Regional Analysis of Hippocampal Activation During Memory Encoding and Retrieval: fMRI Study

Michael D. Greicius,1* Ben Krasnow,1 Jesse M. Boyett-Anderson,1 Stephan Eliez,1 Alan F. Schatzberg,1 Allan L. Reiss,1,2,3 and Vinod Menon1,2,3

1Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine, Stanford, California
2Program in Neurosciences, Stanford University School of Medicine, Stanford, California
3Stanford Brain Research Center, Stanford University School of Medicine, Stanford, California

ABSTRACT: Investigators have recently begun to examine the differential role of subregions of the hippocampus in episodic memory. Two distinct models have gained prominence in the field. One model, outlined by Moser and Moser (Hippocampus 1998;8:608–619), based mainly on animal studies, has proposed that episodic memory is subserved by the posterior two-thirds of the hippocampus alone. A second model, derived by Lepage et al. (Hippocampus 1998;8:313–322) from their review of 52 PET studies, has suggested that the anterior hippocampus is activated by memory encoding while the posterior hippocampus is activated by memory retrieval. Functional magnetic resonance imaging (fMRI) studies have tended to show limited activation in the anteriormost regions of the hippocampus, providing support for the Moser and Moser model. A potential confounding factor in these fMRI studies, however, is that susceptibility artifact may differentially reduce signal in the anterior versus the posterior hippocampus. In the present study, we examined activation differences between hippocampal subregions during encoding and retrieval of words and interpreted our findings within the context of these two models. We also examined the extent to which susceptibility artifact affects the analysis and interpretation of hippocampal activation by demonstrating its differential effect on the anterior versus the posterior hippocampus. Both voxel-by-voxel and region-of-interest analyses were conducted, allowing us to quantify differences between the anterior and posterior aspects of the hippocampus. We detected significant hippocampal activation in both the encoding and retrieval conditions. Our data do not provide evidence for regional anatomic differences in activation between encoding and retrieval. The data do suggest that, even after accounting for susceptibility artifact, both encoding and retrieval of verbal stimuli activate the middle and posterior hippocampus more strongly than the anterior hippocampus. Finally, this study is the first to quantify the effects of susceptibility-induced signal loss on hippocampal activation and suggests that this artifact has significantly biased the interpretation of earlier fMRI studies. Hippocampus 2003;13:164–174. © 2003 Wiley-Liss, Inc.

KEY WORDS: hippocampus, functional imaging, memory, susceptibility artifact

INTRODUCTION

The hippocampus has been the focus of memory research since the seminal neuropsychological studies of patient H.M., who suffered amnesia after a bilateral medial temporal lobe (MTL) resection (Scoville and Milner, 1957). More recently, investigators have begun to explore functional differences between different subregions of the hippocampus. The elucidation of functional specialization within the hippocampus would be critical to a more thorough understanding of the neuroanatomic basis of memory. In addition, this information could have significant clinical implications for disorders involving hippocampal dysfunction (Small et al., 2000). Two separate models describing functional specialization between subregions of the hippocampus have gained prominence in the field.

Moser and Moser (1998) developed a model, based mainly on animal studies, proposing that the anterior third of the hippocampus is functionally distinct from the posterior two-thirds. This hypothesis is derived primarily from observations that the afferent and efferent connections of the anterior (ventral) third of the hippocampus are largely distinct from the connections of the more posterior (dorsal) region. In particular, the anterior...
third has robust efferent connections to the rostral hypothalamus and amygdala (Canteras and Swanson, 1992; Risold and Swanson, 1996). Furthermore, studies of spatial learning in rats (E.I. Moser et al., 1993; M-B. Moser et al., 1995) and monkeys (Colombo et al., 1998) have suggested that the anterior third of the hippocampus may not be required for visuospatial memory. Moser and Moser then adapted this model to the human hippocampus, citing several functional magnetic resonance imaging (fMRI) studies that detected predominantly posterior hippocampal activation during verbal (Fernandez et al., 1998) and visual (Stern et al., 1996; Rombouts et al., 1997) memory tasks. In sum, the Moser and Moser model proposes that the anterior third of the hippocampus is not integral to episodic memory.

