Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe

C. Brock Kirwan¹ and Craig E.L. Stark^{1,2,3}

1 Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, Maryland 21218, USA; ² Department of Neuroscience, Johns Hopkins University, Baltimore, Maryland 21218, USA

The medial temporal lobe (MTL) supports the formation and retrieval of long-term declarative memories, or memories for facts and everyday events. One challenge posed for this type of memory stems from the highly overlapping nature of common episodes. Within cognitive psychology, it is widely accepted that interference between information learned at different times is a major limitation on memory. In spite of several decades of intense research in the fields of interference theory and the neurobiological underpinnings of declarative memory, there is little direct evidence bearing on how the MTL resolves this interference to form accurate memories of everyday facts and events. Computational models of MTL function have proposed a mechanism in which the MTL, specifically the hippocampus, performs pattern separation, whereby overlapping representations are made less similar. However, there is little evidence bearing on how this process is carried out in the intact human MTL. Using high-resolution fMRI, we conducted a set of experiments that taxed behavioral pattern separation by using highly similar, interfering stimuli in a modified continuous recognition task. Regions within the parahippocampal gyrus demonstrated activity consistent with a "recall to reject" strategy. In contrast and critical to performing the task, activity within the hippocampus distinguished between correctly identified true stimulus repetitions, correctly rejected presentations of similar lure stimuli, and false alarms to similar lures. These data support the computational models' assertion that the hippocampus plays a key role in pattern separation.

Declarative memory, or memory for everyday facts and events, is critically dependent on the medial temporal lobe (MTL) structures, including the hippocampus and the surrounding structures of the parahippocampal gyrus (Squire et al. 2004). However, the contributions of the several MTL structures are not well defined. One class of computational models proposes a distinction between hippocampus and cortex according to which the cortex slowly forms mnemonic representations based on the statistical regularities in the environment (e.g., McClelland et al. 1995; O'Reilly and Rudy 2001; Norman and O'Reilly 2003). The hippocampus, on the other hand, is thought to possess key circuitry in the dentate gyrus that is designed to quickly form distinct representations and perform pattern separation to a degree not found in other structures. At the computational level, pattern separation is the process of reducing the average overlap between two representations, thus making similar representations more distinct or orthogonalized in order to afford rapid learning without inducing interference and retrieval errors (e.g., McClelland et al. 1995; Rolls and Treves 1998; O'Reilly and Rudy 2001; Norman and O'Reilly 2003). The establishment of individual representations of potentially highly overlapping input is essential to episodic memory and the process of recollection (Tulving 2002) and to the rapid learning of arbitrary associations in these models. The complement to pattern separation, pattern completion, is the process whereby a previously stored representation is retrieved given a partial or otherwise degraded cue. Computational models suggest that pattern completion and pattern separation interact dynamically and that the hippocampus may require smaller changes in the input in order to perform pattern separation than cortical regions. Thus, given the same set of stimuli, the hippocampus may be more prone to pattern separation, while the cortex may be more prone to pattern completion.

3Corresponding author. E-mail cstark@jhu.edu; fax (410) 516-4478.

Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.663507.

Given the emphasis placed on hippocampal pattern separation in these models, a key question is whether pattern separation can actually be observed in the hippocampus. Several recent electrophysiology (Lee et al. 2004; Leutgeb et al. 2004, 2005, 2007; Wills et al. 2005) and early immediate gene (Vazdarjanova and Guzowski 2004) studies offer evidence for pattern separation processes in the rodent hippocampus. By recording from place cells in the hippocampus, these studies have shown that changes in the environmental cues can lead to both very similar representations of this distorted space across cells (completion) and highly dissimilar, or orthogonalized, representations of this space (separation). Further electrophysiological evidence from single neurons in the MTL of epileptic patients (Viskontas et al. 2006) supports the predictions of computational models by showing that the human hippocampus is characterized by more sparse representations than the parahippocampal gyrus. The authors interpret these findings as consistent with the role of the hippocampus in pattern separation.

Direct evidence that human hippocampal neurons perform pattern separation is difficult to obtain using less invasive techniques. fMRI neuroimaging is noninvasive, but faces several severe challenges. For instance, Viskontas et al. (2006) point out that increases in fMRI activity observed in the hippocampus in response to repeated stimuli may reflect overall decreases in synaptic activity due to active inhibition (Stark and Squire 2001; Caesar et al. 2003). When coupled with fMRI's resolution (even that of high-resolution fMRI), it is difficult to relate changes in the patterns of activity across neurons to changes in the BOLD fMRI signal. Furthermore, while it may be difficult for any technique to cleanly isolate processes of encoding and retrieval (Martin 1999; Schacter and Wagner 1999; Stark and Okado 2003), the severely limited temporal resolution makes this especially difficult as we cannot even directly observe an encoding process that might follow a retrieval failure (a mechanism found in many of the models). Thus, even though many of the technical challenges of performing high-resolution fMRI in the MTL have been well met by recent advancements in imaging techniques, many challenges still remain.

That is not to say that the predictions of the computational models cannot be tested using fMRI, however. Here we present two experiments that test a fundamental hypothesis that the hippocampus possesses the strongest ability to perform pattern separation (e.g., via orthogonalization in the dentate gyrus). We do this using a variant of the continuous recognition task that taxes pattern separation. The behavioral consequence of pattern separation is the ability to mnemonically distinguish between two similar stimuli. In order to behaviorally tax pattern separation, one should be able to present subjects with two similar stimuli (e.g., as a target and a similar lure in a continuous recognition memory paradigm) and determine if the subject can distinguish between them by correctly rejecting the lures as being new and only similar to previously presented items. The ability to reject similar lures critically depends on pattern separation. Incorrectly calling the lures "old" (false alarms) would indicate a behavioral failure of pattern separation and correct rejections of lures would offer behavioral evidence that pattern separation has occurred at a computational level. Critically, one must allow not only "old" and "new" responses, but "similar" responses as well, so that one can assess whether, when subjects fail to respond "old," they have merely forgotten the original item or whether they can retrieve a memory of the original item and know that the present one is similar to, but not the same as the original.

This approach, while allowing us to place demands on pattern separation processes, still does not allow us to cleanly isolate computational pattern separation. Rejection of similar lures depends on adequate representation of differences between the stimuli, which is the definition of computational pattern separation (Rolls and Treves 1998). In addition to the representation of differences between stimuli, correctly rejecting similar lures also depends on the retrieval of details necessary to distinguish lures from targets. Indeed, computational models of MTL functioning (Norman and O'Reilly 2003) predict that under conditions of high overlap between targets and lures, subjects might employ a "recall to reject" strategy. Thus, when presented with a lure, subjects might first retrieve the original representation (through pattern completion processes) and compare that representation to the new one. If the representations match to a high enough degree, the subject will likely respond "old" as if to a target. If the representations do not match, the subject will likely respond "similar" or even "new" and attempt to encode this as a new event, separated from the original event. Therefore, correct rejections of similar lures imply encoding and retrieval of sufficient levels of detail to be diagnostic of previous presentation of similar stimuli. False alarms to lures imply that not enough detail was encoded or retrieved about the original target stimulus.

Thus, the computational processes of pattern separation and pattern completion will likely exist in each of these trial types, albeit in varying degrees. Processes of encoding and retrieval will also exist in each of these trial types, albeit in varying degrees. Here, we have taken the approach of using this variability to assess function. Given the above concerns, we cannot know a priori which contrast, if any, would cleanly isolate a computation such as pattern separation. We can, however, assess where in the MTL we can observe signals that differentiate any of these trial types from each other. That is, we can identify regions with activity that correlates with subjects' ability to recognize or fail to recognize the distinctions between repeated and similar events. If a region or set of regions (e.g., the various subregions of the hippocampus) demonstrates activity that differentiates these trial types, this activity could be used as the basis for the different behaviors of responding "old" to true repetitions, "similar" to lures, and "old" to lures. Thus, we can test the hypothesis that, under high demands of "behavioral" pattern separation, it is the hippocampal region, and not the adjacent structures of the parahippocampal cortex, that provides these signals.

Here, we used high-resolution fMRI in two experiments to scan subjects while they performed a variant of the continuous recognition paradigm. Subjects were shown a series of stimuli and were required to decide if the stimulus had been presented before in the experiment. Subjects were informed that some of the stimuli would be very similar, but not identical to previously presented stimuli ("Lures"). Accordingly, subjects made one of three responses to each stimulus: "new," "similar," or "old." Experiment 1 used objects as stimuli, while Experiment 2 used faces. We hypothesized that MTL responses would be modulated by interstimulus interference as well as by task demands and stimulus type and found that, while many MTL regions showed some such modulation, it was only the signals in the hippocampus that fully differentiated the critical trials from each other.

Materials and Methods

Subjects

Subjects were healthy volunteers and were recruited from the Johns Hopkins University community. Twenty-one right-handed subjects (15 female) gave written informed consent before participating in Experiment 1. The mean age was 22.9 (6.87 SD). One subject was excluded from the analysis because of excessive motion. For Experiment 2, twenty-one healthy right-handed subjects gave written informed consent (9 female). The mean age was 23.53 yr (SD = 4.52). One subject was excluded from analysis because of experimenter error. Five people participated in both Experiments 1 and 2 on different occasions.

Experiment 1

Stimuli consisted of 1109 color photographs of common namable objects. Three hundred eighty-four of the stimuli were similar pairs (192 total pairs) (see Fig. 1A). Similar stimulus pairs were selected as the most similar pair of objects in a larger set of ∼15–

Figure 1. Example stimulus pairs used in Experiment 1 (*A*) and Experiment 2 (*B*). The images shown to participants were in full color.

30 objects per set based on independent normative ratings (using a 7-point Likert scale). The remaining stimuli were unrelated random foils. Stimulus presentation and behavioral data collection were realized with the Cogent 2000 toolbox (http:// www.fil.ion.ucl.ac.uk) for Matlab (The MathWorks, Natick, MA).

Subjects were scanned while performing a continuous recognition task. Stimuli were displayed by means of a backprojection screen one at a time for 2500 msec with a 500-msec intertrial interval. Blocks consisted of 108 total trials. For each block, 32 of the similar sets were randomly chosen along with 44 unrelated foils. Half of the similar sets were randomly assigned to the "repeat" condition, in which only one of the stimuli was shown but at two different time points during the block, and the other half were assigned to the "lure" condition, in which both stimuli from the set were presented after a variable lag. Thus, for each block, there were 44 random (unrelated) foil trials, 32 "first" presentation trials (the first presentation of one of the stimuli from the lure sets), 16 "repeat" trials (the second presentation of the stimulus from the lure sets), and 16 "lure" trials (the presentation of the second member of the lure sets). Trials were presented in a pseudorandom order with the constraint that the mean lag separating the first and repeat or first and lure trials be $~\sim$ 30 trials (actual mean = 31.92 trials or 95.76 sec, SD = 20.96 trials, range = 1–105 trials). Subjects contributed between four and six blocks (mean = 5.85 , SD = 0.49).

Subjects were told that they would see several stimuli during the experiment and their task was to decide if they had seen the stimulus before in that block or not; they were instructed to respond "new" to novel stimuli and "old" to stimuli that they had seen before in that block. Subjects were informed that some of the stimuli would be similar, although not identical to previous stimuli. They were thus given a third response option, "similar," and told to make that response in the case that they believed the current stimulus was similar to something previously seen in that block. Subjects were shown a similar pair (such as those shown in Fig. 1) to demonstrate the lure trials. Subjects were given three MR-compatible response buttons and instructed to respond "new," "old," or "similar" to novel stimuli (foils and first trials), repeat trials, and lure trials, respectively. These instructions were displayed to the subjects before each scan block.

Experiment 2

Experiment 2 was conducted to test the hypothesis that different task demands would differentially affect pattern separation processes in the MTL. The stimuli in Experiment 1 were all namable objects that belonged to different sets. In Experiment 2, we used the same modified continuous recognition paradigm with face stimuli. Faces were chosen because they tend to be nonverbalizable (Diamond and Carey 1986), processed holistically (Yin 1969), and are homogeneous regarding set (Gauthier et al. 2000). Furthermore, the visual regions involved in processing face stimuli are relatively well described (Kanwisher et al. 1997; Haxby et al. 2000).

The stimuli used in Experiment 2 were 823 color portraitstyle photographs of 367 unique individuals. Each person was pictured in two to four different poses. There were several possible differences between pictures of the same person, such as gaze direction, lighting, expression, hairstyle, clothes, or a combination of any of these (see Fig. 1B). All stimuli were obtained from freely available online databases (Martinez and Benavente 1998; Nordstrom et al. 2004; http://lrv.fri.uni-lj.si/facedb.html; http://pics.psych.stir.ac.uk). Independent normative similarity ratings were obtained for the different pictures of the individuals. The average similarity rating was 3.88 (SD = 0.51) on a 7-point Likert scale (1 = very similar, 7 = very dissimilar; *n* = 14).

The behavioral procedure used the same modified continuous recognition paradigm as Experiment 1 with the following exceptions: of the 83 total trials per block, 35 were unrelated foils, 24 were first presentations, 12 were true repeats, and 12 were related lures. The mean lag between the first presentation and a repeat or lure was 19.32 trials (range 1–53, SEM = 2.88). Subjects were once again instructed that they would be asked to

perform a continuous recognition task. They were informed that some of the pictures would repeat ("same person, same picture"), and that some would be similar to a previous picture ("same person, different picture of that person"). Subjects were shown an example of two similar stimuli before the task began and were reminded of the response options between blocks of the experiment.

MRI data acquisition

MRI data acquisition parameters were based on those established in Kirwan et al. (2007). MRI data were collected on a Phillips 3T scanner (Best) equipped with a SENSE (Sensitivity Encoding) head coil at the F.M. Kirby Research Center for Functional Brain Imaging at the Kennedy Krieger Institute (Baltimore, MD). The scan order was as follows: first, a standard MP-RAGE was obtained for each subject (150 oblique axial slices, $1 \times 1 \times 1$ mm voxels). Next, several high-resolution MP-RAGE scans were also obtained (60 oblique axial slices, $0.75 \times 0.75 \times 0.75$ mm voxels). The MP-RAGE scans were later used for anatomical localization and cross-subject alignment. Following the structural MP-RAGE scans, the functional echoplanar images were collected using a high-speed echoplanar single-shot pulse sequence with an acquisition matrix size of 64×64 , an echo time of 30 msec, a flip angle of 70°, a SENSE factor of 2, slice thickness or 1.5 mm with no between-slice gap, and an in-plane acquisition resolution of 1.5×1.5 mm. In Experiment 1, a total of 216 volumes were acquired in each run with a TR of 1.5 sec. In Experiment 2, 166 volumes were collected per scan run. Each volume consisted of 19 oblique axial slices aligned to the long axis of the hippocampus and centered to include the hippocampus and most of the parahippocampal gyrus. Data acquisition began after the fourth image to allow for stabilization of the MR signal. Finally, following the functional data acquisition, one to two more highresolution MP-RAGE scans were obtained. The coregistered and averaged high-resolution MP-RAGE scans were used to delineate the subregions of the hippocampus in the ROI-LDDMM alignment process (see below).

fMRI data analysis

Data analysis was carried out using the Analysis of Functional Neuroimages (AFNI) software (Cox 1996). The data were coregistered in three dimensions to the standard whole-brain anatomical data. Functional data were also coregistered through time to reduce any effects of head motion. Time periods in which a significant motion event (more than 3° of rotation or 2 mm of translation in any direction) occurred, plus and minus one TR, were eliminated from the analysis.

Based on trial type and behavioral response, trials were sorted into one of the following bins: repeated stimuli called "old" (hits), lure stimuli called "old" (lure false alarms), lure stimuli called "similar" (lure correct rejections), and repeated or lure stimuli called "new" (miss). The first presentations of a stimulus from the similar pairs that were called "new" were sorted according to subsequent performance thusly: subsequent hit, subsequent lure false alarm, subsequent lure correct rejection, and subsequent misses. A further bin, foils correctly identified as "new," served as a baseline against which to compare the contrasts of interest. Finally, all other trial types not of immediate interest (including foils and first presentations called "old" or "similar") were also labeled and included in the model, but not included in subsequent analyses. Behavioral vectors were created based on these bins. All of these vectors were used to individually model each participant's functional data using a deconvolution approach based on multiple linear regression (http:// afni.nimh.nih.gov/pub/dist/doc/manuals/3dDeconvolve.pdf). The resultant fit coefficients (β coefficients) represent activity versus baseline for a given time point and trial type in a voxel. The sum of the fit coefficients over the expected hemodynamic response (∼3–12 sec after trial onset) was taken as the estimate of the model of the response to each trial type (relative to the nulltask baseline). In order to account for residual intersubject functional variation, the fit coefficient maps were blurred with a

3-mm FWHM Gaussian kernel that respected the anatomical boundaries of the individual subjects' MTL regions of interest. Briefly, the functional data were first masked with the set of anatomically defined ROIs on an individual subject's basis (see below for the full list of MTL ROIs and their delineation) to exclude all activity outside the ROI. The masked ROIs were then blurred with a 3-mm FWHM Gaussian kernel that respected the edge of the ROIs to not blur in zeros from outside the masked regions. The blurred data were then masked again with the anatomical ROIs so that the blurred activity respected the boundaries of the anatomical ROI. The blurred and masked data from each ROI were then recombined to form one blurred data set. Thus, the blurring occurred within the anatomical ROI, but did not cross into neighboring regions. This was done to account for any residual intersubject functional variability while still respecting the anatomical boundaries defined within the MTL (Kirwan et al. 2007).

Cross-subject alignment

The cross-subject alignment used the ROI-LDDMM method as described in Kirwan et al. (2007). First, all subjects' anatomical and functional scans were normalized to the Talairach atlas (Talairach and Tournoux 1988) using AFNI. Anatomical regions of interest were fully segmented in 3D on the Talairach transformed standard (1 mm^3) MP-RAGE images for the temporal polar, entorhinal, and perirhinal cortices according to the landmarks described by Insausti et al. (1998). 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, as in our previous research (Stark and Okado 2003; Kirwan and Stark 2004; Law et al. 2005; Okado and Stark 2005). The subfields of the hippocampus were also defined bilaterally as the DG/CA3 (dentate gyrus and CA3 field), CA1, and subiculum (SUB) (Wang et al. 2003; Csernansky et al. 2005) following the atlas of Duvernoy (1998) and following our previous work (Kirwan et al. 2007) using the high-resolution (0.75 mm^3) MP-RAGE. The anatomically defined ROIs were then used to calculate the ROI-LDDMM 3D vector field transformation for each subject using a central tendency modal model (or template) developed in our laboratory as the target. The ROI-LDDMM transformation for each individual subject was then applied to the fit coefficient maps for that subject.

This modal model was used not only as the template for the transformation of individual subject anatomical ROI maps, but also as the basis for localization of group data. One feature of this approach is that the anatomical location of activity in group analyses can be identified by consulting the anatomical ROI labels (e.g., right CA1) in this template (or even in each subject).

Results

Behavioral results

Table 1 shows the mean number of trials for each trial type according to the behavioral response in terms of both the raw num-

Table 1. Behavioral response by trial type

ber of trials and the percent of possible trials. In Experiment 1, subjects were highly accurate at identifying novel stimuli (96.61% and 96.68% for the foil and first presentations, respectively). Accuracy was also high for correctly identifying the repeated stimuli as "old" (83.48%). This performance was similar to the performance of subjects in similar versions of the continuous recognition task (Underwood 1965). Responses to the lure stimuli were divided between correctly identifying them as "similar" (54.49%) and incorrectly calling them "old" (33.61%). Mean reaction times (RTs) were similar for the subsequent hits (939 $[SD = 197]$, subsequent lure false alarms $(941 \; [233])$, and subsequent lure correct rejections (929 [222]) (all *P*s > 0.05). RTs for hits (1156 [138]) were faster than for lure false alarms (1266 [153]; $t_{(19)} = -6.07$, $P < 0.0001$) and for lure correct rejections $(1338 \text{ [}141\text{]; } t_{(19)} = -8.38, P < 0.0001$). RTs for lure false alarms were also faster than for lure correct rejections $(t_{(19)} = -3.03)$, *P* < 0.01). There was a significant correlation between the normative similarity rating and the average correct rejection rate for the similar pairs ($r = 0.311$; $F_{(1190)} = 20.35$, $P < 0.001$), indicating that stimulus pairs that were rated as highly dissimilar were more likely to be correctly identified as lures during the experiment.

Behavioral accuracy for the face stimuli in Experiment 2 was similar to, although slightly below that of Experiment 1 (Table 1). Again, RTs for subsequent hits (1132 [153]), subsequent lure false alarms (1130 [191]), and subsequent lure correct rejections (1135 [167]) did not differ statistically (all *P*s > 0.05). RTs for lure correct rejections [1459 [163]) were once again significantly longer than for hits (1345 [177]; $t_{(19)} = -4.38$, $P < 0.001$) or lure false alarms (1378 [153]; $t_{(19)} = -2.54$, $P < 0.05$). The longer reaction times for the lure correct rejections in both experiments is consistent with the hypothesis that subjects adopted a recall to reject strategy for these trials. There was not a significant correlation between the normative similarity ratings for the pairs of faces and the average rate of correctly rejecting lures during the experiment ($r = 0.02$; $F_{(1448)} = 0.107$, n.s.).

fMRI results

Experiment 1

To examine effects at the time of encoding, functional data from the first presentation trials only were subjected to a two-way ANOVA with subjects as a random factor and condition (subsequent hit, subsequent lure false alarm, and subsequent lure correct rejection) as a fixed factor. For this and subsequent analyses, functional ROIs were defined by setting a voxel-wise threshold for the F-map at $P = 0.03$ and a spatial extent threshold of a minimum ROI volume of 40 mm³ (12 voxels) . The use of both a

Mean (SEM) number of trials for each trial type according to stimulus type and behavioral response. Raw numbers for each category as well as percentage of possible responses (e.g., number of "old" responses to repeats relative to total number of repeats) are presented.

more general F-statistic rather than a direct pairwise contrast and of a somewhat liberal spatial extent threshold was meant to reduce voxel selection biases (see Baker et al. 2007). Voxels within each ROI defined in this manner were collapsed to create functionally defined ROIs, and all subsequent pairwise comparisons were assessed with a final alpha threshold of *P* < 0.05. The cluster analysis yielded four regions of activation in the MTL; one in the left hippocampus head, primarily within the CA1, two distinct regions in left parahippocampal cortex, and one in the left hippocampal tail, primarily within CA1.

Figure 2 shows the activity within the functionally defined ROIs during trials in which a stimulus was first presented as a function of subsequent performance. Post hoc *t*-tests revealed reliable differences between encoding trials with subsequent lure correct rejections and subsequent lure false alarms in each of the regions defined by the ANOVA (see Table 2). Activity for subsequent lure correct rejections was greater than that for subsequent lure false alarms in three of the four regions. In the remaining region (posterior left CA1), activity for subsequent lure false alarms was significantly greater than for both subsequent lure correct rejections as well as for subsequent hits. In the anterior CA1 and the two parahippocampal cortex ROIs, activity for subsequent hits and subsequent lure false alarms also differed, with subsequent hits being greater than subsequent lure false alarms (Table 2). Finally, within the left anterior CA1 ROI, activity for subsequent lure correct rejections and subsequent hits differed significantly, with subsequent lure correct rejections being greater. Bear in mind that to correctly reject a lure at time of retrieval, one must encode a sufficient level of detail about the original stimulus in order for it to be distinguished from the new stimulus representation when presented with the lure. Whereas, if one false-alarms to a lure, the representations of the original stimulus and the lure are indistinguishable, thus eliciting a response of "old." This implies inadequate encoding of diagnostic details at the time of encoding for these stimuli. This interpretation is bolstered by the fact that activity for subsequent hits (repeats called "old") was reliably different from subsequent lure false alarms in each of these areas, indicating that activity in these regions was indicative of later retrieval success.

To examine activity ostensibly at the time of retrieval, a second two-way ANOVA was performed on the fMRI data from the repeat and lure trials with subjects as a random factor and condition (hit, lure false alarm, and lure correct rejection) as a fixed factor. In order to account for any global decreases in activity for the repeated presentations of the stimuli, a fourth level of the condition factor was included that corresponded with correct identification of novel foils (the baseline condition). Again,

the resultant F-maps were thresholded with a voxel-wise threshold of $P = 0.03$ and a spatial activation extent of 40 mm³, and subsequent functional-ROI analyses were conducted with a final alpha threshold of *P* < 0.05. This analysis yielded several regions of significant activation, detailed in Figure 3. These included regions in left perirhinal cortex, DG/CA3, SUB, parahippocampal cortex, and right CA1 (tail). The pattern of activity differed across these regions; in parahippocampal and perirhinal cortex, activity for lure false alarms was significantly lower than that for both lure correct rejections and hits, which were identical to each other. In both the left DG/CA3 and right CA1 of the hippocampus, activity for lure false alarms was significantly lower than that for hits, but did not differ reliably from that for lure correct rejections. Activity for hits was significantly greater than that for lure correct rejections in left SUB and right CA1. Thus, in the functional ROIs that fall within the MTL cortical regions (parahippocampal and perirhinal cortex), we observe a consistent pattern of activity in which activation for lure false alarms was significantly lower than that for lure correct rejections or hits. This pattern of activity at retrieval is consistent with that seen during encoding, with more activity in the conditions that require more diagnostic detail from the original stimulus, consistent with a recall to reject interpretation. The pattern of activity in the regions within the hippocampus, however, differed from that observed in the parahippocampal gyrus regions. In the left DG/CA3 region, the difference between lure false alarms and lure correct rejections failed to reach significance $(t_{(19)} = -1.82, P = 0.085)$. In the remaining hippocampal regions, activity for hits was significantly greater than for lure correct rejections (both left SUB and right CA1) and lure false alarms (right CA1) (see Table 2). Thus, it appears that during retrieval, the activity in the regions of the parahippocampal gyrus may reflect recall to reject processes, but activity within the hippocampus appears to be more complex, perhaps reflecting a mix of signals. However, considering activity across hippocampal structures and the structures in the parahippocampal gyrus, it was only the hippocampus that provided signals that differentiated these critical trial types. It is only in the hippocampus that we observed activity that isolated a lure as being similar to but not identical to a previously seen object. Thus, it was only the hippocampus that exhibited activity consistent with overcoming interference.

Experiment 2

There were somewhat fewer lure false alarms for the faces than for the objects (on average 19.9 vs. 30.5 for the faces and objects, respectively). In order to obtain an accurate estimate of activity for this condition, subjects with fewer than 15 total lure false

> alarms (five subjects) were excluded from the subsequent analyses. A twoway ANOVA conducted on the data from the first presentation trials with encoding condition (subsequent hit, subsequent lure false alarm, and subsequent lure correct rejection) as a fixed factor and subjects as a random factor revealed three regions of interest: one centered largely in left subiculum, but also including voxels in left CA1 and DG/CA3, a second in left perirhinal cortex, and a third in right CA1 and DG/CA3. In each of these regions, activity for subsequent lure correct rejections was significantly greater than for both subsequent hits and subsequent lure false alarms (Table 3). In order to correctly reject a subsequent lure, a distinct, detailed represen-

Figure 2. Encoding activity in the MTL for object continuous recognition task (Experiment 1). (*A*) Mean activity within the functionally defined ROIs pictured in *B*. For paired comparison *P*-values, see Table 2. (L) Left; (R) right; (PHC) parahippocampal cortex; (FA) false alarm; (CR) correct rejection.

Table 2. Paired comparison results from Experiment 1

	Conditions	Df	t	P
Objects-Encoding				
L CA1 (head)	Subs hit-subs lure FA	19	4.38	< 0.001
	Subs hit-subs lure CR	19	-3.86	< 0.01
	Subs lure FA-subs lure CR	19	-5.74	< 0.001
L CA1 (tail)	Subs hit-subs lure FA	19	-3.99	< 0.001
	Subs lure FA-subs lure CR	19	3.28	< 0.01
L parahippocampal cortex (1)	Subs hit-subs lure FA	19	6.11	< 0.001
	Subs lure FA-subs lure CR	19	-4.96	< 0.001
L parahippocampal cortex (2)	Subs hit-subs lure FA	19	4.43	< 0.001
	Subs lure FA-subs lure CR	19	-3.14	< 0.01
Objects-Retrieval				
L CA3/DG	Hit-lure FA	19	5.22	< 0.001
	Lure FA-lure CR	19	-1.82	0.085
L perirhinal cortex	Hit-lure FA	19	5.41	< 0.001
R CA1	Hit-lure FA	19	4.79	< 0.001
	Hit-lure CR	19	2.88	< 0.01
L SUB	Hit-lure CR	19	2.74	0.01
L parahippocampal cortex	Hit-lure FA	19	3.06	< 0.01
	Lure FA-lure CR	19	-3.30	< 0.01
L parahippocampal cortex	Hit-lure FA	19	3.81	0.001
	Lure FA-lure CR	19	-3.47	< 0.01

Results of paired comparisons in functionally defined ROIs at both encoding and retrieval for the object stimuli in Experiment 1.

tation of the original stimulus must be established. Failure to do so may result in a weak representation that leads to a false alarm. Thus, the observed pattern of activity (greater activity for subsequent lure correct rejections than for subsequent lure false alarms; Fig. 4) is again consistent with these regions participating in encoding of detailed representations sufficient for later accurate performance.

The data from retrieval trials were subjected to the same ANOVA as the retrieval trials in Experiment 1, with condition as a fixed and subject as a random factor. This analysis yielded three ROIs: the first in left CA1, another in right entorhinal cortex, and the third in left SUB (Fig. 5). Again, consistent with the results of the previous experiment, the regions in hippocampus show a different pattern of activation from those in the adjacent cortex. For the right entorhinal cortical ROI, activity for lure false alarms was significantly lower than that for hits or lure correct rejections (see Table 3), similar to the perirhinal and parahippocampal results from Experiment 1. In the left CA1 and left subiculum, activity for lure correct rejections was significantly greater than that for both hits and lure false alarms. Activity in both regions for hits was, in turn, significantly greater than that for lure false

alarms. While the details of the pattern of activation differed from those of Experiment 1, the basic finding of the hippocampus and not the parahippcoampal gyrus demonstrating signals that isolate the critical trial types is consistent.

Discussion

We scanned subjects using highresolution fMRI while they performed a modified continuous recognition task using either object or face stimuli. The experimental manipulation within each experiment had two conditions: one in which the same stimulus repeated and another in which a similar lure was presented after a similar lag. Subjects were required to indicate if the stimulus was new, if they recognized it as having been presented before in the experiment, or if

it was new but similar to a previously presented stimulus. We sorted first and subsequent presentation trials according to stimulus type and behavioral performance. During first-presentation trials, we observed activity in the parahippocampal gyrus and hippocampus that predicted later success in this task, with significant differences between subsequent lure correct rejections and lure false alarms. This activity may have reflected encoding of episodic details necessary for processes such as recall to reject for both repeat and lure stimuli. During repeat and lure trials, activity in the parahippocampal gyrus was also consistent with this interpretation, with greater activity for lure correct rejections than for lure false alarms. Within the hippocampal region, we observed several distinct patterns of activity that together distinguished hits, lure correct rejections, and lure false alarms. Thus, in a task that places great demands on pattern separation by increasing interfer-

ence, it was the hippocampus that demonstrated activity that was correlated with behavioral performance and distinguished trial types, while this differentiation was not observed in the parahippocampal gyrus. Of course, it is possible that the parahippocampal gyrus contained information that distinguished across all trial types, that was not reflected in an overall change in the BOLD fMRI signal. While we cannot eliminate the possibility that such signals might therefore exist and have gone undetected in the parahippocampal gyrus, such signals were clearly present in the hippocampus.

When examining how participants overcome the interference and exhibit behavioral pattern separation, the most informative contrast is between the lure correct rejections and the lure false alarms. On lure trials, accurate performance (responding "similar") required recognizing that something like the lure stimulus had been presented previously in the experiment and the further step of determining that the prior stimulus was, in fact, not this particular stimulus, but something only similar. This required both establishing an accurate, distinct representation of the stimulus at time of encoding the first presentation of the stimulus, and retrieving that representation when presented

Table 3. Paired comparison results from Experiment 2

	Conditions	Df	t	P
Faces-Encoding				
L SUB	Subs hit–subs lure CR	14	-3.19	< 0.01
	Subs lure FA-subs lure CR	14	-5.44	< 0.001
L perirhinal cortex	Subs hit-subs lure CR	14	-2.52	< 0.05
	Subs lure FA-subs lure CR	14	-3.46	< 0.01
R CA1	Subs hit-subs lure CR	14	-3.19	< 0.01
L parahippocampal cortex	Hit-lure FA	19	3.06	< 0.01
	Lure FA-lure CR	19	-3.30	< 0.01
L parahippocampal cortex	Hit-lure FA	19	3.81	0.001
	Lure FA-lure CR	19	-3.47	< 0.01
Faces-Retrieval				
L CA1	Hit-lure FA	14	2.24	< 0.05
	Hit-lure CR	14	-4.27	< 0.001
	Lure FA-lure CR	14	-5.81	< 0.0001
R entorhinal cortex	Hit-lure FA	14	2.45	< 0.05
	Lure FA-lure CR	14	-3.08	< 0.01
L SUB	Hit-lure FA	14	2.82	0.01
	Hit-lure CR	14	-3.30	< 0.01
	Lure FA-lure CR	14	-4.73	< 0.001

Results of paired comparisons in functionally defined ROIs at both encoding and retrieval for the object stimuli in Experiment 2.

with a similar lure in order to correctly identify it as similar (i.e., recall to reject). Failure of this later step might have resulted in incorrectly endorsing the lure as "old" (lure false alarms).

Here we demonstrate that several MTL regions showed a pattern of activity consistent with "behavioral" pattern separation, with activity for lure correct rejections significantly different from that for lure false alarms during initial and subsequent stimulus presentations. This pattern of activation is consistent with the amount of diagnostic detail needed to drive the behavior in each trial type. The level of diagnostic detail necessary to support performance on hit trials could potentially vary from very little (as in lure false alarms) to a great deal (as in lure correct rejections). Indeed, the pattern of results observed in the parahippocampal gyrus bears this out, as activity for hits and lure correct rejections was indistinguishable during initial and subsequent stimulus presentations. Activation that distinguished between hits and lure correct rejections was observed exclusively in the hippocampus. In Experiment 1, activity associated with hits was numerically the highest in all three hippocampal ROIs, with many of the pairwise contrasts showing significant elevations of hippocampal activity during hit trials. The most parsimonious explanation of this pattern is that these signals reflect not only the amount of information retrieved, but also an additional

"matching" signal indicating a match between the stored representation and the actual stimulus. Overall, within the MTL, activity within the hippocampus but not the parahippocampal gyrus distinguished between the trial types of interest (hits, lure false alarms, and lure correct rejections) at both encoding and retrieval.

Other studies have used highresolution fMRI to investigate activity changes in the hippocampus with new declarative learning (Zeineh et al. 2000, 2003). Zeineh et al. (2003) demonstrated dynamic changes within hippocampal subregions that varied with encoding and retrieval conditions. The present results do not speak to this distinction, as the continuous recognition paradigm

This same continuum of difference between testing environments may not be entirely analogous to the difference between

overtly includes encoding and retrieval demands on each trial. Even when more traditional tests are used that do not overtly mix encoding and retrieval, there is clear evidence that encoding processes are engaged during retrieval tasks (Buckner et al. 2001; Stark and Okado 2003). It is worth noting, however, that activity changes observed in the hippocampus by Zeineh and colleagues is consistent with an interference interpretation; activity declined within the hippocampus during both encoding and retrieval as face–name associations became well learned, presumably correlating with reductions in interstimulus interference.

A key finding presented here is that in both experiments, activity in the hippocampus consistently distinguished between our trial types of interest. The exact pattern of activation within the hippocampus differed between objects and faces. The changes in hippocampal re-

sponses to different stimuli and task demands are not inconsistent with findings in the literature from electrophysiology in rodents. Lee et al. (2004) found that representations in CA1 were less coherent than those in CA3 in the face of shifts in local and distal cues on a circular track. The investigators interpreted these findings as evidence that CA1 performed pattern separation, while CA3 performed pattern completion in these circumstances. Leutgeb et al. (2004), on the other hand, found evidence for pattern separation in CA3 in response to more extensive shifts in testing environment. Guzowski et al. (2004) suggest that this is due to attractor dynamics in CA3, which cause it to respond in a nonlinear fashion as stimulus similarity changes. Data from an early immediate gene study conducted by Vazdarjanova and Guzowski (2004) lend further support for this argument. Rats were exposed to two identical environments, two environments that differed in either their proximal or distal cues, or two environments that differed in both their proximal and distal cues. Cell ensembles in CA3 showed overlap in the identical and smaller change conditions, but the coherence between ensemble activity dropped to chance levels when exposed to two completely different environments.

www.learnmem.org examples and the control of the control

plays a key role over and above adjacent MTL structures when strong demands are placed on pattern separation, such as when high amounts of interference

We thank Di Wu and the staff of the F.M. Kirby Center for Functional Brain Imaging for their assistance in data collection and Ken Norman for his comments on earlier versions of this manuscript. We gratefully acknowledge the use of the resources of the Center for Imaging Science at the Johns Hopkins University in the LDDMM alignment analysis. The funding source was NSF BCS-

Figure 5. Retrieval activity for face stimuli (Experiment 2). (*A*) Mean activation in the ROIs defined by the ANOVA, a subset of which are pictured in *B*. (L) Left; (R) right; (ERC) entorhinal cortex.

face and object stimuli in the present experiment. The decreases in behavioral accuracy for the faces may indicate that the faces were more (not less) similar than the objects, which one might suppose would lead to the different pattern of activation observed in the two experiments. Consistent between the two experiments, however, was the distinction between activity in hippocampal regions that distinguished between trial types, and activity in the parahippcampal gyrus that did not distinguish between hits and lure correct rejections.

Finally, we should note that the data presented here do not show clear differences among subregions of the hippocampus. The pattern of activity in CA1, CA3/DG, and the subiculum was qualitatively similar. As noted above, the data from the rodents suggest that numerous experimental factors can affect where separation is observed in the hippocampus. Given this and the fact that no contrast in the present study can cleanly isolate separation or completion signals, it would be imprudent to conclude that these data are inconsistent with proposed distinctions among the hippocampal subfields. It is quite possible that the present experimental design lacked the ability to discern the finegrained differences among hippocampal subfields while still being able to detect the coarser differences between the hippocampus and the adjacent cortical structures.

Conclusions

Pattern separation and pattern completion processes are required for accurate storage and retrieval of episodic memories. Here we assessed fMRI activity in the MTL that correlated with performance on a task that uses interference to place strong demands on behavioral pattern separation. We demonstrate that during encoding, regions in the hippocampus as well as in the adjacent MTL cortex predict overall encoding success. Likewise, during retrieval, regions in both the hippocampus and the adjacent MTL cortex exhibited activity correlated with overall successful retrieval of information. During both encoding and retrieval, however, the pattern of activity in the hippocampus was more complex than that in the parahippocampal gyrus. While the signals in the parahippocampal gyrus were consistent with basic encoding and retrieval success and the predictions of a "recall to reject" strategy (e.g., differentiating lure false alarms from hits and lure correct rejections), they did not fully differentiate all three trial types of interest (hits, false alarms to lures, and correct rejections of lures). Within the hippocampus, such differentiation was present. Thus, signals were present in the hippocampus that would allow participants to overcome the interference and accurately perform the task (e.g., correctly note that the present stimulus was a true repetition vs. correctly note that it was only similar to a prior stimulus). Therefore, we suggest that the hippocampus

References

Baker, C.I., Hutchinson, T.L., and Kanwisher, N. 2007. Does the fusiform face area contain subregions highly selective for nonfaces? *Nat. Neurosci.* **10:** 3–4.

must be overcome.

Acknowledgments

Buckner, R.L., Wheeler, M.E., and Sheridan, M.A. 2001. Encoding processes during retrieval tasks. *J. Cogn. Neurosci.* **13:** 406–415.

0236431.

- Caesar, K., Gold, L., and Lauritzen, M. 2003. Context sensitivity of activity-dependent increases in cerebral blood flow. *Proc. Natl. Acad. Sci.* **100:** 4239–4244.
- Cox, R.W. 1996. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29:** 162–163.
- Csernansky, J.G., Wang, L., Swank, J., Miller, J.P., Gado, M., McKeel, D., Miller, M.I., and Morris, J.C. 2005. Preclinical detection of Alzheimer's disease: Hippocampal shape and volume predict
- dementia onset in the elderly. *Neuroimage* **25:** 783–792. Diamond, R. and Carey, S. 1986. Why faces are and are not special: An effect of expertise. *J. Exp. Psychol. Gen.* **115:** 107–117.
- Duvernoy, H. 1998. *The human hippocampus.* Springer-Verlag, New York. Gauthier, I., Skudlarski, P., Gore, J.C., and Anderson, A.W. 2000. Expertise for cars and birds recruits brain areas involved in face
- recognition. *Nat. Neurosci.* **3:** 191–197. Guzowski, J.F., Knierim, J.J., and Moser, E.I. 2004. Ensemble dynamics
- of hippocampal regions CA3 and CA1. *Neuron* **44:** 581–584. Haxby, J.V., Hoffman, E.A., and Gobbini, M.I. 2000. The distributed
- human neural system for face perception. *Trends Cogn. Sci.* **4:** 223–233.
- Insausti, R., Juottonen, K., Soininen, H., Insausti, A.M., Partanen, K., Vainio, P., Laakso, M.P., and Pitkanen, A. 1998. MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *AJNR Am. J. Neuroradiol.* **19:** 659–671.
- Kanwisher, N., McDermott, J., and Chun, M.M. 1997. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17:** 4302–4311.
- Kirwan, C. and Stark, C.E.L. 2004. Medial temporal lobe activation during encoding and retrieval of novel face–name pairs. *Hippocampus* **14:** 919–930.
- Kirwan, C.B., Jones, C.K., Miller, M.I., and Stark, C.E.L. 2007. High-resolution fMRI investigation of the medial temporal lobe. *Hum. Brain Mapp.* doi: 10.1002/hbm.20331.
- Law, J.R., Flanery, M.A., Wirth, S., Yanike, M., Smith, A.C., Frank, L.M., Suzuki, W.A., Brown, E.N., and Stark, C.E.L. 2005. Functional magnetic resonance imaging activity during the gradual acquisition and expression of paired-associate memory. *J. Neurosci.* **25:** 5720–5729.
- Lee, I., Yoganarasimha, D., Rao, G., and Knierim, J.J. 2004. Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature* **430:** 456–459.
- Leutgeb, S., Leutgeb, J.K., Treves, A., Moser, M., and Moser, E.I. 2004. Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* **305:** 1295–1298.
- Leutgeb, S., Leutgeb, J.K., Barnes, C.A., Moser, E.I., McNaughton, B.L., and Moser, M.-B. 2005. Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science* **309:** 619–623.
- Leutgeb, J.K., Leutgeb, S., Moser, M., and Moser, E.I. 2007. Pattern separation in dentate gyrus and CA3 of the hippocampus. *Science* **315:** 961–966.
- Martin, A. 1999. Automatic activation of the medial temporal lobe during encoding: Lateralized influences of meaning and novelty.

Hippocampus **9:** 62–70.

- Martinez, A.M. and Benavente, R. 1998. *The AR Face Database*, CVC Technical Report #24. Computer Vision Center, Universitat Autònoma de Barcelona, Barcelona, Spain.
- McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* **102:** 419–457.
- Nordstrom, M.M., Larsen, M., Sierakowski, J., and Stegmann, M.B. 2004. *The IMM Face Database: An annotated dataset of 240 face images.* Informatics and Mathematical Modelling, Technical University of Denmark, Denmark.
- Norman, K.A. and O'Reilly, R.C. 2003. Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach. *Psychol. Rev.* **110:** 611–646.
- Okado, Y. and Stark, C.E.L. 2005. Neural activity during encoding predicts false memories created by misinformation. *Learn. Mem.* **12:** 3–11.
- O'Reilly, R.C. and Rudy, J.W. 2001. Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychol. Rev.* **108:** 311–345.
- Rolls, E.T. and Treves, A. 1998. *Neural networks and brain function.* Oxford University Press, Oxford.
- Schacter, D.L. and Wagner, A.D. 1999. Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. *Hippocampus* **9:** 7–24.
- Squire, L.R., Stark, C.E.L., and Clark, R.E. 2004. The medial temporal lobe. *Annu. Rev. Neurosci.* **27:** 279–306.
- Stark, C.E.L. and Okado, Y. 2003. Making memories without trying: Medial temporal lobe activity associated with incidental memory formation during recognition. *J. Neurosci.* **23:** 6748–6753.
- Stark, C.E.L. and Squire, L.R. 2001. When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proc. Natl. Acad. Sci.* **98:** 12760–12766.
- Talairach, J. and Tournoux, P. 1988. *A co-planar stereotaxic atlas of the human brain.* Thieme Medical, New York.
- Tulving, E. 2002. Episodic memory: From mind to brain. *Annu. Rev. Psychol.* **53:** 1–25.
- Underwood, B.J. 1965. False recognition produced by implicit verbal
- responses. *J. Exp. Psychol.* **70:** 122–129. Vazdarjanova, A. and Guzowski, J.F. 2004. Differences in hippocampal neuronal population responses to modifications of an environmental context: Evidence for distinct, yet complementary, functions of CA3 and CA1 ensembles. *J. Neurosci.* **24:** 6489–6496.
- Viskontas, I.V., Knowlton, B.J., Steinmetz, P.N., and Fried, I. 2006. Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. *J. Cogn. Neurosci.* **18:** 1654–1662.
- Wang, L., Swank, J.S., Glick, I.E., Gado, M.H., Miller, M.I., Morris, J.C., and Csernansky, J.G. 2003. Changes in hippocampal volume and shape across time distinguish dementia of the Alzheimer type from healthy aging. *Neuroimage* **20:** 667–682.
- Wills, T.J., Lever, C., Cacucci, F., Burgess, N., and O'Keefe, J. 2005. Attractor dynamics in the hippocampal representation of the local environment. *Science* **308:** 873–876.
- Yin, R.K. 1969. Looking at upside-down faces. *J. Exp. Psychol.* **81:** 141–145.
- Zeineh, M.M., Engel, S.A., and Bookheimer, S.Y. 2000. Application of cortical unfolding techniques to functional MRI of the human hippocampal region. *Neuroimage* **11:** 668–683.
- Zeineh, M.M., Engel, S.A., Thompson, P.M., and Bookheimer, S.Y. 2003. Dynamics of the hippocampus during encoding and retrieval of face–name pairs. *Science* **299:** 577–580.

Received June 6, 2007; accepted in revised form July 30, 2007.