HIPPOCAMPUS 22:389–398 (2012)

RAPID COMMUNICATION

Evidence for Area CA1 as a Match/Mismatch Detector: A High-Resolution fMRI Study of the Human Hippocampus

Katherine Duncan,[1] Nicholas Ketz,[1,2] Souheil J. Inati,[2] and Lila Davachi[1,2]*

ABSTRACT: The hippocampus is proposed to switch between memory encoding and retrieval by continually computing the overlap between what is expected and what is encountered. Central to this hypothesis is that area CA1 performs this calculation. However, empirical evidence for this is lacking. To test the theoretical role of area CA1 in match/mismatch detection, we had subjects study complex stimuli and then, during high-resolution fMRI scanning, make memory judgments about probes that either matched or mismatched expectations. More than any other hippocampal subfield, area CA1 displayed responses consistent with a match/mismatch detector. Specifically, the responses in area CA1 tracked the total number of changes present in the probe. Additionally, area CA1 was sensitive to both behaviorally relevant and irrelevant changes, a key feature of an automatic comparator. These results are consistent with, and provide the first evidence in humans for, the theoretically important role of area CA1 as a match/mismatch detector. © 2011 Wiley Periodicals, Inc.

KEY WORDS: hippocampus; area CA1; mismatch/mismatch signal; comparator; human; fMRI

INTRODUCTION

Our experiences can be quite similar from day to day, even composed of the same elements, yet our episodic memory system (Tulving, 1987) allows us to form distinct mnemonic representations of these highly overlapping events and, later, to retrieve each event when presented with only a subset of the associated elements. Numerous lesion, electrophysiology, and human imaging studies have provided evidence that the hippocampus plays a key role in both encoding and retrieval (Davachi, 2006; Eichenbaum et al., 2007; Mayes et al., 2007; Squire et al., 2007). However, an important challenge for any theoretical account of the hippocampal memory system is to explain how this system can de-emphasize similarities across events to lay down distinctive memory traces, yet also be able to utilize similarities to reactivate the details of an event. One solution proposed by many models of episodic memory is that the hippocampus is able to shift dynamically between encoding and retrieval modes (Hasselmo and Schnell, 1994; Hasselmo et al., 1995; Vinogradova, 2001; Norman and O'Reilly, 2003; Meeter et al., 2004; Lisman and Grace, 2005). This mode shift is thought to be controlled by calculating how consistent the environment is with previous experience; if an experience is consistent with expectation, then the hippocampal dynamics will favor retrieval or pattern completion, but if the two differ, the hippocampus will shift into a mode that supports the encoding of the unexpected event via pattern separation and enhanced long-term potentiation (LTP) (Hasselmo and Schnell, 1994).

This match/mismatch calculation is thought to be driven not simply by novelty per se, but instead requires a comparison between expectation and what is being experienced (Kumaran and Maguire, 2006, 2007a,b). Expectations are hypothesized to be generated by the recurrent collaterals in area CA3, which use environmental cues to retrieve previous experiences (Hasselmo et al., 1995; Lisman and Grace, 2005). These predictions are then automatically compared to sensory input transferred through the entorhinal cortex (EC) to determine if the environmental input mismatches the expectation (Hasselmo and Schnell, 1994; Lisman and Grace, 2005). The resulting signal is then hypothesized to be integrated with attentional and motivational information likely supported by acetylcholine (Hasselmo and Schnell, 1994) or dopamine (Lisman and Grace, 2005) networks to allow for more controlled processing based on the integrated signals. Consistent with these theories, recent electrophysiological recordings in humans have provided some evidence for these two distinct hippocampal responses; an early automatic signal in response to unexpected events, and a later, more task-related response that correlated with memory performance (Axmacher et al., 2010). Additionally, fMRI studies have highlighted the role of expectations, by demonstrating that the hippocampus responds maximally when some portion of a stimulus is familiar and thus likely to initiate an expectation, while the other portion of the stimulus is unexpected in that context (Kohler et al., 2005; Kumaran and Maguire, 2006, 2007a).

Empirical tests of the precise locus of this signal are, as yet, lacking. In many of the models concerned with this match/mismatch calculation (Hasselmo and Schnell, 1994; Hasselmo et al., 1995; Meeter et al.,

[1] **Department of Psychology, New York University, New York;** [2] **Department of Psychology and Center for Neural Science, New York University, New York**
Change of address for Souheil Inati who is now at: National Institute of Mental Health, Bethesda, MD 20892
Additional Supporting Information may be found in the online version of this article.
Grant sponsor: Lila Davachi; Grant number: NIMH MH074692;
*Correspondence to: Lila Davachi, Department of Psychology, New York University, 6 Washington Place, Rm. 409, New York, NY 10003, USA.
E-mail: lila.davachi@nyu.edu
Accepted for publication 6 January 2011
DOI 10.1002/hipo.20933
Published online 11 April 2011 in Wiley Online Library (wileyonlinelibrary.com).

© 2011 WILEY PERIODICALS, INC.

2004; Lisman and Grace, 2005; Kumaran and Maguire, 2007b) it is hypothesized that area CA1 performs the comparison because it receives direct input from both the EC and area CA3. However, the human studies that have found evidence for match/mismatch signals in the hippocampus have been conducted at too course a spatial resolution to determine which hippocampal subfields contributed to the signal (Kohler et al., 2005; Kumaran and Maguire, 2006, 2007a; Hannula and Ranganath, 2008; Duncan et al., 2009; Axmacher et al., 2010; Dudukovic et al., 2010). Additionally, the results from animal research, which has the spatial resolution to isolate hippocampal subfields, do not paint a clear picture. On the one hand, pyramidal cells in area CA1 have been shown to increase their firing rate when an escape platform appeared in an unexpected location (Fyhn et al., 2002), but these recordings were only made in area CA1 leaving open the possibility that neurons in upstream areas had similar properties. On the other hand, selective lesion of areas CA3 and dentate gyrus (DG) significantly impared rats' ability to detect displaced objects, while animals with CA1 lesions had comparatively preserved performance (Lee et al., 2005), although the surgery was performed prior to learning so the deficits could have resulted from a deficit in mismatch detection or a deficit in the initial encoding.