The second model, based on a meta-analysis of 52 PET studies by Lepage et al. (1998), holds that the anterior hippocampus is activated by encoding and the posterior hippocampus by retrieval. Lepage’s hippocampus in encoding and retrieval (HIPER) model was derived by comparing the activation foci of 22 encoding studies and 32 retrieval studies on a sagittal hippocampal slice from the Talairach and Tournoux atlas (1988). These investigators concluded that the physiological basis for such a distinction was unclear and encouraged further studies that could prospectively test the model. The HIPER model was refuted by Schacter and Wagner (1999) in their review of the functional imaging literature. In this updated meta-analysis, positron emission tomography (PET) studies tended to show anterior and posterior activation with encoding and predominantly posterior activation with retrieval, whereas fMRI studies showed a predominance of posterior MTL activation for encoding. However, there were insufficient fMRI data on hippocampal activation during retrieval to detect a distinct anatomic pattern. Schacter and Wagner (1999) concluded that the limited anterior activation in the MTL studies was due to differences in task design between the PET and fMRI studies. In particular, they posited that tasks involving relational processing such as word pairs (Dolan and Fletcher, 1999) or pictures of people combined with pictures of housing (Henke et al., 1997) might account for the anterior hippocampal activation seen with encoding in some PET studies. A second criticism leveled at the HIPER model was that most studies examined either encoding or retrieval, while few used a within-subjects design to compare encoding and retrieval directly.

Schacter et al. (1999) subsequently carried out a PET study to test the HIPER model prospectively. Efforts were made to control for potential confounding factors by using a within-subjects design and nonrelational stimuli. The analysis was not limited to the hippocampus but showed predominantly posterior MTL activation, including posterior hippocampus, parahippocampal gyrus, and fusiform gyrus for both encoding and retrieval. Direct comparisons of encoding and retrieval also were carried out. The contrast comparing encoding versus retrieval showed posterior MTL activation, while the opposite comparison, retrieval versus encoding, showed no MTL activation. Although this study attempted to clarify the issue of anatomic localization of MTL activation associated with episodic memory, the design did not include a comparison of activation between anatomically specified hippocampal subregions. Furthermore, the group-averaged, voxel-by-voxel analysis employed might eliminate MTL activation occurring in voxels that only a few subjects activated. In addition, the use of PET, in this and most other studies, may not have provided adequate spatial resolution (10 mm in-plane) to examine subregions within the hippocampus, a structure that extends ~40 mm along its anterior-posterior axis (Duvernoy, 1998). Thus, a great deal of uncertainty remains regarding functional distinctions within the hippocampus proper. One potential source of discrepant findings between fMRI studies as well as between fMRI and PET studies pertains to susceptibility artifact (Veltman et al., 2000). This artifact, found only in fMRI, results from abrupt changes in magnetic susceptibility that occur across tissue interfaces such as the border between air-filled sinuses and brain parenchyma or between bone and brain parenchyma. Brain regions closest to such borders are especially susceptible to loss of blood oxygen level-dependent (BOLD) signal due to this artifact. Ojemann et al. (1997) demonstrated that signal attenuation was most prominent in the inferior frontal and inferior temporal regions. On the basis of these findings, Schacter and Wagner (1999) suggested that reduced anterior hippocampal activation in fMRI studies was not due to susceptibility artifact. It is important to note, however, that Ojemann et al. (1997) specified that signal loss was “relative rather than absolute” and decreased with distance away from air–brain or bone–brain borders, so that other temporal lobe regions may still show signal attenuation, albeit less prominent. Ojemann et al. (1997) also cautioned that “it remains unknown exactly how much a local BOLD activation signal would be affected by a superimposed macroscopic signal loss.” Lipschutz et al. (2000) demonstrated that “sensitivity to BOLD effects is directly proportional to signal intensity,” suggesting that BOLD signal detection can be reduced due to susceptibility artifact, even in the absence of complete signal dropout.

Two recent studies have extended the work of Ojemann et al. to show that signal loss due to susceptibility artifact can have significant effects on MTL activation. Devlin et al. (2000) compared PET and fMRI directly, using a nearly identical semantic task, and found that anteromedial temporal activation was detected with PET, but not with fMRI. Cordes et al. (2000) showed susceptibility-induced signal loss in the parahippocampal and amygdala regions while a subject mentally rehearsed a gymnastics routine. Because the hippocampus rises superiorly from anterior to posterior, one would expect greater susceptibility-induced signal loss in the anterior (inferior) relative to the posterior (superior) hippocampus. It is plausible therefore that the relative lack of anterior hippocampal activation and the preponderance of posterior hippocampal activation in fMRI studies of memory is due to this pervasive artifact. No study to date, however, has directly examined the effects of susceptibility artifact on hippocampal activation.

In this study, we investigated regional differences in fMRI activation of the hippocampus during encoding and retrieval. In particular, data were examined within the context of the Moser and Moser and Lepage models discussed above. This study includes several methodological improvements in comparison with previous studies. A within-subjects design was used to examine encoding and retrieval of nonrelational stimuli. The same control condition was used across the encoding and retrieval experiments, to facilitate a direct comparison. In addition to the standard voxel-
by-voxel analysis, we report a more specific, quantitative region-of-interest (ROI) analysis focusing on statistical comparisons between hippocampal subregions. We have previously used a similar approach to investigate regional differences in novelty, memory, and spatial processing within MTL structures (Menon et al., 2000). To assess Lepage’s HIPER model, in which PET activations of encoding mapped to the anterior half of the hippocampus and retrieval activation to the posterior half, we divided our hippocampal ROIs into anterior and posterior halves. To assess the Moser and Moser model, positing a functional difference between the anterior third and the posterior two-thirds of the hippocampus, we divided our hippocampal ROIs into anterior, middle, and posterior thirds. Finally, and perhaps most importantly, we include an analysis of susceptibility artifact and its effect on BOLD signal detection in the hippocampus.

MATERIALS AND METHODS

Subjects

Fourteen healthy, right-handed subjects (six males and eight females, aged 18–48 years) participated in the study after giving informed consent. Two subjects were eliminated before fMRI analysis because their behavioral data indicated that they performed at or below chance on either the encoding or the retrieval task. Analysis of fMRI data was performed on the remaining 12 subjects (five males and seven females, aged 18–48 years, mean age 26 ± 8).

Stimuli

The stimuli for the encoding condition were 40 unique visually presented nouns. In the retrieval condition, the stimuli were 32 of the same words plus 16 new words. The stimuli for the control condition consisted of the same 2 nouns alternating repeatedly.

Design and Procedure

The encoding task consisted of epochs of eight words each presented independently for 2.5 s, with a 0.5-s interstimulus interval. Each of the five encoding epochs was followed by a control epoch in which either of two words was presented in a random order over the eight stimulus periods. The same two words were used for all five control epochs. Each epoch was preceded by an instruction screen shown for 4 s. In addition, a 24-s rest period with fixation occurred before the onset of the first cycle, at the midway point (after the third encoding epoch), and after the last control epoch. The task order for retrieval can be abbreviated as F-R-C-R-C-F-R-C-R-C-R-C-F, again, where F is fixation, R is retrieval, and C is control. All 12 subjects performed the identical retrieval task.

The first five subjects underwent a slightly modified version of the encoding protocol that differed only in two ways: (1) the middle fixation epoch was placed after the third control epoch, rather than before it (F-E-C-E-C-E-C-F); and (2) they were not shown an instruction screen at the beginning of each epoch, but received verbal instructions at the beginning of the protocol.

Stimulus Presentation

The tasks were programmed using Psyscope (http://poppy.psy.cmu.edu/psyscope) on a Macintosh (Sunnyvale, CA) notebook computer. Onset of scanning and task were synchronized using a TTL pulse delivered to the scanner timing microprocessor board from a CMU Button Box microprocessor connected to the Macintosh with a serial cable. Stimuli were presented visually at the center of a screen using a custom-built magnet compatible projection system (Resonance Technology, CA).

fMRI Acquisition

Images were acquired on a 3-tesla (T) GE Signa scanner using a standard GE whole head coil. The scanner runs on an LX platform, with gradients in Mini-CRM configuration (35 mT/m, SR 200 mT/m/s), and has a MagneX 3-T 80-cm magnet. A custom-built head holder was used to prevent head movement; 28 axial slices (4 mm thick, 0.5 mm skip) parallel to the ACPC line and covering the whole brain were imaged with a temporal resolution of 2 s, using a T2*-weighted gradient echo spiral pulse sequence (TR = 2,000 ms, TE = 30 ms, flip angle = 89° and 1 interleave) (Glover and Lai, 1998). The field of view was 200 × 200 mm², and the matrix size was 64 × 64, giving an in-plane spatial resolution of 3.125 mm. To reduce blurring and signal loss arising from field inhomogeneities, a shimming protocol was used before acquiring functional MRI scans (Spielman et al., 1998). The protocol uses a short spiral acquisition to obtain the field maps and downloads the resistive shims automatically. To aid in lo-
calization of functional data, a high-resolution T1-weighted spoiled grass gradient recalled (SPGR) 3D MRI sequence with the following parameters was used: TR = 35 ms; TE = 6 ms; flip angle = 45°; 24-cm field of view; 124 slices in the sagittal plane; 256 × 192 matrix; acquired resolution = 1.5 × 0.9 × 1.2 mm. The images were reconstructed as a 124 × 256 × 256 matrix with a 1.5 × 0.9 × 0.9-mm spatial resolution. Structural and functional images were acquired in the same scan session.

Region of Interest

Our hippocampal ROI was based on the protocol outlined by Kates et al. (1997). Each subject’s structural MRI was first normalized in standard Talairach space. Using the normalized structural image allowed for a more reliable comparison with the normalized functional data minimizing, for example, the effects of changes in head position between the structural and functional scans. The hippocampus was delineated on coronal slices strictly perpendicular to the anterior commissure–posterior commissure (AC–PC) axis. The anterior slice of the hippocampus is most easily defined by the superior shift of the temporal horn of the lateral ventricles to the point at which the temporal horn turns and points superiorly and is positioned so that the amygdala is located superior and the hippocampus inferior to it. Posteriorly, the hippocampus fuses with the fornix and is measured until it is no longer visible. The medial border is defined by the ambient cistern. The inferior border is marked by the collateral white matter, the subiculum, and the parahippocampal gyrus. The lateral border is marked by the temporal horn of the lateral ventricles and more superiorly by the white matter. To test the HIPER model, each individual’s hippocampal ROIs were then split into anterior and posterior hippocampal ROIs. This was done by dividing the full ROI at its midpoint along the anterior-posterior axis. To test the Moser and Moser model, each full hippocampal ROI was divided into thirds along the anterior-posterior axis. Thus, in the ROI analysis of the HIPER model each individual subject had four hippocampal ROIs: left anterior, right anterior, left posterior, and right posterior. For the Moser and Moser model analysis, each individual had six hippocampal ROIs: left anterior, left middle, left posterior, right anterior, right middle, and right posterior.

In addition to the individual ROIs, we used the same protocol to draw bilateral hippocampal ROIs on a normalized brain averaged from all 12 subjects. The left hippocampal ROI had an anterior-posterior extent of 32 mm beginning at a Talairach y-coordinate of −8 and ending at −40 with a midpoint of −24. The right hippocampal ROI had an anterior-posterior extent of 34 mm, beginning at −6 and extending to −40 with a midpoint of −23.

Data Analysis

fMRI data from each subject were analyzed using Statistical Parametric Mapping (SPM99) (http://www.fil.ion.ucl.ac.uk/spm). Before statistical analysis, images were corrected for movement using least-squares minimization without higher-order corrections for spin history, normalized to stereotactic Talairach coordinates (Talairach and Tournoux, 1988), resampled every 2 mm using sinc interpolation and smoothed with a 4-mm Gaussian kernel to eliminate spatial noise. For each subject, voxel-wise activation during each experimental condition compared with the control condition was determined, using multivariate regression analysis with correction for temporal autocorrelations in the fMRI data (Friston et al., 1995). Confounding effects of fluctuations in global mean were removed by proportional scaling, and low-frequency noise was removed with a high-pass filter (0.5 cpm). A regressor waveform for each condition, convolved with a 6-s delay Poisson function accounting for delay and dispersion in the hemodynamic response, was used to compute voxel-wise t-statistics, which were then normalized to z-scores to provide a statistical measure of activation independent of sample size. Using random-effects analysis (Holmes and Friston, 1998; Friston et al., 1999a), brain activation for each of the experimental conditions, contrasted with the control condition, was determined: (1) encoding, (2) retrieval. These activation images were then directly compared. Given the hypothesis-driven nature of our study, we used a threshold of z > 1.67 (P < 0.05) without spatial correction (i.e., no minimal cluster size) to identify significantly activated voxels. The right and left hippocampal ROIs from the group brain were used as a mask on the group-averaged activation image, generating a map of activation within the hippocampus only. This masked activation image was then superimposed on the group-averaged structural image.

For the HIPER model, we calculated the percentage of voxels activated (P < 0.05 without spatial correction) in each individual’s four hippocampal ROIs. While both the percentage of voxels activated (Gabrieli et al., 1997) and the absolute number of voxels activated (Small et al., 2001) have been used as measures in hippocampal activation studies, a statistical comparison of the two ROI techniques (Constable et al., 1998; Fig. 6, p 297) has shown that measuring the percentage of voxels activated is the more robust approach. These data were then entered into an analysis of variance (ANOVA) to examine the effects of three factors on hippocampal activation: the three-way ANOVA examined task (encoding/retrieval) × anterior-posterior (A/P) location × hemisphere (left/right) interactions. A similar analysis was performed with the hippocampal ROIs divided into thirds (anterior, middle, and posterior), to examine the Moser and Moser model.

Artifact Quantification

To quantify the amount of signal loss due to susceptibility artifact, we first calculated the average voxel intensity for each individual’s four ROIs from a T2* image averaged over the fixation epochs in the retrieval paradigm. We chose the fixation epochs to ensure that differences in voxel intensity were not due to activation during the experimental task, but rather to baseline differences in T2* signal. These average voxel intensities were used to examine A/P location by Hemisphere interactions among all 12 subjects using a two-way ANOVA. Then, for each subject, we graphed the intensity of every hippocampal voxel against its subsequent activation t-score taken over the averaged retrieval epochs (Fig. 1). Each subject had an individual voxel intensity threshold (ranging across the group from 432 to 652, arbitrary scale), below which there was no activation but above which all voxels were equally likely to be activated (i.e., there was no correlation between voxel value and t-score). We refer to this threshold as intensity threshold for detection of activation (ITDA). The ITDA results from a threshold masking procedure applied by SPM to remove “non-brain” voxels.
In this procedure, all voxels below a given T2* signal intensity are assigned a \( t \)-score of 0 in the subsequent statistical analysis. The threshold masking function is determined using each individual’s mean T2* signal intensity (rather than a fixed T2* intensity across all individuals), accounting for the variance in the ITDA values across individuals. It should be noted that threshold masking is a widely used procedure to remove “non-brain” voxels, employed not only by SPM, but by alternative fMRI software programs, such as analysis of functional neuroimages (AFNI), as well.

We then compared the percentage of voxels exceeding the ITDA in each of the four ROIs, using another two-way ANOVA examining A/P location by hemisphere interactions. Finally, to determine the effect of signal loss on the percentage of voxels activated, we performed the ROI analyses again, this time restricting them to voxels that exceeded the ITDA calculated for each individual.

**RESULTS**

**Behavioral Task Performance**

Behavioral data demonstrated that 12 of the 14 subjects performed both the encoding and retrieval tasks correctly (i.e., better than chance). One subject responded correctly to only 45% of the
encoding judgments. A separate subject responded correctly to only 42% of the retrieval judgments. Both subjects were excluded from fMRI analysis, as we could not be certain that they were performing the encoding and retrieval tasks correctly. The remaining 12 subjects had accuracy rates on the encoding semantic judgment (man-made or not man-made) ranging from 87.5% to 100%, with a mean of 95% and an SD of 4.1%. In the retrieval task, accuracy rates on the “new” versus “old” judgment (combining correctly identified old words and correctly rejected foils) ranged from 79.2% to 91.6% correct, with a mean of 85.4% and an SD of 4.7%. There were no significant differences in the encoding or retrieval task performances between the five subjects who underwent the slightly modified encoding protocol (see Materials and Methods) and the remaining seven subjects.

**Hippocampal Activation**

Figures 2 and 3 show group hippocampal activation for two comparisons: encoding versus the control condition and retrieval versus the control condition. During encoding, activation was seen bilaterally both anterior and posterior to the midpoint of the hippocampus (y = −23), although the bulk of the activated voxels appears to be at or just posterior to the midpoint (Fig. 2). During retrieval, activation was observed mainly on the right side, both anterior and posterior to the midpoint with a slight posterior predominance (Fig. 3).

We then performed two additional group comparisons: encoding versus retrieval and retrieval versus encoding. The encoding
versus retrieval comparison activated only 29 hippocampal voxels (of 909 total voxels in the bilateral hippocampi) compared with 300 voxels which were activated when the encoding and control conditions were compared. These 29 voxels were spread diffusely across the anterior and posterior regions of the hippocampus, mainly in the left hemisphere. The retrieval versus encoding comparison activated only six voxels compared with the 106 voxels, which were activated when the retrieval and control conditions were compared. These six voxels were near the midpoint of the hippocampus in the right hemisphere.

Artifact Quantification

Figure 4 illustrates the effect of susceptibility-induced voxel intensity attenuation on BOLD signal detection in the hippocampus. To quantify differences in voxel intensity between hippocampal subregions, we performed a two-way ANOVA examining the effects of A/P location and hemisphere on the average voxel intensity (during fixation) of each of four ROIs (left anterior, right anterior, left posterior, right posterior). This analysis showed a significant main effect of A/P location (F (1,11) = 28.89, P = 0.0002)—the anterior ROIs had significantly lower average voxel intensities than the posterior ROIs. The main effect of hemisphere showed a trend towards significance (F (1,11) = 4.73, P = 0.052), in which the left hippocampus had lower average voxel intensities than the right. There was no significant interaction between A/P location and hemisphere (F (1,11) = 0.39, P = 0.55).

We then examined the effect of A/P location and hemisphere on our ability to detect BOLD activation. We performed a similar two-way ANOVA (A/P location × hemisphere), using the percentage of voxels above the ITDA as the dependent variable (see Materials and Methods; see also Fig. 1). There was a main effect of A/P location (F (1,11) = 10.31, P = 0.008). The posterior ROIs had significantly more voxels exceeding the ITDA than were found for the anterior ROIs. We did not detect a main effect of hemisphere (F (1,11) = 1.95, P = 0.19) or an interaction between A/P location and hemisphere (F (1,11) = 2.05, P = 0.18).

ROI Analysis

HIPER model

The HIPER model was tested by dividing each hippocampal ROI into half. For each of the four ROIs (left anterior, right anterior, left posterior, right posterior), we calculated the percentage of activated voxels. These percentages were then used as the dependent variable in a three-way ANOVA that examined task × A/P location × hemisphere interactions. There were no main effects or interactions (Table 1). When the analysis was limited to artifact-free voxels whose intensity exceeded the ITDA, there were still no main effects nor any interactions (Table 1).

Moser and Moser model

The Moser and Moser model was tested by dividing each hippocampal ROI into thirds. We then performed a three-way ANOVA (task × A/P location × hemisphere) using three A/P locations (anterior, middle, and posterior). There were no main effects or interactions among the three factors (Table 2). When the analysis was limited to artifact-free voxels whose intensity exceeded the ITDA, the ANOVA still did not yield any significant main effects or interactions (Table 2).

Based on our voxel-by-voxel analysis, which demonstrated a preponderance of activation in the middle third of the hippocampus, moderate activation in the posterior third, and lowest activation in the anterior third, we hypothesized that activation in the posterior third might be diluting a significant difference in activation between the anterior and middle thirds when all three A/P locations were combined in the ANOVA. We therefore performed a post-hoc three-way ANOVA (task × A/P location × hemisphere) comparing only the anterior and middle thirds. This analysis showed a significant main effect for A/P location, where the middle third had 21% of voxels activated and the anterior third had 16% of voxels activated (F (1,11) = 7.73, P = 0.018, η² = 0.89). There were no other main effects and no interactions (Table 3). When the ANOVA was limited to artifact-free voxels whose intensity exceeded the

<table>
<thead>
<tr>
<th>Effect</th>
<th>Uncorrected ROIs</th>
<th>Corrected ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td>DF</td>
<td>F</td>
</tr>
<tr>
<td>A/P location</td>
<td>1,11</td>
<td>0.05</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1,11</td>
<td>2.02</td>
</tr>
<tr>
<td>Task × A/P location</td>
<td>1,11</td>
<td>0.1</td>
</tr>
<tr>
<td>Task × hemisphere</td>
<td>1,11</td>
<td>0.54</td>
</tr>
<tr>
<td>A/P location × hemisphere</td>
<td>1,11</td>
<td>1.39</td>
</tr>
<tr>
<td>Task × A/P location × hemisphere</td>
<td>1,11</td>
<td>0.01</td>
</tr>
</tbody>
</table>

ANOVA, analysis of variance; A/P, anterior-posterior; ROI, region of interest.

*ANOVA was run with uncorrected ROIs and then repeated with ROIs corrected for susceptibility-induced signal loss.
ITDA, the percentage of activated voxels in the anterior third increased slightly to 17%, while the percentage of activated voxels in the middle third remained essentially unchanged at 21.2%. The main effect of A/P location was reduced to a non-significant trend ($F(1,11) = 4.67, P = 0.054, \eta^2 = 0.82$). There were no other main effects or interactions (Table 3).

### DISCUSSION

Our study demonstrates that activation occurs across the full anterior-posterior extent of the hippocampus (defined in this study as the hippocampal formation distinct from entorhinal and parahippocampal cortices) for both encoding and retrieval tasks. The results do not support an anterior-posterior activation differential for encoding versus retrieval and thus, fail to support the HIPER model proposed by Lepage and colleagues. This model was supported by neither voxel-by-voxel analysis nor ROI analysis (performed both with and without a correction for susceptibility artifact). The findings offer minimal support for the Moser and Moser model in that the episodic memory tasks activated the posterior two-thirds of the hippocampus to a slightly greater extent than the anterior third. Finally, and perhaps most importantly, our data suggest that susceptibility artifact has played a significant, confounding role in the interpretation of previous hippocampal activation studies.

In testing the HIPER model, we followed guidelines, suggested by Schacter et al (1999), for minimizing potential confounds in tasks designed to compare encoding and retrieval. In particular, we used a within-subjects design, nonrelational stimuli, and an analysis that included direct comparisons of encoding and retrieval. Our results showed greater anterior hippocampal activation during both encoding and retrieval than was found in the Schacter PET

### TABLE 2.

Results of Three-Way ANOVA With Factors Task (Encoding, Retrieval), A/P Location (Anterior Third, Middle Third, and Posterior Third), and Hemisphere (Left, Right)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Uncorrected ROIs</th>
<th>Corrected ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DF</td>
<td>F</td>
</tr>
<tr>
<td>Task</td>
<td>1,11</td>
<td>0.03</td>
</tr>
<tr>
<td>A/P location</td>
<td>2,22</td>
<td>2.06</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1,11</td>
<td>1.84</td>
</tr>
<tr>
<td>Task × A/P location</td>
<td>2,22</td>
<td>0.79</td>
</tr>
<tr>
<td>Task × hemisphere</td>
<td>1,11</td>
<td>0.46</td>
</tr>
<tr>
<td>A/P location × hemisphere</td>
<td>2,22</td>
<td>2.29</td>
</tr>
<tr>
<td>Task × A/P location × hemisphere</td>
<td>2,22</td>
<td>0.33</td>
</tr>
</tbody>
</table>

ANOVA, analysis of variance; A/P, anterior-posterior; ROI, region of interest.

*The ANOVA was run with uncorrected ROIs and then repeated with ROIs corrected for susceptibility-induced signal loss.

### TABLE 3.

Results of Three-way ANOVA With Factors Task (Encoding, