Behavioral/Systems/Cognitive

Medial Temporal Lobe Function and Recognition Memory: A Novel Approach to Separating the Contribution of Recollection and Familiarity

Zhuang Song,[1] **Annette Jeneson,**[2] **and Larry R. Squire**[1,2,3,4]

Departments of [1]Psychiatry, [2]Psychology, and [3]Neurosciences, University of California, San Diego, California 92093, and [4]Veterans Affairs San Diego Healthcare System, San Diego, California 92161

Human neuroimaging studies of recognition memory have often been interpreted to mean that the hippocampus supports recollection but not familiarity. This interpretation is complicated by the fact that recollection-based decisions are typically associated with stronger memories than familiarity-based decisions. Some studies of source memory controlled for this difference in memory strength and found that hippocampal activity during learning predicted subsequent item memory strength while recollection-based memory (performance on source memory questions) was held at chance. This result suggests that the hippocampus is important for familiarity. However, a difficulty with this approach is that when source memory is assessed by asking specific, task-relevant source memory questions, participants who fail to answer the prescribed questions might nevertheless have available other (task-irrelevant) source information. Accordingly, successful item memory could still be associated with recollection. The present study used a novel method to assess item memory and source memory. Instead of responding to specific source questions, participants rated their source memory strength based on any information about the learning episode that was available to them. When subsequent source memory strength was held constant at the lowest possible level, we identified regions bilaterally in hippocampus, as well as in perirhinal cortex, where activity during learning increased as subsequent item memory increased in strength. In addition, activity in cortical regions (including prefrontal cortex) was related to source memory success independently of item memory strength. These findings suggest that activity in the hippocampus is related to the encoding of familiarity-based item memory, independent of subsequent recollection-based success.

Introduction

Recognition memory refers to the ability to judge that an item has been encountered previously and depends on the integrity of the medial temporal lobe (the hippocampus, the dentate gyrus, the subicular complex, and the adjacent perirhinal, entorhinal, and parahippocampal cortices) (Squire et al., 2007). This ability is thought to consist of two components: recollection and familiarity (Atkinson and Juola, 1974; Mandler, 1980; Jacoby, 1991). Recollection involves remembering specific contextual details about a previous learning episode, and familiarity involves simply knowing that an item was encountered previously.

Neuroimaging studies have often been taken to mean that medial temporal lobe structures might differentially and uniquely support these two psychological constructs. For example, in studies of source memory, correctly identifying both the study item as well as some information about the learning episode (e.g., that you made a pleasant/unpleasant judgment when you saw the study item) is thought to reflect recollection, whereas identifying only the item information is thought to reflect familiarity. A typical finding is that hippocampal activity is higher for source correct (recollection-based) decisions than for source incorrect (familiarity-based) decisions. In addition, perirhinal activity has been detected for familiarity-based decisions (for review, see Eichenbaum et al., 2007; Diana et al., 2007; Ranganath, 2010). These findings have given support to the idea that the hippocampus selectively supports recollection, whereas the perirhinal cortex supports familiarity (Brown and Aggleton, 2001).

A difficulty with this conclusion is that recollection-based decisions are typically associated with strong memories, whereas familiarity-based decisions are typically associated with weaker memories (Wixted and Squire, 2011). In the case of source memory studies, strong item memory is associated with more accurate source memory than weak item memory is (Slotnick and Dodson, 2005; Gold et al., 2006; Mickes et al., 2009). Thus, if the intention is to isolate recollection and familiarity, it is important to control for memory strength (Smith et al., 2011). Neuroimaging studies of source memory that have controlled for the memory strength of source correct and source incorrect decisions have found that the hippocampus is important for familiarity as well as for recollection (Kirwan et al., 2008; Wais et al., 2010).
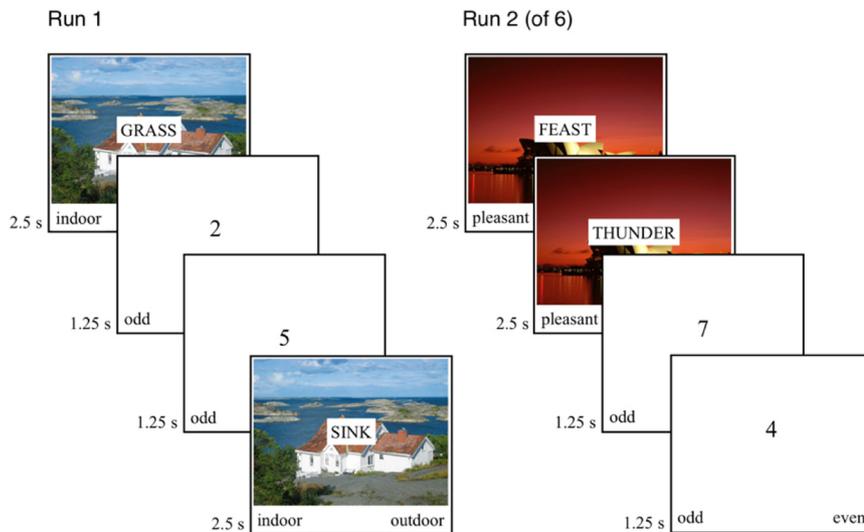
Run 1

Run 2 (of 6)

Figure 1. During scanning, participants saw 360 words in six runs (60 words per run, 2.5 s per word). Each word was presented against a background picture that was unique to each run. In half of the runs, participants decided whether the item was pleasant or unpleasant (pleasant/unpleasant judgment). In the other half of the runs, participants decided whether the item would typically be encountered indoors or outdoors (indoor/outdoor judgment). The words were intermixed with baseline trials (1.25 s per digit) where participants indicated whether a digit was odd or even.

Yet there is a difficulty with this approach as well. Whenever source memory is assessed by asking specific, task-relevant questions, participants who fail to answer the prescribed source memory questions might nevertheless have available other, task-irrelevant source information about the study episode (Yonelinas and Jacoby, 1996; Kafkas and Migo, 2009). Accordingly, item memory could be associated with recollection even when the intent is to assess brain activity related to familiarity alone.

In the present study, we used a novel method to assess brain activity related to item memory and source memory that avoids this difficulty. Instead of responding to a specific source question, participants rated their source memory strength based on any information about the learning episode available to them. We performed two analyses. First, we examined activity in the medial temporal lobe as a function of subsequent, familiarity-based item memory strength while recollection-based source memory strength was held constant at the lowest possible level. Second, we examined activity during learning related to the subsequent strength of source memory while item memory was held constant at a high level.

Materials and Methods

Participants. Sixteen right-handed volunteers recruited from the University community gave written informed consent before participation. One person was excluded due to marked head motion during scanning, so the data analysis was based on 15 participants (mean age, 27 years; range, 20–38 years; seven male).

Materials. The stimuli were 720 nouns with a mean frequency of 27 (range, 1–198) and concreteness ratings >500 (mean, 573) obtained from the Medical Research Council Psycholinguistic Database (Wilson, 1988). Half of the words were assigned to six 60-word study lists, and the other half of the words provided foils for the retrieval test. The assignment of words to the study and test conditions was counterbalanced across participants.

Procedure. Participants were scanned in six separate runs (~2 min delay between runs), during which the 360 target words were presented (2.5 s per word). To create a rich study context, each word was presented together with a picture of a natural scene. Participants were told that the picture "was just background" and that their memory for the words would later be tested. A different picture was used for each 60-word run

(Fig. 1). Also, in half of the runs, participants decided whether the word was pleasant or unpleasant (pleasant/unpleasant judgment). In the other half of the runs, they decided whether the word would typically be encountered indoors or outdoors (indoor/outdoor judgment). These two encoding questions alternated from run to run. The order of the six background pictures and the order of the two encoding questions were mixed across participants. Participants were encouraged to respond during the 2.5 s presentation time, and responses were collected via an MR-compatible button box (Current Designs). An odd/even digit task (Stark and Squire, 2001) was intermixed with word presentation and served as a baseline against which the hemodynamic response was estimated. For the digit task, participants saw a digit (1–8) for 1.25 s and indicated by button press whether the digit was odd or even. Digit task trials (80 trials per run) were pseudorandomly intermixed with the encoding trials with two constraints: each run began and ended with at least 10 digit trials (mean = 10.9) and the number of digit trials placed after each word was an even number from 0 to 10 so as to fit within the 2.5 s repetition time (TR; see fMRI imaging, below). Word presentation was most often followed by 0 digit trials (66.5% of the time), and was followed by 2, 4, 6, 8, 10 digit trials 21.9%, 6.6%, 3.3%, 1.3%, and 0.4% of the time, respectively. The mean interval between words was 1.25 s (range, 0–12.5 s). A short practice block was given before scanning to ensure that participants understood the task.

After scanning (~15 min delay), participants completed a recognition memory test. They saw all 360 words from the scan session (targets) and 360 new words (foils) one at a time in random order. For each word, participants made a recognition confidence judgment on a scale from 1 to 6 (1, definitely new; 2, probably new; 3, guess new; 4, guess old; 5, probably old; 6, definitely old). For each word rated as old, they were also asked to rate how much they remembered about their encounter with the word on the list (source memory) using a six-point scale (1, remember nothing; 6, remember almost every detail). There was no time limit for responses. Participants were encouraged to use the entire range of both the item memory scale and the source memory scale. For source memory ratings, participants were instructed to base their rating on all the detail they could remember from the time that they saw the word (for example, whether they were asked to make a pleasant/unpleasant or an indoor/outdoor judgment about the word, which decision they made, whether the word reminded them of something or someone, whether they saw the word early or late in the experiment, which picture was in the background, and whether the word made them think of other words on the list).

The confidence ratings for the item and source questions were used to back sort each of the study trials according to subsequent item memory strength and source memory strength. For the item memory question, responses to study words with a rating of 1, 2, or 3 represented misses of high, medium, and low confidence, respectively; responses with a rating of 4, 5, or 6 to study words represented hits of low, medium, and high confidence, respectively. For the source memory question, ratings from 1 to 6 were combined into three categories to obtain a sufficient number of trials at each level. Responses with a rating of 5 or 6 were designated as high-strength source memories; responses with a rating of 3 or 4 were designated as medium-strength source memories, and responses with a rating of 1 or 2 were designated as low-strength source memories.

Before testing, participants completed a practice block to ensure that they understood the instructions and the rating scales. During practice, participants were asked after each source memory rating to describe what they remembered from the time that they saw the word on the list. This

allowed the experimenter to identify any potential misunderstandings and to provide feedback on how to use the source memory scale.

After the postscan recognition memory test was completed, participants received a surprise verification test for source memory. All the target words that had been correctly endorsed in the recognition memory test (item ratings of 4, 5, or 6) were presented in random order, and participants were asked (1) to indicate whether they had made a pleasant/unpleasant or an indoor/outdoor judgment about that word during study, and (2) which one of two images had been presented as background for the word during study (i.e., they made a forced-choice decision between the true background image and one of the other five images from the other runs). Memory for the background picture on the surprise verification test of memory was no better than chance: 49.5 ± 1.4%, 51.3 ± 1.5%, and 52.4 ± 4.2% for low, medium, and high source memory ratings, respectively, and these data were not considered further.

fMRI imaging. Imaging was performed on a 3 T GE Healthcare scanner at the Center for Functional MRI (University of California, San Diego). Functional images were acquired using a gradient-echo, echo-planar, T2*-weighted pulse sequence [TR, 2500 ms; 100 TRs/run; echo time (TE), 30 ms; flip angle, 90°; matrix size, 64 × 64; field of view, 22 cm]. The first five TRs were discarded to allow for T1 equilibration. Forty oblique coronal slices (slice thickness = 5 mm, 0 gap) were acquired perpendicular to the long axis of the hippocampus and covering the whole brain. After the six functional runs, a high-resolution structural image was acquired using a T1-weighted, fast spoiled gradient-echo pulse sequence (TE, 3.0 ms; flip angle, 8°; matrix size, 256 × 192; field of view, 25 cm; 172 slices; 1 mm slice thickness).

fMRI data analysis. fMRI data were analyzed using the AFNI suite of programs (Cox, 1996). Functional data were corrected for field inhomogeneity with field mapping data collected before functional scanning, coregistered in three dimensions with the whole-brain anatomical data, slice-time corrected, and coregistered through time to reduce effects of head motion. Large motion events, defined as TRs in which there was >0.3° of rotation or >0.6 mm of translation in any direction, were excluded from the deconvolution analysis by censoring the excluded time points, but without affecting the temporal structure of the data. We also excluded the TRs immediately preceding and following the motion-contaminated TRs.

Behavioral vectors were created that coded each study trial according to the outcome of the subsequent item and source confidence ratings. Confidence ratings <4 (misses) were coded with a single behavioral vector. Item confidence ratings of 4, 5, and 6 (hits) had three vectors each, one for each category of source memory responses (low, medium, or high confidence). Trials in which there was no response to the word (mean 0.8 trials per participant) were modeled but then excluded from further analysis. The behavioral vectors and six vectors that coded for motion (three for translation and three for rotation) were used in deconvolution analyses of the fMRI time series data. This method does not assume a shape of the hemodynamic response, and the fit of the data to the model was estimated for each time point independently. The resultant fit coefficients (β-coefficients) represent activity versus baseline in each voxel for a given time point and for each of the trial types. This activity was summed over the expected hemodynamic response (2.5–15 s after trial onset) and taken as the estimate of the response to each trial type (relative to the digit task baseline).

Initial spatial normalization was accomplished using each participant's structural MRI scan to transform the data to the atlas of Talairach and Tournoux (1988). Statistical maps were also transformed to Talairach space, resampled to 2 mm isotropic, and smoothed using a Gaussian filter (4 mm full-width at half maximum) that respected the anatomical boundaries of the regions of the medial temporal lobe (MTL) defined for each individual participant. Specifically, the smoothing was performed within each of the anatomically defined MTL regions, but, to prevent activity in one region (e.g., parahippocampal cortex) from being blurred into another, adjacent region (e.g., hippocampus), smoothing was not extended beyond the edges of the MTL regions. This was accomplished by creating a separate mask for each region, smoothing the data within that mask, and then recombining the smoothed data. The Talairach-transformed data were used in the whole-brain analyses.
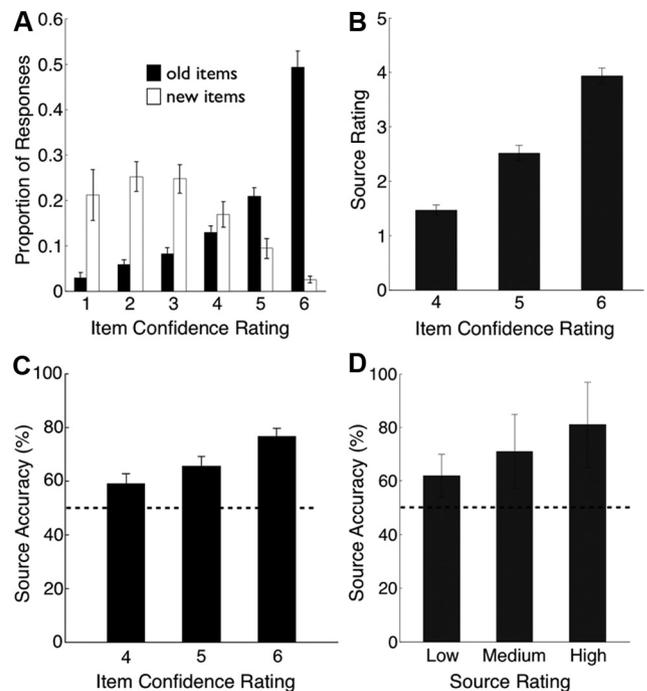

Figure 2. *A–D*, Behavioral performance on the postscan recognition memory test (*A*, *B*) and the surprise source memory verification test (*C*, *D*). *A*, Mean proportion of studied (old) and unstudied (new) items endorsed at each confidence level for the item (old/new) memory decision (1, definitely new; 6, definitely old). *B*, For words declared old, participants were asked to rate how much they remembered from the time that they saw the word on the list (1, remember nothing; 6, remember almost every detail). Source memory ratings increased with increasing item confidence ratings. After the recognition memory test, participants received a surprise verification test for source memory. For words declared old, participants decided which encoding question had been associated with each item. *C*, Source memory accuracy (i.e., memory for the encoding question) increased with increasing item memory confidence. *D*, Source memory accuracy also increased with increasing source memory ratings from 1 to 6 (low, ratings of 1 or 2; medium, ratings of 3 or 4; high, ratings of 5 or 6). Dotted lines indicate chance performance. Error bars indicate SEM.

A region of interest (ROI) alignment technique was used to improve cross-participant alignment of the MTL and thus increase statistical power for group analyses. The alignment technique used was conceptually similar to the ROI large deformation diffeomorphic mapping technique (Miller et al., 2005; Kirwan et al., 2007) and was performed using the Diffeomorphic Demons tool for MedINRIA software (ROI-Demons) (Yassa and Stark, 2009). The first step in this approach is to define anatomical regions of interest for each subject. Anatomical regions of interest were manually segmented in 3D on the Talairach-transformed anatomical images for the hippocampus; temporal polar; and entorhinal, perirhinal, and parahippocampal cortices. Temporal polar, entorhinal, and perirhinal cortices were defined according to the landmarks described by Insausti et al. (1998a). The caudal border of the perirhinal cortex was defined as 4 mm caudal to the posterior limit of the gyrus intralimbicus as identified on coronal sections (Insausti et al., 1998a). The parahippocampal cortex was defined bilaterally as the portion of the parahippocampal gyrus caudal to the perirhinal cortex and rostral to the splenium of the corpus callosum (Insausti et al., 1998b). Using ROI-Demons, the anatomically defined ROIs for each individual participant were then used to normalize each subject's set of ROIs to a previously defined template for each structure (Kirwan et al., 2007). Diffeomorphic mapping techniques such as ROI-Demons have an advantage over other flat-mapping techniques: the spatial transformation of structures maintains the relationships between voxels. This transformation was then applied to the statistical maps; all MTL analyses were performed on the ROI-Demons-transformed data.

Following individual deconvolution analysis, parameter estimate maps from each participant were entered into group-level analyses. The
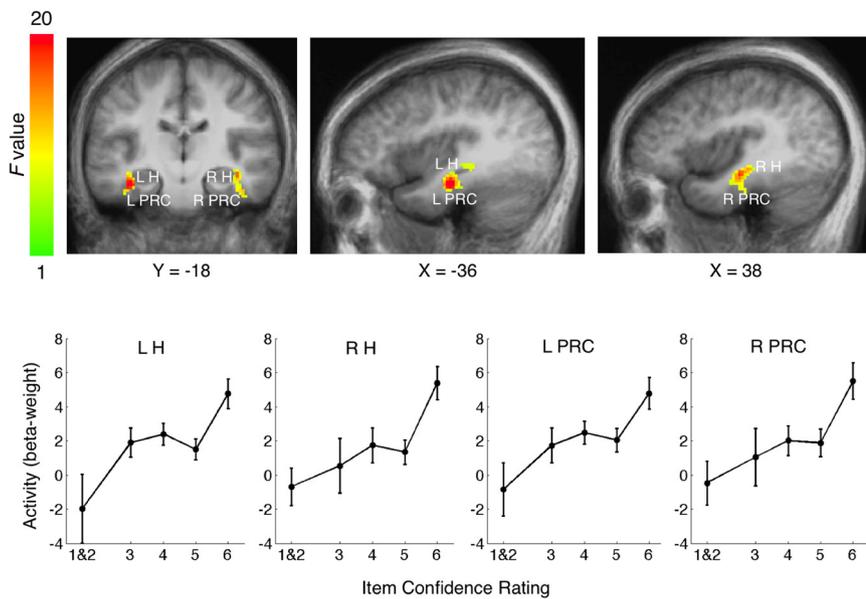
Figure 3. During learning, activity bilaterally in the hippocampus (L H and R H) and perirhinal cortex (L PRC and R PRC) increased as a function of subsequent item memory strength (confidence rating 1–6). Source memory strength was held constant at a low level by limiting the analysis to items missed on the old/new test (item confidence rating 1–3) and items recognized and given a low source memory rating (1 or 2). Error bars indicate ±SEM.

statistical map was thresholded at a voxelwise *p* value of <0.01. For the MTL analyses, group statistical maps were masked using the MTL template from the ROI-Demons alignment procedure to include only regions of the MTL. A cluster-correction technique was used to correct for multiple comparisons, and Monte Carlo simulations were used to determine how large a cluster of voxels was needed to be statistically meaningful ($p < 0.05$) within the volume of the MTL (minimum cluster extent of 17 contiguous voxels, 136 μl) (Forman et al., 1995; Xiong et al., 1995). For whole-brain analyses, the minimum cluster extent was 45 voxels (360 μl).

Results

Behavioral performance
Figure 2*A* shows the mean proportion of responses for targets (studied words) and foils (new words) on the item recognition portion of the postscan old/new recognition test. Overall, participants scored 77.2 ± 2.5% (mean ± SEM) correct on the old/new recognition test (hit rate, 83.2 ± 2.1%; false alarm rate, 28.8 ± 4.8%; $d' = 1.70 ± 0.18$). Source memory ratings (i.e., ratings of how much the participant reported to remember from the time that the word was presented) increased correspondingly with item confidence ratings (Fig. 2*B*). Source accuracy on the surprise verification test also increased in correspondence with item memory confidence (58.9 ± 3.7%, 65.5 ± 3.5%, and 76.5 ± 2.9% for item confidence ratings of 4, 5, and 6, respectively; Fig. 2*C*). Note, however, that the source accuracy of those items given a low source rating (Fig. 2*D*, leftmost bar) did not increase as item confidence increased (58.0 ± 3.2%, 62.1 ± 3.3%, and 60.7 ± 4.4% source accuracy for item confidence ratings of 4, 5, and 6, respectively, a finding not illustrated in the Fig. 2). Last, accuracy on the surprise verification test that assessed memory for the encoding question increased correspondingly with source memory ratings (Fig. 2*D*): 61.8 ± 2.0%, 70.9 ± 3.7%, and 80.6 ± 4.2% for low (ratings 1–2), medium (ratings 3–4), and high (ratings 5–6) source memory ratings, respectively.

fMRI analyses
We first examined the relationship between fMRI activity during learning and the strength of subsequent item memory when the

strength of subsequent source memory was held constant at the lowest possible level. We then examined the relationship between fMRI activity during learning and the strength of subsequent source memory when the strength of subsequent item memory was held constant at a high level.

For the first analysis, source memory was held constant at the lowest possible level by limiting the analysis to missed items (item confidence ratings 1–3) and items that were recognized but associated with low source memory rating (ratings 1 or 2). Items with source ratings of 1 or 2 were combined to obtain a sufficient number of trials at each level of item confidence ratings. A linear trend analysis (item confidence ratings 1–6) restricted to the medial temporal lobe (Kirwan et al., 2008) identified clusters bilaterally in both hippocampus and perirhinal cortex (left hippocampus and perirhinal cortex: $x = -41, y = -21, z = -8$; $F_{(1,14)} = 9.87$; 1040 μl; right hippocampus and perirhinal cortex: $x = 37, y = -17, z = -14$; $F_{(1,14)} = 13.53$; 816 μl). Activity at study in these regions increased as a function of the strength of subsequent item memory (Fig. 3). A repeated-measures ANOVA (again, for item confidence ratings 1–6 when source memory was low) identified clusters in nearly the same locations as illustrated in Figure 3, except that activity in right perirhinal cortex was not detected (left hippocampus and perirhinal cortex: $x = -37, y = -17, z = -16$; $F_{(4,70)} = 6.70$; 736 μl; right hippocampus: $x = 37, y = -19, z = -8$; $F_{(4,70)} = 6.97$; 456 μl). Both the linear trend analysis and the ANOVA also identified clusters in bilateral temporopolar cortex (linear trend: left side: $x = -35, y = 27, z = -20$; $F_{(1,14)} = 18.95$; 1216 μl; right side: $x = 49, y = 19, z = -18$; $F_{(1,14)} = 19.34$; 1800 μl), where activity decreased as a function of subsequent item memory strength. We also conducted the same linear trend analysis for the whole brain, and the results are presented in Table 1.

Although low source memory ratings were the lowest possible ratings, source memory was not completely absent when that rating was used. Thus, on the surprise verification test, memory for the encoding question was measurably above chance even when the source memory rating was low (61.8 ± 4.4% correct, $p < 0.05$; Fig. 2*D*). However, as mentioned earlier, it is important to note that, even though source memory on the verification test was above chance when source memory ratings were low, source memory remained constant (at ~60%) as item memory strength increased from 4 to 5 to 6. Thus, activity associated with increasing item memory strength was presumably due to increasing item familiarity, not increasing recollection.

To reduce still further the contribution of source memory to recognition memory judgments, we removed a subset of those items that were associated with accurate performance on the verification test. This procedure brought performance on the source verification test to chance levels for those items that were recognized and also associated with a low source memory rating (removal of 25.9 ± 5.1 trials per subject). The resulting source accuracy scores across all items that were recognized and associated with a low source memory rating were 52.6 ± 3.0%, 49.8 ± 1.8%, and 52.7 ± 3.4% correct for items given old/new confidence ratings of 4, 5,

Table 1. Linear trend analysis for the whole-brain data (as in Fig. 3)

Direction of trend	Region	Talairach coordinates			F value (peak)	Cluster size (μl)
		x	y	z		
Positive	L superior occipital gyrus	−33	−85	24	15.84	1336
	B cerebellum	−7	−55	−52	15.99	1088
		9	−85	−36	14.41	576
		−13	−85	−36	17.98	440
	B middle occipital gyrus	−51	−79	6	13.46	680
		41	−89	8	12.82	392
	L cuneus	−17	−91	42	14.56	432
Negative	L basal ganglia/insula	−31	−9	6	21.52	16584
	B precuneus	13	−57	40	23.74	15760
	B superior/medial frontal gyrus	−17	21	44	18.53	12336
	R superior/middle temporal gyrus	45	9	−18	22.16	7576
	R supramarginal gyrus	35	−47	40	17.88	6208
	R thalamus	25	−15	6	18.37	3304
	B cerebellum	35	−45	−50	19.59	2040
		−31	−47	−36	17.40	1784
		−37	−39	−30	16.91	672
	R middle frontal gyrus	23	11	40	15.47	1640
	L basal ganglia	−23	17	10	12.97	1136
	R precentral gyrus	61	11	10	17.03	800
	L postcentral gyrus	−59	−29	22	11.78	728
		−23	−31	46	14.68	720
	L insula	−45	−19	18	13.58	648
	R superior temporal gyrus	59	−47	14	11.96	648
	L superior frontal gyrus	−19	55	18	13.61	504
	R putamen	23	−9	18	16.54	456
	R superior frontal gyrus	9	13	66	13.95	368

Areas where activity during learning varied (positively or negatively) as a function of the subsequent strength of item memory (item confidence rating 1– 6). Source memory strength was held constant at a low level by limiting the analysis to missed items (confidence rating 1–3) and items recognized with low-strength source memory (source memory rating 1 or 2). Talairach coordinates indicate the location of the voxel that had the peak F value from the linear trend analysis. L, Left; R, right; B, bilateral.

Table 2. Linear trend analysis for the whole-brain data (as in Fig. 4)

Direction of trend	Region	Talairach coordinates			F value (peak)	Cluster size (μl)
		x	y	z		
Positive	L angular gyrus	−49	−59	34	29.91	9952
	B medial/superior frontal gyrus	−13	49	34	22.58	3704
		−7	37	54	19.20	2600
	L lentiform nucleus	−19	−5	2	17.82	1960
	R caudate	13	9	6	11.23	568
	L anterior cingulate gyrus	−21	17	30	14.16	568
	L posterior cingulate gyrus	−5	−47	44	17.85	448
	R precentral gyrus	21	−21	62	17.88	416
	R cerebelum	41	−47	−50	18.65	400
	L middle frontal gyrus	−37	15	38	16.63	392
Negative	L inferior parietal	−53	−33	38	19.01	880
	L lingual gyrus	−23	−61	−2	18.74	704
	R insula	41	−25	−2	19.44	584
	L middle temporal gyrus	−53	−63	0	10.07	456

Areas where activity during learning varied (positively or negatively) as a function of the subsequent strength of source memory (low, medium, and high). Item memory strength was held constant at the highest level (confidence rating 6). Talairach coordinates indicate the location of the voxel that had the peak F value from the linear trend analysis. L, Left; R, right; B, bilateral.
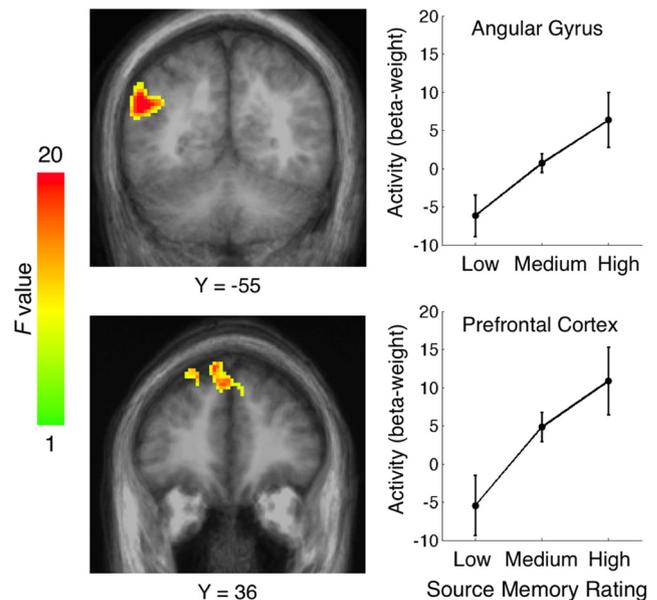

Figure 4. During learning, activity in angular gyrus and superior/medial prefrontal cortex increased as a function of subsequent source memory ratings (low, medium, high). Item memory strength was held constant by limiting the analysis to studied items that were subsequently endorsed correctly with high confidence (item confidence rating 6). Error bars indicate ±SEM. See Table 2 for other areas of the whole brain where activity varied in relation to source memory strength.

and 6, respectively. With these items, we then performed a linear trend analysis as described above and identified clusters bilaterally in hippocampus and perirhinal cortex (Fig. 3). The results of a repeated-measures ANOVA were also similar to the results illustrated in Figure 3, except that the right perirhinal cortex was not identified.

For the second analysis, we examined activity during learning as a function of subsequent source memory strength. For this analysis, the strength of item memory was held constant at the highest level (confidence rating of 6). A linear trend analysis (low, medium, and high source memory ratings) restricted to the medial temporal lobe did not identify regions with significant activation. Indeed, no single voxel was detected in the hippocampus for this analysis. A whole-brain analysis identified a number of regions (Table 2). The largest activations occurred in angular gyrus, medial prefrontal cortex, and superior prefrontal cortex, where activity increased in relation to subsequent source memory strength (Fig. 4). Similarly, a repeated-measures ANOVA restricted to the medial temporal lobe did not identify regions with significant activation. A whole-brain analysis identified a number of regions in prefrontal cortex and temporal cortex, similar to the findings from the linear trend analysis (Table 2).

Discussion

Participants made pleasant/unpleasant or indoor/outdoor judgments for 360 words in the scanner. After scanning, they made old/new item confidence judgments (1–6) for these 360 words and 360 new words. For items rated old, participants also rated (1–6) how much they remembered about their encounter with the word on the list. Finally, after the postscan recognition test, they completed a surprise source memory verification test for items correctly declared old. This test assessed memory for the encoding question (pleasant/unpleasant or indoor/outdoor judgment) that was asked during the item presentation at study. We performed two different analyses of the fMRI data, using both the item confidence ratings and the source memory ratings to assess encoding-related activity, first in relation to subsequent item memory strength and then in relation to subsequent source memory strength.

Item memory strength

The objective was to assess activity during learning related to subsequent item memory strength independently of subsequent source memory strength. We identified activity at study in both hippocampus and perirhinal cortex bilaterally that increased as a

function of subsequent item memory strength when subsequent source memory strength was held constant at the lowest possible level (Fig. 3). Although low source memory ratings were as low as possible (i.e., ratings of 1 and 2), source memory was not completely absent under these conditions. Thus, on the surprise verification test, memory for the encoding question was measurably above chance even when the source memory rating was low (61.8 ± 4.4% correct, $p < 0.05$; Fig. 2D). Importantly, however, the logic of our analysis is not dependent on source memory being wholly absent across large variations in the strength of item memory. Instead it is dependent on source memory being low and constant across large variations in the strength of item memory (because, in that case, the variations in item memory strength are not attributable to variations in source memory). Critically, for item confidence ratings of 4, 5, and 6, source accuracy for low-confidence source decisions was 58.0%, 62.1%, and 60.7%, respectively. The results in Figure 3 thus suggest that subsequent source memory success is not necessary to evoke hippocampal activity during learning. Instead, even when the contribution of source memory was at the lowest possible level, hippocampal activity during learning increased as subsequent item memory increased in strength.

In an earlier similar study, Kirwan et al. (2008) also identified hippocampal activity at study that was related to item memory at encoding and was independent of source memory. However, a concern in Kirwan et al. (2008), as well as in other previous studies of source memory, is that the estimate of source memory strength was based on how well participants could answer a specific, task-relevant source memory question. Yet, when participants failed to answer any specific source question, they might nevertheless have had available other, task-irrelevant source memory about their encounter with the item (Yonelinas and Jacoby, 1996; Kafkas and Migo, 2009). Accordingly, even though the intent might be to assess brain activity in relation to item memory alone (in the absence of successful recollection), it is possible that item memory could still be associated with recollection. The present study introduced a novel procedure to address this problem. Participants were instructed to rate their source memory based on any kind of source information available to them about the item they had studied. The idea is that a low overall source rating provides a more reliable way of identifying weak source memory than an incorrect answer to one of many possible source questions.

Note that the generic source memory rating provided by participants was not simply a proxy for the score they obtained on the surprise source memory test about the encoding question. Participants were specifically instructed to base their source judgment on any information they could recollect about a study item [much in the same way that the Remember instruction is used in the Remember-Know procedure (Rajaram, 1996)]. Furthermore, the test about the encoding question came only after all the source ratings had been collected and served simply to verify that the source ratings provided by participants had validity and related measurably to objective features of the study episode.

Our finding that activity in the hippocampus increased as a function of subsequent item memory strength even when subsequent source memory strength was held at the lowest possible level suggests that the hippocampus is important for the encoding of item familiarity. Nonetheless, other studies have not observed familiarity-based activity in the hippocampus, and such findings have often been interpreted to mean that the hippocampus is important for supporting recollection rather than familiarity (for review, see Brown and Aggleton, 2001; Eichenbaum et al., 2007). Yet

many of these studies have involved comparisons between strong, recollection-based memory and weaker, familiarity-based memory (Squire et al., 2007; Wixted and Squire, 2011), and recent work suggests that fMRI activity in the hippocampus is best detectable when memory is strong (Song et al., 2011a). Furthermore, when strong, recollection-based memories are compared with strong, familiarity-based memories, the evidence suggests that the hippocampus supports both recollection and familiarity (Kirwan et al., 2010; Wais et al., 2010; Smith et al., 2011; Song et al., 2011b).

Source memory strength

We assessed activity during learning as a function of subsequent source memory strength when item memory strength was held constant at the highest level (rating of 6). It was not possible to examine activity related to differences in source memory strength while holding item memory constant at lower levels (ratings of 4 and 5) because there was an insufficient number of trials associated with these lower levels of item memory strength at each of the three levels of source memory ratings (low, medium, and high). Holding subsequent item memory strength constant reduces the influence of activity related to differences in item memory strength and makes it possible to consider the influence of source memory itself.

Interestingly, this analysis did not identify regions in the medial temporal lobe. In contrast, a whole-brain analysis did identify a number of regions, including angular gyrus, superior prefrontal cortex, and medial prefrontal cortex, where activity increased as a function of subsequent source memory strength when item memory strength was held constant at the highest level (Table 2, Fig. 4).

The results from the whole-brain analyses are consistent with previous studies that have related activity in angular gyrus and adjacent temporoparietal junction to source memory retrieval (Yonelinas et al., 2005; Daselaar et al., 2006; Uncapher et al., 2010) and activity in medial prefrontal cortex to source memory encoding (Cansino et al., 2002; Kirwan et al., 2008). Activity related to source memory is often assessed by comparing activity associated with correct source memory decisions to activity associated with incorrect source memory decisions. With this contrast, several studies have reported greater activity in the hippocampus (but not perirhinal or entorhinal cortex) for correct versus incorrect source memory decisions (Davachi et al., 2003; Ranganath et al., 2004; Kensinger and Schacter, 2006; Uncapher et al., 2006; Park et al., 2008). This finding has often been interpreted to mean that the hippocampus, but not perirhinal cortex, is important for recollection-based source memory. However, memory for items associated with correct source memory decisions is typically stronger than memory for items that are associated with incorrect source memory decisions (Fig. 2C) (Slotnick and Dodson, 2005; Gold et al., 2006; Mickes et al., 2009). Accordingly, it is useful to eliminate this difference in memory strength before assessing activity related to source memory success (Kirwan et al., 2008; Wais et al., 2010), as we have done here.

Conclusion

We conducted two independent analyses to explore activity in the medial temporal lobe during encoding. First, we assessed activity as a function of the subsequent strength of item memory while source memory strength was held constant at the lowest possible level. Second, we assessed activity as a function of the subsequent strength of source memory while item memory was held constant at the highest level. Participants rated source memory strength

based on any source information available to them. Source memory was not based on any task-specific question as it was in most previous studies. The key finding was that, when subsequent source memory strength was held constant at the lowest possible level, we identified regions bilaterally in hippocampus, as well as in perirhinal cortex, where activity during learning increased as subsequent item memory increased in strength. This finding, together with earlier results (Kirwan et al., 2008), suggests that activity in the hippocampus, as well as in the perirhinal cortex, is related to the encoding of familiarity-based item memory, regardless of subsequent recollection-based success. In addition, activity in cortical regions (including angular gyrus and prefrontal cortex) was related to source memory success independently of item memory strength.

References

Atkinson RC, Juola JF (1974) Search and decision processes in recognition memory. In: Contemporary developments in mathematical psychology (Krantz DH, Atkinson RC, Suppes P, eds), pp 243–290. San Francisco: Freeman.

Brown MW, Aggleton JP (2001) Recognition memory: what are the roles of the perirhinal cortex and hippocampus? Nat Rev Neurosci 2:51–61.

Cansino S, Maquet P, Dolan RJ, Rugg MD (2002) Brain activity underlying encoding and retrieval of source memory. Cereb Cortex 12:1048–1056.

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173.

Daselaar SM, Fleck MS, Cabeza R (2006) Triple dissociation in the medial temporal lobes: recollection, familiarity, and novelty. J Neurophysiol 96:1902–1911.

Davachi L, Mitchell JP, Wagner AD (2003) Multiple routes to memory: distinct medial temporal lobe processes build up item and source memories. Proc Natl Acad Sci U S A 100:2157–2162.

Diana RA, Yonelinas AP, Ranganath C (2007) Imaging recollection and familiarity in the medial temporal lobe: a three-component model. Trends Cogn Sci 11:379–386.

Eichenbaum H, Yonelinas AP, Ranganath C (2007) The medial temporal lobe and recognition memory. Annu Rev Neurosci 30:123–152.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med 33:636–647.

Gold JJ, Smith CN, Bayley PJ, Shrager Y, Brewer JB, Stark CE, Hopkins RO, Squire LR (2006) Item memory, source memory, and the medial temporal lobe: concordant findings from fMRI and memory-impaired patients. Proc Natl Acad Sci U S A 103:9351–9356.

Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, Laakso MP, Pitkänen A (1998a) MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. Am J Neuroradiol 19:659–671.

Insausti R, Insausti AM, Sobreviela MT, Salinas A, Martínez-Peñuela JM (1998b) Human medial temporal lobe in aging: anatomical basis of memory preservation. Microsc Res Tech 43:8–15.

Jacoby LL (1991) A process dissociation framework: separating automatic and intentional uses of memory. J Mem Lang 30:513–541.

Kafkas A, Migo EM (2009) Familiarity and recollection in the medial temporal lobe. J Neurosci 29:2309–2311.

Kensinger EA, Schacter DL (2006) Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. J Neurosci 26:2564–2570.

Kirwan CB, Jones CK, Miller MI, Stark CE (2007) High-resolution fMRI investigation of the medial temporal lobe. Hum Brain Mapp 28:959–966.

Kirwan CB, Wixted JT, Squire LR (2008) Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. J Neurosci 28:10541–10548.

Kirwan CB, Wixted JT, Squire LR (2010) A demonstration that the hippocampus supports both recollection and familiarity. Proc Natl Acad Sci U S A 107:344–348.

Mandler G (1980) Recognizing: the judgment of previous occurrence. Psychol Rev 87:252–271.

Mickes L, Wais PE, Wixted JT (2009) Recollection is a continuous process: implications for dual-process theories of recognition memory. Psychol Sci 20:509–515.

Miller MI, Beg MF, Ceritoglu C, Stark C (2005) Increasing the power of functional maps of the medial temporal lobe by using large deformation diffeomorphic metric mapping. Proc Natl Acad Sci U S A 102:9685–9690.

Park H, Uncapher MR, Rugg MD (2008) Effects of study task on the neural correlates of source encoding. Learn Mem 15:417–425.

Rajaram S (1996) Perceptual effects on remembering: recollective processes in picture recognition memory. J Exp Psychol Learn Mem Cogn 22:365–377.

Ranganath C (2010) A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. Hippocampus 20:1263–1290.

Ranganath C, Yonelinas AP, Cohen MX, Dy CJ, Tom SM, D'Esposito M (2004) Dissociable correlates of recollection and familiarity within the medial temporal lobes. Neuropsychologia 42:2–13.

Slotnick SD, Dodson CS (2005) Support for a continuous (single-process) model of recognition memory and source memory. Mem Cognit 33:151–170.

Smith CN, Wixted, JT, Squire LR (2011) The hippocampus supports both recollection and familiarity when memories are strong. J Neurosci, in press.

Song Z, Wixted JT, Smith CN, Squire LR (2011a) Different nonlinear functions in hippocampus and perirhinal cortex relating functional MRI activity to memory strength. Proc Natl Acad Sci U S A 108:5783–5788.

Song Z, Wixted JT, Hopkins RO, Squire LR (2011b) Impaired capacity for familiarity after hippocampal damage. Proc Natl Acad Sci U S A 108:9655–9660.

Squire LR, Wixted JT, Clark RE (2007) Recognition memory and the medial temporal lobe: a new perspective. Nat Rev Neurosci 8:872–883.

Stark CE, Squire LR (2001) When zero is not zero: the problem of ambiguous baseline conditions in fMRI. Proc Natl Acad Sci U S A 98:12760-12766.

Talairach J, Tournoux P (1988) A co-planar stereotaxic atlas of the human brain. New York: Thieme Medical.

Uncapher MR, Otten LJ, Rugg MD (2006) Episodic encoding is more than the sum of its parts: an fMRI investigation of multifeatural contextual encoding. Neuron 52:547–556.

Uncapher MR, Hutchinson JB, Wagner AD (2010) A roadmap to brain mapping: toward a functional map of human parietal cortex. Neuron 67:5–8.

Wais PE, Squire LR, Wixted JT (2010) In search of recollection and familiarity signals in the hippocampus. J Cogn Neurosci 22:109–123.

Wilson MD (1988) The MRC psycholinguistic database: machine readable dictionary, version 2. Behav Res Methods Instrum Comput 20:6–11.

Wixted JT, Squire LR (2011) The medial temporal lobe and the attributes of memory. Trends Cogn Sci 15:210–217.

Xiong JH, Gao JH, Lancaster JL, Fox PT (1995) Clustered pixels analysis for functional MRI activation studies of the human brain. Hum Brain Mapp 3:287–301.

Yassa MA, Stark CE (2009) A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe. Neuroimage 44:319–327.

Yonelinas AP, Jacoby LL (1996) Noncriterial recollection: familiarity as automatic, irrelevant recollection. Conscious Cogn 5:131–141.

Yonelinas AP, Otten LJ, Shaw KN, Rugg MD (2005) Separating the brain regions involved in recollection and familiarity in recognition memory. J Neurosci 25:3002–3008.