To address these inconsistent results, the current study used high-resolution fMRI as a complementary method to investigate match/mismatch signals arising from distinct human hippocampal subfields. We predicted that a match/mismatch signal should scale with the number of changes present in an image. Furthermore, since this response is hypothesized to reflect the automatic comparison of expectations and the environment (Kumaran and Maguire, 2007b), this signal should reflect the total amount of change in the environment regardless of behavioral relevance. To test this hypothesis, we developed a paradigm with several key features. First, we presented subjects with images of well-learned rooms that contained various combinations of different kinds of changes (see Fig. 1 for details of the task) so that regions that have a general response to changes, both task relevant and irrelevant, could be identified. Second, we presented these probe images after subjects were cued with the name of the room. This allowed them to develop an expectation by, in essence, bringing a part or all of the image representation online prior to comparing it with the probe image. Finally, we used probe images that were largely composed of old features so that each probe could be used as a retrieval cue. This means that the signals associated with recollection should be similar across conditions and that response differences should be largely related to the degree to which an image matches expectations. Together, these features allowed us to identify which hippocampal subfield(s), if any, exhibits responses that would be consistent with match/mismatch detection.

MATERIALS AND METHODS

Subjects

Twenty-four (10 female) right-handed native English speakers with normal or corrected to normal vision participated in the study. Two subjects were removed from the analysis because of below chance performance on one of the tasks (Layout Task for one subject and Furniture Task for the other), one subject was removed for excessive movement, and one subject was removed because of poor data quality. The remaining 20 subjects had a mean age of 25.4 with a range of 21 to 34. Consent was obtained in a manner approved by the Institutional Review Board at New York University.

Design

Participants performed two memory tasks (the Furniture and Layout Tasks) while in the scanner. In both tasks, the name of a familiar room (see Training Procedures) was presented as a cue and was followed by an image that contained 0, 1, or 2 changes in the layout along with 0, 1, or 2 changes to the individual pieces of furniture. In the Furniture Task, participants were asked to indicate if all of the furniture was the same as in the studied image and ignore changes to the layout, while in the Layout Task the participants were asked to indicate if the layout was the same as the studied image and ignore changes to individual pieces of furniture. These manipulations result in a 2 (Task: Furniture, Layout) by 3 (Relevant change: 0, 1, 2) by 3 (Irrelevant change: 0,1, 2) factorial design. The crossing of relevant and irrelevant changes produces the nine types of visual stimuli that were used in each task.

Stimuli

In total, 286 unique stimuli (16 for training only) were generated using Punch! Home Design Software (Kansas City, MO). The nine types of changes (0, 1, or 2 relevant $\times$ 0, 1, or 2 irrelevant changes) were applied to 30 distinctive rooms to generate the 270 experimental stimuli. To produce relevant and irrelevant manipulations, changes were made to the layout of the rooms by swapping the locations of two pieces of furniture and to the furniture in the rooms by replacing a piece of furniture with a piece of the same type, e.g. replacing a bed with a different bed (see Figure 1 for examples). Each original room had a distinctive layout that allowed all the furniture to be clearly visible after each manipulation. Additionally, relevant and irrelevant changes were equally likely to occur at each spatial position across rooms and nearly equally likely within a room. Stimuli were constrained such that relevant and irrelevant changes never occurred simultaneously to the same object. Stimuli in both the training and experimental sessions were presented using the psychophysics toolbox (version PTB-3) (Brainard, 1997) for Matlab (MathWorks, Sherborn, MA).

Training Procedures

Participants were trained on the 30-room stimuli both during a separate session $\sim$24 h before the scanning session and again immediately before the scan. Each room was given an identifying name that included the name of the family who owned it along with the type of room (for example, Johnson's Living Room). In the first training session, subjects performed

four tasks with the images. In the first task they were exposed to each room along with its name and were asked to rate their liking of the furniture and layout. The following three memory tasks required them to identify the 30 rooms with a 90% accuracy criterion for at least two consecutive cycles. In the first memory task they were presented with each room and were asked to identify the name of the owner. In the second task the background cues were removed from the images so that memory performance was only dependent on the furniture. In the third task, half of the image, either top, bottom, left, or right was covered with a black box to encourage the participants to distribute their attention across multiple pieces of furniture. Participants performed the second and third task again during the second session that took place right before the scanned portion. The training on the first day generally took participants 1.5 to 2 h to complete and ~30 min to complete on the second day.

Participants were given instructions for both experimental tasks after they finished the training and then completed nine practice trials per task (one with each type of manipulation). To ensure they understood the task, participants were asked to tell the experimenter what response they were going to make and why before pressing the response button. At the end of each practice trial, the originally studied room was briefly presented so that the participants could get a sense of the types of manipulations that could occur. During the initial calibration scans the participants passively viewed the 30 rooms (5 s presentation) to refresh their memories and to expose them to the images in the new setting.

Behavioral Procedures

Each trial began with the presentation of the room name along with the name of the task (Furniture or Layout) for 1.5 s. This was followed by 1 s of blank screen and the presentation of a probe image for 4 s during which a response was made. In each image, 0, 1, or 2 pieces of furniture could be changed (e.g., replacing a bed with another bed) and 0, 1, or 2 swaps in location could be made (e.g., the bed moving to the position previously occupied by the dresser and the dresser moving to the bed's previous location). In the Furniture Task, participants were instructed to indicate "yes" with their left middle finger when the probe image contained all the original pieces of furniture regardless of their locations and "no" with their left index finger if there were 1 or 2 changed pieces of furniture. Similarly, in the Layout Task, participants were instructed to indicate "yes" with their middle finger when all pieces of furniture were in their original location, disregarding any changes to the individual pieces of furniture, and "no" with their index finger if there were 1 or 2 swaps between pairs of objects. A variable ITI followed the probe presentation and lasted between 1 and 13.5 s. During all ITI periods greater than 1 s, participants performed a secondary task where they indicated the direction of an arrow that randomly alternated between left and right every 0.5 s. This task was selected to reduce hippocampal activity between trials and, thus, improve our estimate of its response to experimental stimuli (Stark and Squire, 2001). Each run ended in 12.5 s (5 TRs) of the active baseline task to allow for the full response to the last experimental trial to be fully recorded. Additionally, the first 12.5 s of each run did not include any trials to allow time for signal normalization.

The experimental tasks were split across 10 runs (alternating between the Layout and Furniture Tasks) and subjects received performance feedback at the end of each run containing the percent of matches and changes correctly detected. At the end of the sixth run, the anatomical scan was run (11 min) and subjects were re-presented with the original images to release them from the interference that accumulates when viewing manipulated images.

Each of the 10 runs contained 27 trials, resulting in a total of 135 trials per task. There were an equal number of trials with each type of perceptual change in a run, which resulted in 45 trials with each level of relevant (0, 1, or 2) and irrelevant (0, 1, or 2) change per task. The presentation sequences and ITI lengths were determined by the "Optseq2" algorithm (available at: http://surfer.nmr.mgh.harvard.edu/optseq/) to optimize the design for subsequent estimation of condition specific effects. Participants were run on one of the sequences in pairs such that they saw the same sequence of images but performed different tasks on them (e.g., subject 1a would start with the Layout Task and subject 1b would see the same stimuli but start with the Furniture Task).

fMRI Methods

All scanning was performed using a 3T Siemens Allegra magnetic resonance imaging (MRI) system with a whole-head coil. Functional data was collected using a zoomed high-resolution echo-planar pulse (EPI) sequence similar to that used in Olman et al. (2009) (TR = 2,500 μs, TE = 49, FOV = 192 × 96, 26 interleaved slices, distance factor of 20%, 1.5 × 1.5 × 2 mm voxel size, 10 runs, 125 volumes per run). The oblique coronal slices were aligned perpendicular to the hippocampal long axis. The field of view was reduced in the phase encode direction to reduce the total read out time thus minimizing distortions and artifacts (Olman et al., 2009). Saturation bands were used to suppress signal for tissue superior and inferior to the region of interest. A T1-weighted high-resolution MPRAGE sequence (magnetization-prepared rapid-acquisition gradient echo) (1 × 1 × 1 voxel size, 176 sagital slices) was collected following the completion of the sixth experimental run and was used as a high-resolution anatomical scan. Visual stimuli were projected onto a screen that was viewed through a mirror attached to the subject's head-coil.

Separate anatomical regions of interest (ROIs) were drawn for the DG/CA2/CA3, the CA1, and the subicular subfields of each participant's hippocampi. Before drawing, the T1-anatomical images were resampled to 0.75 mm isotropic voxels using a bicubic interpolation algorithm and then processed with the SUSAN nonlinear noise reduction tool in FSL to improve image quality. A single observer (KD) drew ROIs for all participants using a similar procedure to Kirwan et al. (2007).

The coronal plates from Duvernoy (1998) were matched to the most similar coronal slice in each hippocampus and used as templates, and then the remaining slices were filled in to form continuous regions. This process was performed while viewing slices from each plane along with a three-dimensional rendering to ensure that the drawing respected all anatomical boundaries (See Supporting Information Figure 2A for example ROIs).

Preprocessing of functional data was conducted with FSL (FMRIB Software Library; available online at: http://www.fmrib.ox.ac.uk/fsl). The first five volumes (12.5 s) were discarded to allow for signal normalization and the slicetimer function was used to correct for differences in slice acquisition timing. Next, the EPI volumes were motion corrected across runs by aligning to the last functional image recorded before the MPRAGE (the final volume from the sixth run) using the MCFLIRT function. One subject had three TRs of sudden movement (3 mm displacement) in a single run so these volumes were excluded from all subsequent analyses. All other subjects included in the analyses did not move more than 2 mm in any direction over the course of the experiment. Finally, subject specific anatomical ROIs were used as masks for a 3-mm Gaussian smoothing kernel so that only data from the same hippocampal subfield would be smoothed with itself. Like other smoothing procedures, this process improves the signal-to-noise (S/N) ratio and cross-subject registration, but it has the additional benefit of maintaining spatial specificity.

Cross-subject registration was accomplished using the ROI-Demons tool (Yassa and Stark, 2009). First, all subjects' anatomical images were aligned to their mean EPI image and the calculated transformation parameters were applied to the anatomical ROIs. Next, aligned anatomical images were transformed into Talairach space and the resulting affine transformations were applied to the aligned ROIs. Diffeomorphic transformations were then calculated (MedINRIA; available at: www-sop.inria.fr/asclepios/software/MedINRIA/about.php) to bring each subject's Talairached ROIs into registration with the first subject's. These transformations were applied to the first-level parameter maps, which had been resampled to the anatomical resolution (1 mm isotropic). These aligned parameter maps were included in the second-level models. Supporting Information Figure 2 includes example ROIs and overlap metrics for the transformed ROIs.

FMRI Statistical Analysis

Functional data were analyzed using a general linear model (GLM) implemented with 3dDeconvolve in AFNI (Cox, 1996). Separate run and linear drift regressors were included in first-level models as effects of no interest. The cue and probe portions of a trial were separately modeled as boxcars lasting for the duration of the event (1.5 s for the cue and 4 s for the probe) convolved with a double gamma function to approximate the hemodynamic response. A statistical threshold of $P < 0.005$ and 14 contiguous voxels (three original voxels) was used to determine which clusters were statistically reliable.

Total change model

The Total Change Model was used to identify voxels that were linearly modulated by the total number of changes in a stimulus, regardless of the relevance of the changes. The model contained an event regressor for all correct trials in each of the tasks along with a single parametric regressor for each task that tracked the total number of changes (both relevant and irrelevant) in a trial. For example, if a probe during the Furniture task contained 2 furniture (i.e., relevant) and 1 layout (i.e., irrelevant) change, it would be coded with a "3." Separate regressors of no interest were included for incorrect trials, linear drift, and run number.

Main effect analyses

These models separately estimated responses to each condition so that differences that did not follow the predicted linear patterns could be evaluated. The Change Model separately estimated the response for each combination of relevant and irrelevant change in each task. The cue periods were binned according to task and accuracy to form four regressors. All incorrect probe periods in a task were included as single regressors. The correct probe periods were further divided into each cell of the design for an additional 18 regressors (three levels of relevant change by three levels of irrelevant change by two levels of task). Parameter estimates from these regressors were included in two second-level models that estimated the effects of relevant and irrelevant change using a 3×3 analysis of variance (ANOVA) for each task. Clusters of voxels that displayed main effects of relevant and irrelevant change were identified in group parameter maps. This model also provided the estimates for the visualization of condition specific responses in regions identified by parametric models.

Regions that predicted the accuracy of a response were isolated using a separate model. It contained a single regressor for cue periods and four regressors for probe periods that were binned according to accuracy and task. Probe parameter estimates were included in a second-level ANOVA model that estimated main effects of task and accuracy. Additionally, contrasts for simple effects of accuracy for each task were included.

Interaction analysis

To directly compare the response to change across hippocampal subfields, all voxels (no cluster extent) that displayed a main effect of either task relevant or irrelevant change in either task (from the above Change Model) were selected. The parameter estimates for each change condition (total number of changes = 0 through 4) were then averaged across voxels within a given hippocampal subfield. These estimates were submitted to a Generalized Estimating Equation (GEE) model that included the total number of changes and the subfields (CA1 as reference category) as factors along with their interaction. Importantly, the total number of changes was included as a continuous measure so that the model would estimate how well fit each region was by a linear response to change.

Linear change score

The Linear Change Score was developed so that the linearity of each region's response to the total number of changes could be directly compared. We used correlation values as our measure of fit. Each subject's response was correlated with the total number of changes in each task resulting in a set of correlation values for each hippocampal subfield. If a task resulted in a negative correlation with a region's response, we multiplied the correlation values by -1 for that task and subfield. This was done because we wanted to have a score that reflected the linear fit but was agnostic to the direction of the correlation. Finally, all statistical tests were performed on Fisher *Z*-transformed correlation values.

RESULTS

Behavioral Results

During the experiment, subjects were cued to make one of two decisions on probe images of rooms that contained 0–2 manipulations to individual pieces of furniture (e.g., replacing a bed with a different bed) along with 0–2 changes to the layout of the furniture (e.g., swapping the locations of a dresser and a table). In the Furniture Task, subjects were asked to indicate if the pieces of furniture were the same as in the studied images, regardless of their layout, and in the Layout Task, subjects were asked to indicate if the layout was the same as in the studied image, regardless of changes to individual pieces of furniture. The average accuracy and response times for each task can be found in the Supporting Information Figure 1. A 3 (Relevant change) $\times$ 3 (Irrelevant change) $\times$ 2 (Task) repeated measures ANOVA was conducted for accuracy. A small but reliable main effect of the task was found ($F_{(1,19)} = 15.6$, $P < 0.001$), with the average performance on the Layout Task being 6% higher than the Furniture Task. There was a main effect of relevant change ($F_{(2,38)} = 51.5$, $P < 0.001$) with planned comparisons showing higher performance on trials with 0 and 2 relevant changes than on trials with 1 relevant change ($t_{(19)} = 5.2$, $P < 0.001$ and $t_{(19)} = 12.1$, $P < 0.001$ respectively). Importantly, there was no interaction between task and number of relevant changes, $F < 1$, indicating that the manipulation of the task relevant changes had a similar effect across tasks. There was a main effect for irrelevant change ($F_{(2,38)} = 4.8$, $P < 0.05$) with planned comparisons indicating a significant advantage on trials with 0 irrelevant changes compared to those with 2 irrelevant changes ($t_{(19)} = 3.5$, $P < 0.01$). There was also an interaction between task and irrelevant changes with the addition of irrelevant changes causing a more extreme decrement in performance in the Furniture than in the Layout Task. Finally, there was an interaction between relevant and irrelevant changes such that the improved accuracy that resulted from no irrelevant changes was only evident for trials that also included no relevant changes, i.e., for images that were identical to the studied image.

A 3 (Relevant change) by 3 (Irrelevant change) by 2 (Task) repeated measures ANOVA was also conducted for response times on accurate trials. Overall, the results were similar to those found in the accuracy analysis. The main effects for all variables were significant ($F_{(1,19)} = 9.2$, $P < 0.005$ for task; $F_{(2,38)} = 38.0$, $P < 0.001$ for relevant change; $F_{(2,38)} = 22.5$, $P < 0.001$ for irrelevant change). Responses took on average 38 μs longer in the Furniture task, indicating that there was no speed accuracy tradeoff between tasks. Response times were faster as the number of relevant changes increased with significant differences between all levels (all $P < 0.001$), consistent with a strategy of searching for changes. Conversely, response times were slower as the number of irrelevant changes increased with 2 irrelevant changes taking significantly longer than 0 or 1 (both $P < 0.001$), reflecting the time required to resolve interference from the irrelevant dimension.

Imaging Results

General match/mismatch signal

To isolate voxels that display a response consistent with a general match/mismatch signal, we used a diffeomorphic subfield alignment procedure (Yassa and Stark, 2009) registering subject parameter maps according to hand-drawn subfield ROIs, in effect bringing each subject's hippocampal subfields into alignment (see Methods for details). We modeled the data using a simple parametric model that coded for the total number of changes (0, 1, or 2 relevant changes plus 0, 1, or 2 irrelevant changes) in each task. For example, if an image contained 2 relevant changes in the Furniture Task (i.e., changes to individual pieces of furniture) and 1 irrelevant change (i.e., a change to the layout), that trial would be coded with a "3." Importantly, this model would be able to isolate voxels that have an increased response to more changes as well as voxels that have a maximal response to images that match the studied scene with a decreasing response to more changes. Allowing for both increases and decreases in response to change is a key feature of our analysis because whether a match/mismatch signal reflects the degree of match (Hasselmo and Schnell, 1994) or mismatch (Lisman and Grace, 2005) is not consistent across the proposed models. Both types of signal have also been reported across fMRI studies, with some finding greater hippocampal responses to probes that mismatch expectations (Kumaran and Maguire, 2006, 2007a; Duncan et al., 2009) and others to probes that match expectations (Hannula and Ranganath, 2008; Duncan et al., 2009; (Dudukovic et al., 2010).

Using this parametric approach, we found that the only hippocampal subregion to display a general match/mismatch signal was area CA1, consistent with comparator models of the hippocampus (Hasselmo and Schnell, 1994; Hasselmo et al., 1995; Meeter et al., 2004; Lisman and Grace, 2005; Kumaran and Maguire, 2007b). Figure 2a displays the cluster of voxels in the anterior portion of area CA1 that exhibits a parametric increase in BOLD response as the total number of changes increases during the Furniture Task. Importantly, no other

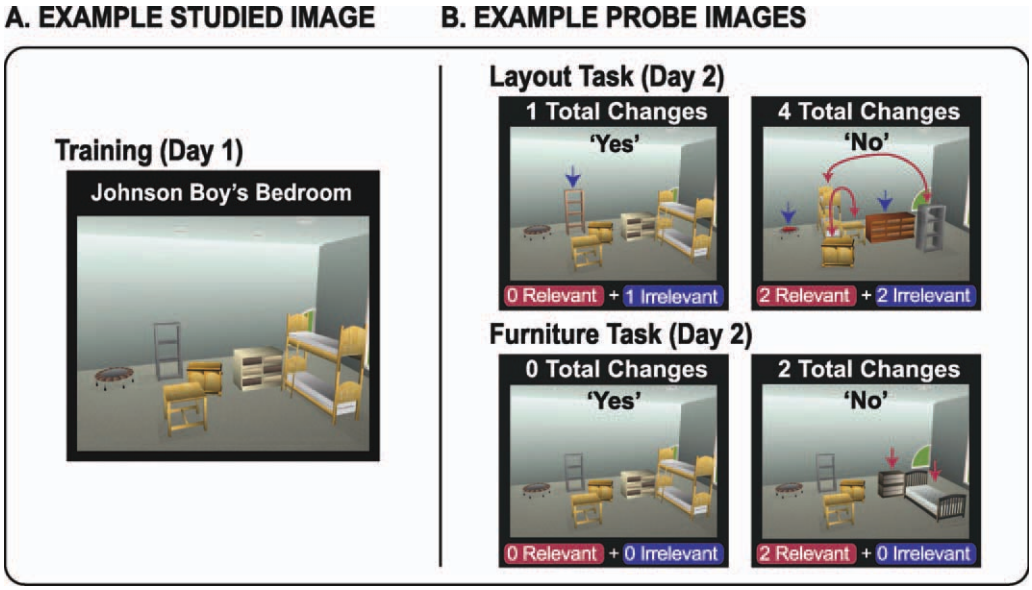

FIGURE 1. Experimental design. A. Example of a studied image. On day 1, subjects studied 30 distinctive rooms until they could reliably identify the name of each room. Subjects were retrained at the beginning of day 2, immediately before the experimental scan. B. Example probe images. On day 2, subjects performed the two experimental tasks, the Layout Task and the Furniture Task, in the MRI scanner. Each trial began with the presentation of a cue that consisted of the name of the room that would be presented next along with a task reminder. A probe image appeared 2.5 s later that could contain 0–2 furniture changes along with 0–2 layout changes. While performing the Layout Task, subjects were asked to indicate if the room had the same layout ("yes") or not ("no") while disregarding any changes to individual pieces of furniture. While performing the Furniture Task, subjects were asked to indicate if the room had the same furniture ("yes") or not ("no") while disregarding any changes to the layout. This resulted in probe images that contained 0, 1, or 2 task relevant changes (indicated in red) along with 0, 1, or 2 task irrelevant change (indicated in blue) for a total of 0 through 4 changes. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

clusters in the hippocampus were revealed to show this kind of general response to change.

Specificity of area CA1

To more directly compare the CA1 activity with activation patterns seen in other hippocampal subfields, we next selected voxels in all hippocampal subfields that were in any way sensitive to change, but which were not statistically biased to show a particular pattern of responses across conditions. Specifically, we first identified voxels that displayed a main effect of task relevant or irrelevant change in either task. We removed our cluster extent threshold for this identification because there were no clusters outside of area CA1 that displayed significant main effects. The parameter estimates from these "change sensitive" voxels were then submitted to a Generalized Estimating Equa-

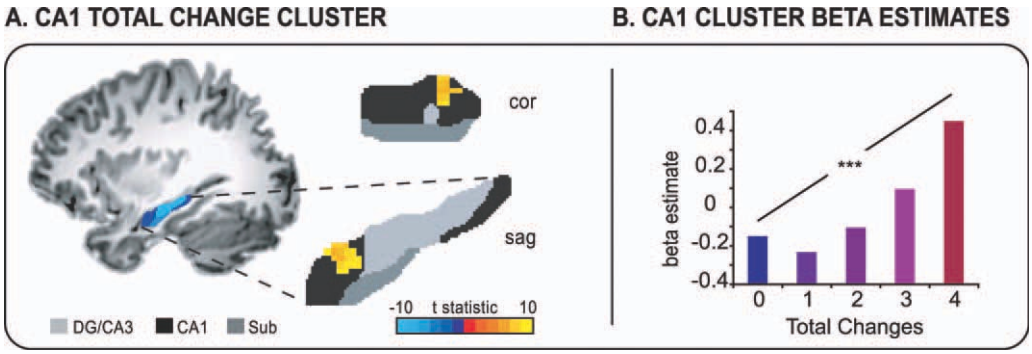

FIGURE 2. CA1 total change response. A. Group parameter map of voxels showing a significant parametric modulation according to the total number of changes during the Furniture Task. A reference whole-brain sagittal slice is included to the left with the hippocampal ROIs overlaid on it. On the right, parameter maps are displayed on the group mode hippocampal ROIs in both coronal and sagittal planes. Color indicates the value of the *t*-statistic. B. Group parameter estimates from the voxels contained in the Total Change Cluster. Estimates for each number of changes are plotted to illustrate the parametric effect that was used to isolate the voxels. All parameters are estimated using correct trials only. ***$P < 0.001$ for the linear trend. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

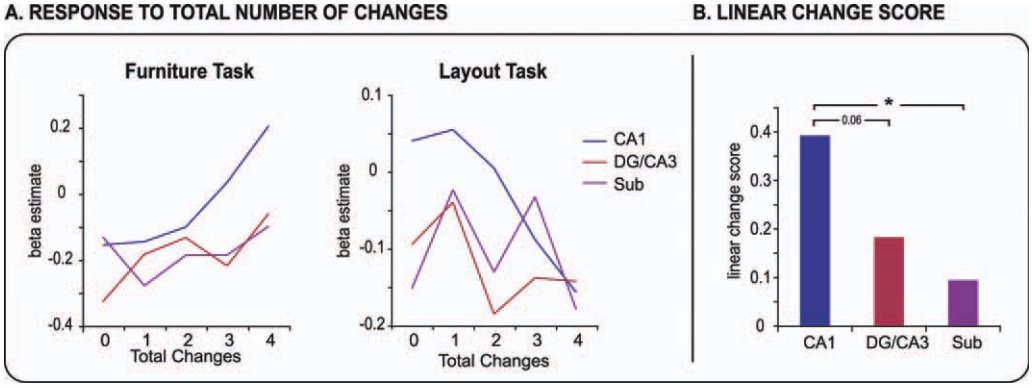

FIGURE 3. Total change responses for all hippocampal regions. A. Responses to the total number of changes for each subfield. Estimates for each number of changes are averaged across all voxels which displayed a main effect for either task relevant or irrelevant change in either task at a $P < 0.005$ level (no cluster extent). Area CA1 was the only region to show a significant linear trend in either task ($P < 0.05$ for both tasks; $P > 0.12$ for all other subfields). There was a significant interaction between region and total number of changes in the Furniture Task ($P < 0.05$) and a marginal interaction in the Layout Task ($P = 0.11$). All parameters are estimated using correct trials only. B. Linear change score for each region. This score reflects the correlation between the response on correct trials in that region and the total number of changes (see Methods for more details). *$P < 0.05$. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

tion (GEE) model that included both the total number of changes (0 through 4) as a linear predictor along with the hippocampal subfield (CA1, DG/CA2/CA3 and subiculum) to determine if there was an interaction between region and response to change. There was a significant interaction between subfield and number of changes in the Furniture Task ($P < 0.05$) driven by area CA1 being the only region to show a significant linear response to the total number of changes ($P < 0.05$; $P > 0.12$ in all other areas). There was also a marginal interaction between change and region for the Layout Task ($P = 0.11$) that was also driven by area CA1 being the only region to show a significant linear response to the total number of changes ($P < 0.05$; $P > 0.54$ in all other areas, see Fig. 3a).

Finally, we wanted to directly compare how well fit each region was by the total number of changes across tasks. To do this we developed a linear change score that combined the correlation between the response in each hippocampal region and the total number of changes across tasks as a measure of how well fit each region was to the linear effect of change (see Methods for details). When the linear change scores for each region were submitted to a repeated measures ANOVA, there was a significant main effect of region ($F_{(2,38)} = 3.74$, $P < 0.05$), that was driven by a significantly greater linear change score in area CA1 than in subiculum ($F_{(1,19)} = 6.0$, $P < 0.05$) and a marginally greater score in area CA1 than in area DG/CA2/CA3 ($P = 0.06$) (Fig. 3b). Together, the linear effects provide evidence that the response in area CA1, more than any other hippocampal subfield, tracks the total amount of change in an image, consistent with its hypothesized role in match/mismatch detection.

Task-related response

In addition to the general response to changes, a more task-related mnemonic signal may be expressed by additional regions of the hippocampus, as has been found in recent work (Duncan et al. 2009; Axmacher et al., 2010). To identify clusters showing this type of response we looked for voxels that displayed a main effect for task relevant change in each task. As

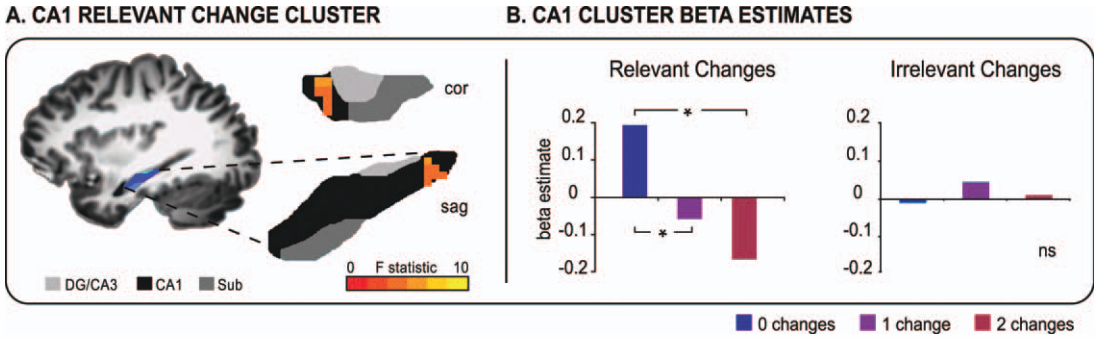

FIGURE 4. Task-related responses. A. Parameter map of voxels showing a task relevant main effect during the Layout Task overlaid on the modal segmentation of hippocampal subfields. A reference whole-brain sagittal slice is included to the left with the hippocampal ROIs overlaid on it. Color indicates the value of the *F*-statistic for the main effect of task relevant change during the Furniture Task. B. Average [*beta*] estimates from the Layout Task cluster displayed in panel A. All parameters are estimated using correct trials only. *$P < 0.05$. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

mentioned above, the only subfield to display a significant main effect of relevant change was area CA1. Figure 4 displays a cluster that is selectively sensitive to relevant change in the Layout Task. This cluster was located in the body of area CA1, and displayed an increased response for task relevant layout matches, with 0 relevant changes being greater than 1 or 2 relevant changes ($t_{(19)} = 3.0$, $P < 0.01$ and $t_{(19)} = 4.6$, $P < 0.001$, respectively). It was not, however, modulated by the degree of irrelevant change in the same task. Interestingly, these voxels also displayed a trend towards greater activity on accurate compared to inaccurate trials during the Layout Task ($t_{(19)} = 1.79$, $P = 0.09$) and a more directed analysis revealed a cluster showing a significant accuracy effect located close to this same region (Supp. Info., Fig. 4). Taken together, these results suggest that this match enhancement signal may be involved in the memory decision and further shows that area CA1 can express both a general and more selective response to change.

We were also able to confirm the general change region using this approach. A cluster that was overlapping with the general change region emerged from this analysis as showing a main effect for relevant change in the Furniture Task. This cluster also displayed a marginal effect for task irrelevant changes in the Furniture Task (linear trend: $F_{(1,19)} = 3.4$, $P = 0.08$) (Supp. Info., Fig. 3), consistent with our interpretation that these voxels signal the overall amount of change in each stimulus.

Finally, anterior-posterior regions of CA1 do appear to show a sensitivity to change in different ways. The relevant change region in posterior CA1, which displayed a match enhancement response was identified while subjects performed the task which probed spatial aspects of the scene, however the anterior region, which showed a general response to all change, was identified while subjects focused on nonspatial, object information in the scene. This anterior–posterior distinction is consistent with previous human work (Pihlajamaki et al., 2004) and is broadly consistent with the rodent literature which has demonstrated that the dorsal hippocampus, corresponding to the human posterior hippocampus, is differentially responsible for spatial memory [see (Fanselow and Dong, 2010) for a recent review]. Although the rodent literature has traditionally treated the ventral hippocampus as being differentially involved in emotional processes (Moser and Moser, 1998), the results of a recent study are in line with the human literature in suggesting that neurons in this region may also code for nonspatial local environmental cues (Royer et al., 2010). Consistent with this interpretation, two clusters in the anterior hippocampus were found to correlate with performance in the Furniture Task.

DISCUSSION

The current study was conducted to test hippocampal area CA1's theoretical role in match/mismatch detection. Although these models have been influential in the conceptualization of hippocampal mechanisms, many of the specific aspects remain to be tested. Thus, it remains a critical goal to assess the validity of CA1's theoretical role in performing a match/mismatch calculation. To this end, we conducted a high-resolution fMRI study during which participants performed a task that establishes an expectation by cueing participants with the name of a previously well-studied room. Next, they were presented with a probe image that either matched the studied room or contained various numbers of changes, some of which were task relevant and others task irrelevant. We predicted that a region supporting a match/mismatch calculation would be sensitive to the total number of changes in each probe image, regardless of their behavioral relevance. Area CA1 was the only hippocampal subfield to reliably display this type of response. Furthermore, we demonstrated that area CA1 had a significantly more linear response to the total number of changes than any other hippocampal region.

The results of this study provide much needed empirical evidence for the theoretically important role of area CA1 as a match/mismatch detector. Previous high-resolution fMRI studies have recorded from hippocampal subfields while old and new stimuli were presented (Zeineh et al., 2003; Eldridge et al., 2005; Kirwan et al., 2007; Kirwan and Stark, 2007; Bakker et al., 2008; Preston et al., 2009; Viskontas et al., 2009; Yassa and Stark, 2009; Dudukovic et al., 2010). These studies did not, however, reveal that area CA1 played any special role in discriminating between old and new stimuli. Two key aspects of our design likely led to the differentiation of area CA1's response from other subfields. First, most of the prior paradigms did not allow for a clear expectation to develop on each trial. Instead, previous studies presented new objects (Eldridge et al., 2005; Kirwan and Stark, 2007; Viskontas et al., 2009), faces (Zeineh et al., 2003; Kirwan and Stark, 2007; Preston et al., 2009), scenes (Preston et al., 2009), or words (Zeineh et al., 2003; Eldridge et al., 2005; Viskontas et al., 2009), either in a continuous stream or intermixed with old items. Although these studies were well designed to answer questions about encoding or retrieval success, they were not designed to detect a signal related to how well a presented stimulus matches or mismatches an expectation since there was nothing in the stimulus, or preceding the stimulus, to set up an expectation. There is one recent study, however, that did establish an expectation (Dudukovic et al., 2010) but they also did not find a special role for area CA1 in discriminating between novel and repeated probes. Instead, they found that all subfields displayed an enhanced response for probes that matched a maintained image. Unlike our paradigm, the mismatching probes used in this study were completely novel, and so could not serve as a memory cue. This means that their results are equally consistent with a signal that reflects the reactivation of recent experience or a match/mismatch calculation. Indeed, the authors interpreted their findings as evidence for the former. By including the development of an expectation and by always using probes that could serve as retrieval cues, our design allowed us to demonstrate that area CA1, more so than any other subfield, generates responses that are consistent with match/mismatch detection.

The match/mismatch signals observed in the current study not only distinguished between images that did or did not con-

tain change, but also represented the amount of change in a graded, roughly linear, manner. This is consistent with rodent electrophysiological data acquired while the animals explored environments that were manipulated versions of recently explored environments. The results of these studies suggest that, as compared with areas DG and CA3, neurons in area CA1 represent change in a linear manner (Guzowski et al., 2004, Lee et al., 2004; Leutgeb et al., 2004; Vasdarjanova and Guzowski, 2004). One additional study recorded from area CA1 while rats found an escape platform in an unexpected location (Fyhn et al., 2002), a manipulation more analogous to those performed in the current study. They found that area CA1 shifted its representation of the unexpected location by recruiting more pyramidal neurons while decreasing the response in a population of interneurons. While these studies have primarily measured shifts in neuronal representations to characterize hippocampal subfields' sensitivity to change, there is some evidence that these representational changes are accompanied by changes in neuronal firing rate (Leutgeb et al., 2005), which would be more readily translated into the BOLD response.

It is noteworthy that all previous fMRI studies that have found evidence for hippocampal match/mismatch signals have used tasks with a working memory structure (Kumaran and Maguire, 2006, 2007a; Hannula and Ranganath, 2008; Duncan et al., 2009; Dudukovic et al., 2010) where a matching or mismatching probe is presented within seconds of a cue stimulus. By contrast, the current study asks subjects to make a match/mismatch judgment by comparing a retrieved long-term memory representation, one well-learned across preceding days, to a currently presented stimulus. The hippocampus is proposed to continuously make comparisons based on previously encoded events, so arguably a long-term memory paradigm may be more representative of the situations in which the hippocampus acts as comparator than traditional working memory paradigms. Additionally, by extending the evidence for match/mismatch signals to a long-term memory task, we were able to demonstrate that these responses are not tied to the structure of working memory paradigms but are, rather, more likely to reflect the comparison of past experience with the present.

The theoretical importance of the match/mismatch signal lies in its proposed ability to adaptively direct the hippocampus towards the mnemonic processes that are best suited to the current situation (Hasselmo et al., 1994, 1995; Lorincz and Buzaki, 2000; Lisman and Otmakhova, 2001; Vinogradova, 2001; Meeter et al., 2004; Lisman and Grace, 2005; Kumaran and Maguire, 2007b). Some evidence for the consequences of this detection have been found in electrophysiological recordings in humans which reported an early automatic hippocampal response followed by a later more task-related response (Axmacher et al., 2010). Similarly, in addition to the automatic total change response reported here, we also identified a cluster in area CA1 whose response was only sensitive to task relevant manipulations and was also correlated with task performance. This more task-related signal may correspond to the memory processing that occurs subsequent to the automatic match/mismatch detection. Although fMRI does not allow us to determine the relative timing of these signals, we were able to add a critical piece to the emerging conceptualization of hippocampal novelty detection and memory by isolating these signals to area CA1.

This study also serves as an example of how high-resolution fMRI can contribute unique evidence for theoretical models of hippocampal function (Carr et al., 2010). Although this method is relatively new, it has already served an important role in testing other aspects of hippocampal models (Bakker et al., 2008). This approach gives us a window into the intact hippocampus, allowing us to observe the operations performed within a subfield when all efferent and afferent connections are intact. This is critical when trying to distinguish between distinct operations performed by highly interconnected subfields. Furthermore, the ability to study these processes in humans not only provides evidence directly from the system which we are trying to model, but also allows us to more readily control intentional factors, such as goals and motivations, which are hypothesized to play a key role in modulating hippocampal processing (Lisman and Grace, 2005).

REFERENCES

Axmacher N, Cohen MX, Fell J, Haupt S, Dumpelmann M, Elger CE, Schlaepfer TE, Lenartz D, Sturm V, Ranganath C. 2010. Intracranial EEG correlates of expectancy and memory formation in the human hippocampus and nucleus accumbens. Neuron 65:541–549.

Bakker A, Kirwan CB, Miller M, Stark CE. 2008. Pattern separation in the human hippocampal CA3 and dentate gyrus. Science 319: 1640–1642.

Carr VA, Rissman J, Wagner AD. 2010. Imaging the human medial temporal lobe with high-resolution fMRI. Neuron 65:298–308.

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

Davachi L. 2006. Item, context and relational episodic encoding in humans. Curr Opin Neurobiol 16:693–700.

Dudukovic NM, Preston AR, Archie JJ, Glover GH, Wagner AD. 2011. High-resolution fMRI reveals match enhancement and attentional modulation in the human medial temporal lobe. J Cogn Neurosci 23:670–682.

Duncan K, Curtis C, Davachi L. 2009. Distinct memory signatures in the hippocampus: Intentional states distinguish match and mismatch enhancement signals. J Neurosci 29:131–139.

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

Eldridge LL, Engel SA, Zeineh MM, Bookheimer SY, Knowlton BJ. 2005. A dissociation of encoding and retrieval processes in the human hippocampus. J Neurosci 25:3280–3286.

Fanselow MS, Dong HW. 2010. Are the dorsal and ventral hippocampus functionally distinct structures? Neuron 65:7–19.

Fyhn M, Molden S, Hollup S, Moser MB, Moser E. 2002. Hippocampal neurons responding to first-time dislocation of a target object. Neuron 35:555–566.

Guzowski JF, Knierim JJ, Moser EI. 2004. Ensemble dynamics of hippocampal regions CA3 and CA1. Neuron 44:581–584.

Hannula DE, Ranganath C. 2008. Medial temporal lobe activity predicts successful relational memory binding. J Neurosci 28:116–124.

Hasselmo ME, Schnell E. 1994. Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region CA: Computational modeling and brain slice physiology. J Neurosci 14:3898–3914.

Hasselmo ME, Schnell E, Barkai E. 1995. Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. J Neurosci 15:5249–5262.

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, Stark CE. 2007. Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. Learn Mem 14:625–633.

Kohler S, Danckert S, Gati JS, Menon RS. 2005. Novelty responses to relational and non-relational information in the hippocampus and the parahippocampal region: A comparison based on event-related fMRI. Hippocampus 15:763–774.

Kumaran D, Maguire EA. 2006. An unexpected sequence of events: Mismatch detection in the human hippocampus. PLoS Biol 4:e424

Kumaran D, Maguire EA. 2007a. Match mismatch processes underlie human hippocampal responses to associative novelty. J Neurosci 27:8517–8524.

Kumaran D, Maguire EA. 2007b. Which computational mechanisms operate in the hippocampus during novelty detection? Hippocampus 17:735–748.

Lee I, Yoganaasimha D, Ra G, Knierim JJ. 2004. Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. Nature 430:456–459.

Lee I, Hunsaker MR, Kesner RP. 2005. The role of hippocampal subregions in detecting spatial novelty. Behav Neurosci 119:145–153.

Leutgeb S, Leutgeb J, Treves A, Moser MB, Moser E. 2004. Distinct ensemble codes in hippocampal areas CA3 and CA1. Science 305:1295–1298.

Leutgeb S, Leutgeb J, Barnes CA, Moser EI, McNaughton BL, Moser MB. 2005. Independent codes for spatial and episodic memory in the hippocampal neuronal ensembles. Science 309:619–623.

Lisman JE, Otmakhova NA. 2001. Storage, recall, and novelty detection of sequences by the hippocampus: elaborating on the SOCRATIC model to account for normal and aberrant effects of dopamine. Hippocampus 11:551–568.

Lisman JE, Grace AA. 2005. The hippocampal-VTA loop: Controlling the entry of information into long-term memory. Neuron 46:703–713.

Lorincz A, Buzaki G. 2000. Two-phase computational model training long-term memories in the entorhinal-hippocampal region. Ann NY Acad Sci 911:83–111.

Mayes A, Montaldi D, Migo E. 2007. Associative memory and the medial temporal lobes. Trends Cogn Sci 11:126–135.

Meeter M, Murre JMJ, Talamini LM. 2004. Mode shifting between storage and recall based on novelty detection in oscillating hippocampal circuits. Hippocampus 14:722–741.

Miller EK, Desimone R. 1994. Parallel neuronal mechanisms for short-term memory. Science 263:520–522.

Moser MB, Moser EI. 1998. Functional differentiation in the hippocampus. Hippocampus 8:608–619.

Naber PA, Witter MP, Lopes Silva FH. 2000. Networks of the hippocampal memory system of the rat. The pivotal role of the subiculum. Ann NY Acad Sci 911:392–403.

Norman KA, O'Reilly RC. 2003. Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. Psychol Rev 110:611–646.

Olman CA, Davachi L, Inati S. 2009. Distortion and signal loss in the medial temporal lobe. PLoS ONE 4:e8160

Preston AR, Bornstein AM, Hutchinson JB, Gaare ME, Glover GH, Wagner AD. 2009. High-resolution fMRI of content-sensitive subsequent memory responses in human medial temporal lobe. J Cogn Neurosci 22:156–173.

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 USA 98:12760–12766.

Tulving E. 1987. Multiple memory systems and consciousness. Hum Neurobiol 6:67–80.

Vazdarjanova A, Guzowski JF. 2004. Differences in hippocampal neuronal population responses to modifications of an environmental context: Evidence of distinct, yet complementary, functions of CA3 and CA1 ensembles. J Neurosci 24:6489–6496.

Vinogradova OS. 2001. Hippocampus as comparator: Role of the two input and two output systems of the hippocampus in selection and registration of information. Hippocampus 11:578–598.

Viskontas IV, Carr VA, Engel SA, Knowlton BJ. 2009. The neural correlates of recollection: Hippocampal activation declines as episodic memory fades. Hippocampus 19:265–272.

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.

Zeineh MM, Engel SA, Thompson PM, Bookheimer SY. 2003. Dynamics of the hippocampus during encoding and retrieval of face-name pairs. Science 299:577–580.