. 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.

Parameter estimates for each event type of interest were derived for each subject. The estimates were entered into the second stage analysis in which each subject served as a random effect in a repeated measures ANOVA. The ANOVA used a common error term for all effects and was corrected for nonsphericity using a correction factor that was estimated after pooling over all suprathreshold voxels (Henson & Penny, 2003). The ANOVA of the main effect of number of presentations as well as all subsidiary pairwise contrasts (for details, see Results) were conducted within this ANOVA model.

To evaluate the reliability of effects within the MTL, a mask was created by manually tracing the MTL of the across-subjects mean-normalized 3D anatomical image (based on standard anatomical landmarks; Insausti et al., 1998) and then smoothing the mask with a 3-mm FWHM Gaussian kernel using MRIcron (Rorden, Karnath, & Bonilha, 2007). Unless otherwise noted, clusters were accepted as significant only if the number of contiguously activated voxels (two-tailed threshold of $p < .005$) exceeded a corrected cluster-level threshold of $p < .05$. The corresponding cluster extent thresholds were 19 voxels for the area within the MTL mask and 29 voxels for the remainder of the imaging volume, as determined by Monte Carlo simulations implemented in AlphaSim (Ward, 2000). The results of the ANOVA were used to identify in an unbiased manner all clusters where activity differed between the events of interest. The outcome of the ANOVA was then employed as a mask (ROI) within which specific directional contrasts were implemented (see below). These contrasts were thresholded at $p < .005$ (one-tailed), again with a corrected cluster extent threshold of $p < .05$. All coordinates are reported in MNI space.

Following Johnson et al. (2008), we attempted to distinguish between regions demonstrating categorical and graded new/old effects by estimated two differently weighted contrasts (categorical: $+3 -1 -1 -1$ and $-3 +1 +1 +1$ for correctly judged first, second, third, and fourth presentations, respectively; graded: $+3 +1 -1 -3$ and $-3 -1 +1 +3$). Preliminary analyses revealed that these two sets of weights gave rise to virtually identical results (approximately 90% of the voxels identified by the two contrasts overlapped). Therefore, we reported below the findings derived

from the graded new/old contrasts only. To identify regions where activity varied with response accuracy, we contrasted the activity elicited by correctly versus incorrectly judged third and fourth presentations (too few errors were made to second presentations to allow an analysis by accuracy). Finally, to assess whether hippocampal activity is positively related to memory strength (see Introduction), we searched for regions where old item activity increased as a function of presentation number by applying the contrast weights -1, 0, and 1 to correctly judged second, third, and fourth presentations, respectively.

To further elucidate significant effects, subject-specific parameter estimates for events of interest were extracted from the voxels of each cluster that exceeded the statistical thresholds described above. The parameter estimates were averaged across voxels to a give mean value for each cluster. Because we were unable to identify any clusters where activity varied according to behavioral accuracy, estimates for correctly and incorrectly endorsed items presented for third and fourth times were each averaged to give a single value.

In light of the findings of Huijbers, Pennartz, and Daselaar (2010; for details, see Discussion), we performed a second set of analyses that asked whether activity elicited by old items varied according to the number of items intervening between successive presentations (interitem lag effects). We performed separate analyses on items presented for the second, third, and fourth times. In each item category, we contrasted the activity elicited by the items according to whether they were repeated after a relatively short (lags of 3–14 items) or a relatively long (lags of 19–34 items) interitem lag.

RESULTS

Behavioral Performance

Table 1 shows the mean proportion of correct responses and the mean RTs to each item category (i.e., first through to fourth presentations). A one-way repeated measures ANOVA of the proportion data revealed a significant effect, $F(1.9, 29.0) = 29.53, ps < .001$.[1] Pairwise tests revealed that judgments for new items were more accurate than the judgments for any other condition (all $ps < .001$). In addition, the accuracy for items presented for the second time was higher than for items presented for the third or fourth

Table 1. Mean $\pm$ *SD* Proportions of Correct Responses and Associated RTs to Each Item Category

Presentation Number	Proportions	RTs (msec)
First	0.99 ± 0.02	813 ± 136
Second	0.84 ± 0.10	1041 ± 123
Third	0.68 ± 0.15	1244 ± 125
Fourth	0.68 ± 0.15	1188 ± 159

times (both ps < .005). Accuracy for these latter two conditions did not significantly differ.

A one-way repeated measures ANOVA of the RTs also revealed a significant effect, $F(2.3, 35.1) = 116.77$, ps < .001. Pairwise tests revealed that responses to new items were significantly faster than those to any other type of event (all ps < .001). In addition, RTs to correctly endorsed second presentations were faster than the responses to items presented for the third or fourth time (all ps < .001). RTs for correct responses to these latter items did not differ.

To evaluate the effects of interitem lag on behavioral performance, we separated the mean proportions of correct responses and associated RTs according to presentation number (second, third, fourth) and interitem lag (short vs. long, see Methods) (Table 2). A two-way repeated measures ANOVA (factors of presentation number and lag) on the accuracy scores revealed a significant main effect of item presentation, $F(1.4, 21.6) = 9.35$, p < .005, again reflecting greater accuracy for second than for subsequent presentations. There was no significant lag effect, nor did lag interact with presentation number. ANOVA of the RTs also gave rise to a main effect of number of presentations, $F(1.61, 24.11) = 34.64$, p < .001, but no other significant effects.

fMRI Findings

We first analyzed the fMRI data in a manner unbiased with respect to specific directional hypotheses to identify clusters where activity differed between the item categories. We accomplished this with a one-way ANOVA of the main effect of number of presentations. Using the outcome of the ANOVA as a mask, we then searched for regions where there was an old versus new effect or where activity elicited by old items varied according to presentation number (see Methods).

Brain Regions Demonstrating a Main Effect of Number of Presentations

As is evident in Figure 1, the global ANOVA identified effects in bilateral hippocampus and parahippocampal cortex as well as in extra-MTL regions, including bilateral parietal cortex, precuneus, right lateral occipito-temporal cortex, left retrosplenial/posterior cingulate cortex, and cuneus.

Outcome of Specific Contrasts

Medial-Temporal Lobe. Graded new > old effects were observed in bilateral hippocampus and parahippocampal cortex (Figure 2A, Table 3). As is evident from Figure 2B, there was a tendency in all regions for a stepwise decrease in activity from first to fourth presentations (Figure 2B). No MTL regions were identified where activity was greater for old than for new items. Nor were there any MTL regions where old item activity was enhanced as a function of number of presentations (this remained the case even when the threshold for this contrast was relaxed to p < .05).[2]

Regions outside the MTL. Regions demonstrating graded new > old effects included right occipito-temporal cortex and left retrosplenial/posterior cingulate cortex (Figure 3A, Table 3). There was a tendency for a stepwise decrease in activity from first to fourth presentations in these regions (Figure 3B).

Graded old > new effects were evident in bilateral IPS, precuneus, and left cuneus (Figure 3A, Table 3). As is illustrated in Figure 3B, activity in these regions tended to increase as a function of number of presentations.

Effects of Interitem Lag

We did not detect any lag effects in the MTL that came near to our criteria for statistical significance. For fourth presentations only, we identified a single retrosplenial cluster (MNI coordinates: $x = -8$, $y = -53$, $z = 23$, Z value = 4.06, 29 voxels), where activity was greater for the long relative to short lags.

Effects of Response Accuracy

The contrast between third and fourth presentations as a function of response accuracy failed to identify any effects.

DISCUSSION
Behavioral Performance

We assume that, as was directly demonstrated by Johnson et al. (2008), memory strength of the test items increased with the number of prior presentations. Classification of new items in the present study was near ceiling, indicating that subjects had no difficulty in detecting "unstrengthened" situational unfamiliar items. The accuracy of the

Table 2. Mean ± *SD* Proportions of Correct Responses and Associated RTs According to Presentation Number and Interitem Lag

Presentation Number/Interitem Lag	Proportions		RTs (msec)	
	Short	*Long*	*Short*	*Long*
Second	0.87 ± 0.11	0.83 ± 0.12	1027.4 ± 115.9	1074.0 ± 150.7
Third	0.68 ± 0.16	0.68 ± 0.16	1251.5 ± 123.2	1256.8 ± 150.9
Fourth	0.70 ± 0.18	0.70 ± 0.14	1179.7 ± 153.2	1185.6 ± 182.4

Figure 1. Brain regions demonstrating a main effect of number of presentations. Effects ($p < .005$, two-tailed) are displayed on the SPM maximal intensity projection images.
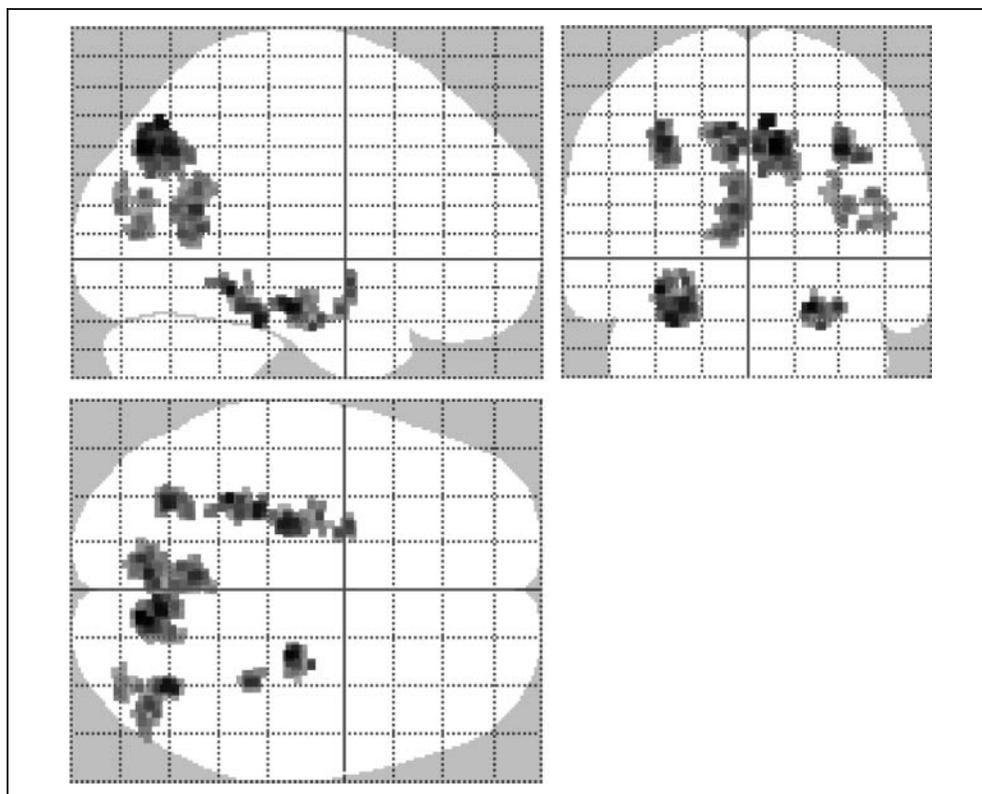
judgments accorded old items on their second presentation was substantially lower (around 80% correct) but was lower still on the two subsequent presentations (around 65% correct in each case). Thus, it would appear that subjects were rarely able recollect how many times an old test item had been presented, especially after its second presentation. This finding accords with the subjective experience of subjects undertaking this task, who reported that they quickly lost track of the number of times items have been presented and selected their responses on the basis of how familiar the item felt (cf. Chalmers, 2005; Shiffrin, 2003; Hintzman, 1988). However, we cannot rule out the possibility that recollection supported accurate responding on at least some proportion of trials.

Figure 2. MTL regions demonstrating graded new > old effects. (A) MTL regions exhibiting a decline in activity as a function of presentation number (indicated in blue). Effects ($p < .005$) are overlaid on coronal sections of the across-subjects mean-normalized anatomical image. (B) Bar graphs showing the mean parameter estimates (with standard errors) for each item category derived from four MTL clusters (a–d). MTL = medial-temporal lobe; L = left; R = right; PHG = parahippocampal gyrus.
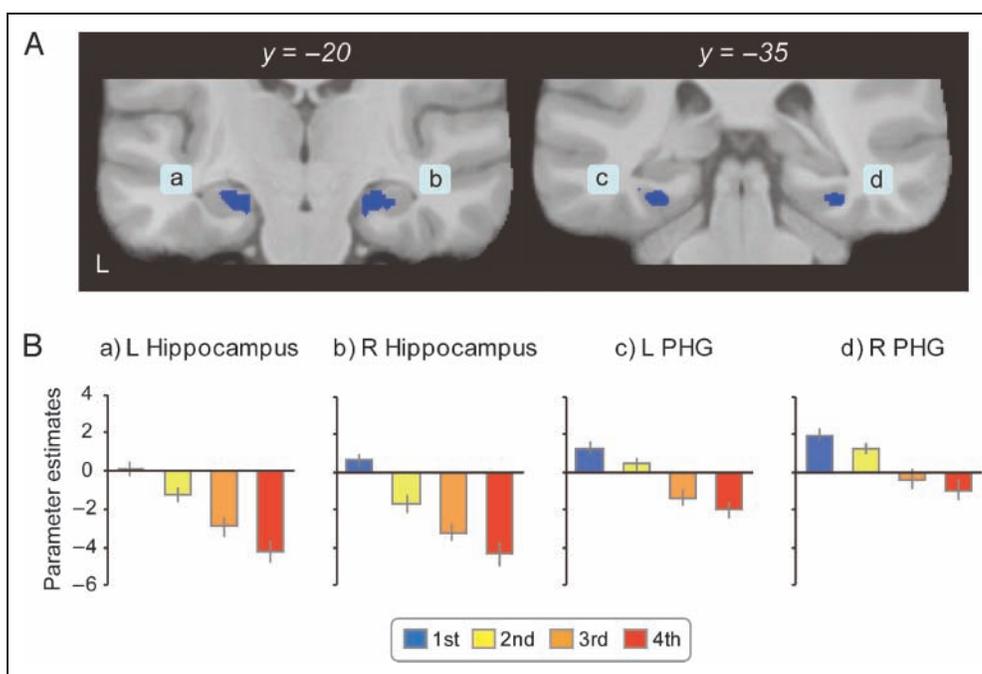
Table 3. Brain Regions Demonstrating Graded New > Old and Old > New Effects

Brain Region	L/R	No. Voxels	x	y	z	Z Score	Brain Region	L/R	No. Voxels	x	y	z	Z Score
Graded New > Old Effects							*Graded Old > New Effects*						
MTL regions													
Hippocampus	L	106	−24	−21	−17	5.06							
Hippocampus	R	89	23	−20	−17	5.70							
Parahippocampal cortex	L	141	−32	−42	−11	5.48							
Parahippocampal cortex	R	43	33	−35	−17	4.60							
Parahippocampal cortex	R	26	24	−39	−14	3.79							
Extra-MTL regions													
Occipito-temporal cortex	R	57	36	−83	18	3.72	IPS	L	194	−30	−65	41	5.07
Occipito-temporal cortex	R	64	45	−74	23	3.81	IPS	R	166	33	−65	39	5.4
Retrosplenial/posterior cingulate cortex	L	281	−5	−54	18	4.69	Precuneus	L	151	−6	−72	47	4.98
							Precuneus	R	416	9	−75	41	6.53
							Cuneus	L	39	−11	−77	9	3.82

Z values refer to the peak of the activated cluster. Coordinates (in MNI space) are rounded to the nearest mm. L = left; R = right; MTL = medial-temporal lobe.

Figure 3. Extra-MTL regions demonstrating graded new > old and old > new effects. (A) Cortical regions showing graded new > old effects (blue) and old > new effects (red). Effects ($p < .005$) are overlaid on sections of the across-subjects mean-normalized anatomical image. (B) Mean parameter estimates (with standard errors) for each item category derived from clusters in left retrosplenial (a), left precuneus (b), and left IPS (c). MTL = medial-temporal lobe; L = left; IPS = intraparietal sulcus.
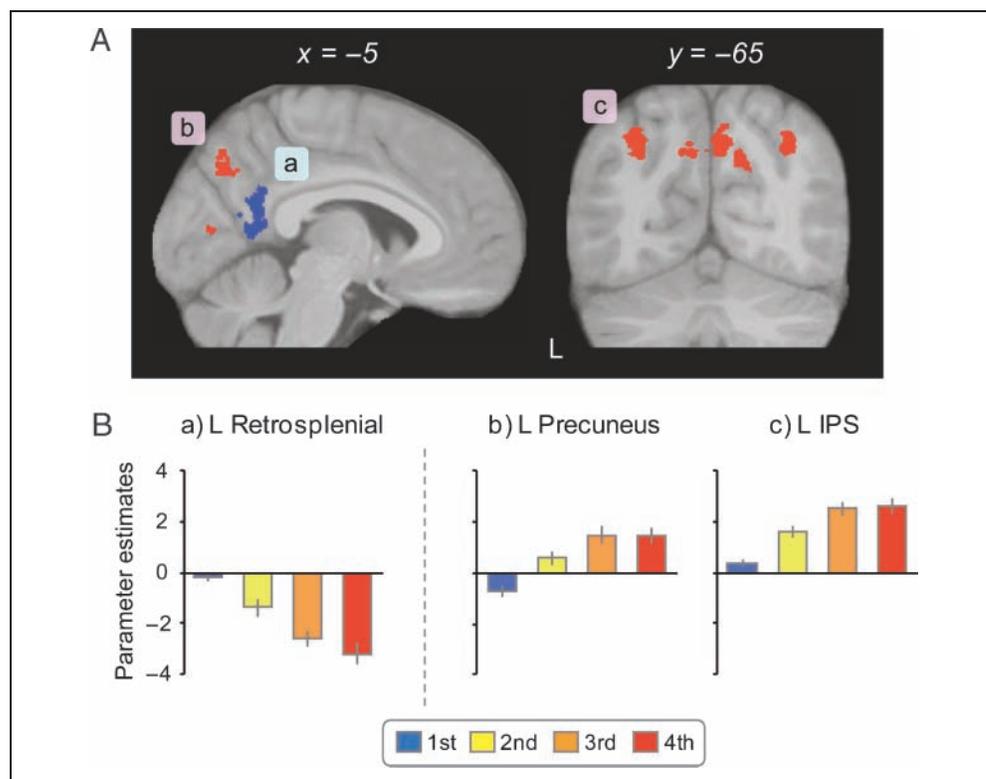
Retrieval-related Activity in the MTL

There were robust differences in the activity elicited by different categories of test items in both the hippocampus and the parahippocampal cortex. Crucially, these differences exclusively took the form of new > old effects; even at a much relaxed statistical threshold, we failed to find any MTL voxels where old item activity increased as a function of number of presentations. Thus, the present results join the findings of Johnson et al. (2008) in opposing the idea that retrieval-related hippocampal activity is positively associated with the strength of an undifferentiated memory signal (Wais et al., 2010; Wais, 2008; Squire et al., 2007; note that this conclusion stands regardless of whether responding in the present task was based largely upon the assessment of familiarity, as we assume, or upon recollection also). As in the study of Suzuki et al. (in press; see Introduction), the experimental task employed here required subjects to discriminate between different classes of old test item. Evidently, this task requirement is not sufficient to reverse the findings reported by Johnson et al. We hypothesized that the crucial factor is whether the memory signal employed to discriminate between old test items carries episodic information.

Our failure to find evidence for a positive association between hippocampal activity and memory strength is complemented by the findings from another recent fMRI study of continuous recognition memory (Huijbers et al., 2010; for similar but weaker findings, see Brozinsky, Yonelinas, Kroll, & Ranganath, 2005) as well as those of Daselaar et al. (2006). In the study of Huijbers et al. (2010), items were repeated once only, and the critical manipulation was the interitem lag between first and second presentations, which ranged from 0 to 31 intervening items. Longer lags were associated with lower recognition accuracy and longer response latencies, indicating that memory strength weakened as the interval between successive presentations increased. No hippocampal regions were reported where activity was negatively associated with lag, however; rather, activity in bilateral anterior hippocampus (and retrosplenial cortex) was reported to covary positively with lag (there was a hint of a similar effect in the present study for the retrosplenial cortex, but not for the hippocampus; see Results). The authors conjectured that their findings indicated that as lag increased, recognition became increasingly reliant on a hippocampally mediated recollection signal. Regardless of the validity of this account, the findings of Huijbers et al. add weight to the view that retrieval-related hippocampal activity is not positively correlated with memory strength (see also Diana, Yonelinas, & Ranganath, 2010).

Employing a study–test design, Daselaar et al. (2006) identified a region of the left anterior hippocampus where activity was negatively and linearly associated with recognition confidence (a small volume correction within a 3-mm sphere centered on the peak of that effect, MNI coordinates: $x = -19$, $y = -8$, $z = -16$, revealed a reliable,

$p < .005$, negatively graded effect in the present study also). In as much as confidence is a proxy for strength (Wixted, 2007), these findings, like the present results, suggest that activity in at least some hippocampal regions is a negative function of memory strength. In contrast to the present findings, however, Daselaar et al. also identified a parahippocampal region where activity was positively associated with recognition confidence. It is unclear which of the many differences between that study and the current one is responsible for this discrepancy. Another study reporting a positive relationship between memory strength (operationalized by recognition confidence) and MTL activity is that of Wais et al. (2010). These authors reported that studied test items attracting highly confident "old" judgments elicited enhanced hippocampal activity relative to items incorrectly endorsed as new (misses), regardless of the level of accuracy of a subsequent source memory judgment. If it is assumed that inaccurate source judgments signify the failure of recollection, these findings are indicative of a positive relationship between hippocampal activity and memory strength in the absence of recollection. As Wais et al. note, however, their findings are equally consistent with the assumption that hippocampal effects on trials associated with inaccurate source judgments were driven by recollection of task-irrelevant study details ("noncriterial recollection"; Yonelinas & Jacoby, 1996) rather than by familiarity strength.

As already noted, we identified robust new > old effects throughout the hippocampus in the present study. Reminiscent of the findings of Johnson et al. (2008), a substantial proportion of the voxels demonstrating these effects showed a graded response profile (specifically, 98% of the voxels identified by the initial ANOVA exhibited a graded profile according to the follow-up parametric analysis), such that activity declined in a linear fashion as a function of number of presentations (Figure 2B). This profile is indicative of a process sensitive to the relative familiarity (or novelty) of the test items rather than one that is responsive merely to whether an item is new or old. One possibility, consistent with prior accounts of hippocampal new > old effects (e.g., Nyberg, 2005; Stark & Okado, 2003), is that these findings reflect the graded modulation of a novelty-sensitive encoding process. That is, encoding processes that are engaged maximally by situationally novel stimulus events are progressively "dampened" as the event becomes more familiar. Alternatively, the present findings can be taken as evidence that retrieval-related hippocampal activity is indeed correlated with memory strength (or, more accurately, familiarity strength) but that the correlation is negative rather than positive. From this perspective, new > old effects could be construed as a hippocampally derived memory signal capable of supporting differential responding on the basis of the test items' study histories, much as has been argued for analogous effects in perirhinal cortex[3] (e.g., Aggleton & Brown, 2006; Henson, Cansino, Herron, Robb, & Rugg, 2003). Although Johnson et al. offered a number of reasons why hippocampal new > old effects

likely do not constitute such a memory signal, we are unaware of any evidence that speaks directly to this issue.

Retrieval-related Activity in Cortical Regions outside the MTL

In addition to the MTL, differences in the activity elicited by the different classes of test items were also evident in posterior cortical regions, including regions that have been implicated previously in successful recognition memory (Vilberg & Rugg, 2009; Shannon & Buckner, 2004; for reviews, see Ciaramelli et al., 2008; Skinner & Fernandez, 2007; Wagner et al., 2005; Rugg & Henson, 2002). Thus, graded new > old effects were evident in the left retrosplenial/posterior cingulate cortex, along with the right lateral occipito-temporal cortex, whereas graded old > new effects were found in the bilateral mid-IPS, precuneus, and right medial occipital cortex. Two general points arise from these findings. First, the dissociation between retrosplenial/posterior cingulate effects on the one hand and the medial and superior parietal effects on the other provides compelling evidence for functional heterogeneity between these two sets of regions (see also Huijbers et al., 2010), despite the fact that they have sometimes been described as belonging to a common "retrieval-success" network (e.g., Wagner et al., 2005; Shannon & Buckner, 2004). Second, whereas robust old > new effects were evident in the mid-IPS (see below), these effects did not extend to more lateral and posterior parietal regions—which were within the field of view of our functional imaging volume—where old > new effects have been reported to be associated selectively with successful recollection (e.g., Yonelinas et al., 2005; Wheeler & Buckner, 2004; for a review, see Vilberg & Rugg, 2008). Although null results should be interpreted cautiously, the insensitivity of this "recollection-sensitive" parietal region to the study status of the test items supports the assumption (see above) that old items elicited little or no recollection.

The graded new > old effects observed in the retrosplenial cortex and adjacent posterior cingulate parallel the effects observed in the hippocampus (and parahippocampal cortex). The retrosplenial cortex in particular is strongly interconnected with the hippocampus and parahippocampal cortex (Vann, Aggleton, & Maguire, 2009; Kahn, Andrews-Hanna, Vincent, Snyder, & Buckner, 2008; Kobayashi & Amaral, 2007) and has been proposed as a component (along with diencephalic structures) of an "extended hippocampal system" that selectively supports episodic memory (Aggleton, 2008). From this perspective, the parallel effects observed in these regions in the present study suggest that the different components of this system responded in concert to the level of familiarity (or novelty) of the test items. As noted above, the functional significance of these responses remains to be ascertained.

Graded new > old effects were also evident in lateral occipito-temporal cortex, on the border between Brodmann areas 39 and 19 (Figure 3A). Although it is possible that, like the retrosplenial cortex, the region is functionally coupled with the hippocampus, this seems unlikely in view of the lack of prior evidence for its involvement in memory. An alternative possibility is that the occipito-temporal new > old effects are an example of "repetition suppression"— the attenuation of item-elicited activity that is often held to be a neural correlate of repetition priming (Henson, 2003).

Old > new effects in the vicinity of the mid-IPS have been reported in numerous previous studies of recognition memory (Vilberg & Rugg, 2008). Although there is some consensus that IPS old/new effects are elicited in the absence of recollection (Cabeza, Ciaramelli, Olson, & Moscovitch, 2008; Vilberg & Rugg, 2008), their functional significance is debated. One possibility is that the effects reflect the role of this region as a mnemonic "accumulator," such that the activity elicited by a test item is proportional to the amount of evidence that the item is old (Donaldson, Wheeler, & Petersen, 2010; Vilberg & Rugg, 2009; Wagner et al., 2005). The present finding that IPS activity increases with memory strength (hence the amount of evidence that an item is old) is clearly compatible with this proposal. Graded old > new effects were also evident in the bilateral precuneus (Figure 3B). Like the IPS, old > new effects in this region are consistently reported in studies of recognition memory. On the basis of their finding that retrieval-related precuneus activity paralleled that in the IPS, Donaldson et al. (2010) proposed that like the latter region, the precuneus also acted as a mnemonic accumulator. Clearly, the present findings are also consistent with this proposal.

Conclusions

We identified robust hippocampal new > old effects in a continuous recognition task that required discrimination between the test items on the basis of number of prior presentations. These effects showed a graded response profile, such that their activity declined linearly as a function of number of presentations. We were unable to identify any hippocampal region where activity increased with number of study presentations. Together with prior findings (Suzuki et al., in press; Johnson et al., 2008), the present results strongly suggest that retrieval-related hippocampal activity is not positively correlated with memory strength.

Notes

1. The degrees of freedom for ANOVAs were adjusted for non-sphericity using the Greenhouse–Geisser correction.

2. A reviewer noted that it would be of interest to report the pattern of effects within the right hippocampal region identified as "recollection sensitive" in our previous study (Suzuki et al., in press). Mean parameter estimates within a 3-mm sphere centered on the peak voxel of the region (MNI coordinates: $x = 26$, $y = -11$, $z = -21$) demonstrated robust, graded new > old effect (estimates of 0.04, -1.08, -1.36, -2.00 for first to fourth presentations respectively, $p < .001$ two-tailed for the contrast between first and fourth presentations).

3. The high-resolution fMRI protocol employed in this study yielded low signal return in ventral anterior MTL in the vicinity of the perirhinal cortex, likely because of the vulnerability of this region to magnetic susceptibility artifact.

REFERENCES

Aggleton, J. P. (2008). EPS Mid-Career Award 2006. Understanding anterograde amnesia: Disconnections and hidden lesions. *Quarterly Journal of Experimental Psychology (Colchester), 61,* 1441–1471.

Aggleton, J. P., & Brown, M. W. (2006). Interleaving brain systems for episodic and recognition memory. *Trends in Cognitive Sciences, 10,* 455–463.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage, 38,* 95–113.

Brozinsky, C. J., Yonelinas, A. P., Kroll, N. E. A., & Ranganath, C. (2005). Lag-sensitive repetition suppression effects in anterior parahippocampal gyrus. *Hippocampus, 15,* 557–561.

Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: An attentional account. *Nature Reviews Neuroscience, 9,* 613–625.

Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cerebral Cortex, 12,* 1048–1056.

Chalmers, K. A. (2005). Basis of recency and frequency judgements of novel faces: Generalised strength or episode-specific memories? *Memory, 13,* 484–498.

Ciaramelli, E., Grady, C., & Moscovitch, M. (2008). Top–down and bottom–up attention to memory: A hypothesis (AtoM) on the role of the posterior parietal cortex in memory retrieval. *Neuropsychologia, 46,* 1828–1851.

Cocosco, C., Kollokian, V., Kwan, R. S., & Evans, A. (1997). Brainweb: Online interface to a 3D MRI simulated brain database. *Neuroimage, 5,* S425.

Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2006). Triple dissociation in the medial temporal lobes: Recollection, familiarity, and novelty. *Journal of Neurophysiology, 96,* 1902–1911.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2010). Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *Journal of Cognitive Neuroscience, 22,* 1808–1818.

Dobbins, I. G., Rice, H. J., Wagner, A. D., & Schacter, D. L. (2003). Memory orientation and success: Separable neurocognitive components underlying episodic recognition. *Neuropsychologia, 41,* 318–333.

Donaldson, D. I., Wheeler, M. E., & Petersen, S. E. (2010). Remember the source: Dissociating frontal and parietal contributions to episodic memory. *Journal of Cognitive Neuroscience, 22,* 377–391.

Duzel, E., Habib, R., Rotte, M., Guderian, S., Tulving, E., & Heinze, H. J. (2003). Human hippocampal and parahippocampal activity during visual associative recognition memory for spatial and nonspatial stimulus configurations. *Journal of Neuroscience, 23,* 9439–9444.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30,* 123–152.

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience, 3,* 1149–1152.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *Neuroimage, 7,* 30–40.

Friston, K. J., Glaser, D. E., Henson, R. N. A., Kiebel, S., Phillips, C., & Ashburner, J. (2002). Classical and bayesian inference in neuroimaging: Applications. *Neuroimage, 16,* 484–512.

Henson, R. N. A. (2003). Neuroimaging studies of priming. *Progress in Neurobiology, 70,* 53–81.

Henson, R. N. A., Cansino, S., Herron, J. E., Robb, W. G. K., & Rugg, M. D. (2003). A familiarity signal in human anterior medial temporal cortex? *Hippocampus, 13,* 301–304.

Henson, R. N. A., & Penny, W. (2003). *ANOVAs and SPM: Technical report.* London: Wellcome Department of Imaging Neuroscience, University College London.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551.

Huijbers, W., Pennartz, C. M. A., & Daselaar, S. M. (2010). Dissociating the "retrieval success" regions of the brain: Effects of retrieval delay. *Neuropsychologia, 48,* 491–497.

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

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Johnson, J. D., Muftuler, L. T., & Rugg, M. D. (2008). Multiple repetitions reveal functionally and anatomically distinct patterns of hippocampal activity during continuous recognition memory. *Hippocampus, 18,* 975–980.

Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z., & Buckner, R. L. (2008). Distinct cortical anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *Journal of Neurophysiology, 100,* 129–139.

Kobayashi, Y., & Amaral, D. G. (2007). Macaque monkey retrosplenial cortex: III. Cortical efferents. *Journal of Comparative Neurology, 502,* 810–833.

Nyberg, L. (2005). Any novelty in hippocampal formation and memory? *Current Opinion in Neurology, 18,* 424–428.

Rombouts, S. A., Barkhof, F., Witter, M. P., Machielsen, W. C., & Scheltens, P. (2001). Anterior medial temporal lobe activation during attempted retrieval of encoded visuospatial scenes: An event-related fMRI study. *Neuroimage, 14,* 67–76.

Rorden, C., Karnath, H. O., & Bonilha, L. (2007). Improving lesion-symptom mapping. *Journal of Cognitive Neuroscience, 19,* 1081–1088.

Rugg, M. D., & Henson, R. N. A. (2002). Episodic memory retrieval: An (event-related) functional neuroimaging perspective. In A. E. Parker, E. L. Wilding, & T. Bussey (Eds.), *The cognitive neuroscience of memory encoding and retrieval* (pp. 3–37). Hove, UK: Psychology Press.

Rugg, M. D., Henson, R. N. A., & Robb, W. G. (2003). Neural correlates of retrieval processing in the prefrontal cortex during recognition and exclusion tasks. *Neuropsychologia, 41,* 40–52.

Shannon, B. J., & Buckner, R. L. (2004). Functional-anatomic correlates of memory retrieval that suggest nontraditional processing roles for multiple distinct regions within posterior parietal cortex. *Journal of Neuroscience, 24,* 10084–10092.

Shiffrin, R. M. (2003). Modeling memory and perception. *Cognitive Science, 27,* 341–378.

Skinner, E. I., & Fernandez, M. A. (2007). Neural correlates of recollection and familiarity: A review of neuroimaging and patient data. *Neuropsychologia, 45,* 2163–2179.

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: A new perspective. *Nature Reviews Neuroscience, 8,* 872–883.

Stark, C. E. L., & Okado, Y. (2003). Making memories without trying: Medial temporal lobe activity associated with incidental memory formation during recognition. *Journal of Neuroscience, 23,* 6748–6753.

Suzuki, M., Johnson, J. D., & Rugg, M. D. (in press). Recollection-related hippocampal activity during continuous recognition: A high-resolution fMRI study. *Hippocampus.*

Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience, 10,* 792–802.

Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia, 46,* 1787–1799.

Vilberg, K. L., & Rugg, M. D. (2009). 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, 45,* 562–571.

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences, 9,* 445–453.

Wais, P. E. (2008). fMRI signals associated with memory strength in the medial temporal lobes: A meta-analysis. *Neuropsychologia, 46,* 3185–3196.

Wais, P. E., Squire, L. R., & Wixted, J. T. (2010). In search of recollection and familiarity signals in the hippocampus. *Journal of Cognitive Neuroscience, 22,* 109–123.

Ward, B. D. (2000). *Simultaneous inference for fMRI data.* Milwaukee: Biophysics Research Institute, Medical College of Wisconsin.

Weis, S., Specht, K., Klaver, P., Tendolkar, I., Willmes, K., Ruhlmann, J., et al. (2004). Process dissociation between contextual retrieval and item recognition. *NeuroReport, 15,* 2729–2733.

Wheeler, M. E., & Buckner, R. L. (2004). Functional-anatomic correlates of remembering and knowing. *Neuroimage, 21,* 1337–1349.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114,* 152–176.

Woodruff, C. C., Johnson, J. D., Uncapher, M. R., & Rugg, M. D. (2005). Content-specificity of the neural correlates of recollection. *Neuropsychologia, 43,* 1022–1032.

Yonelinas, A. P., & Jacoby, L. L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. *Consciousness and Cognition, 5,* 131–141.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., & Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience, 25,* 3002–3008.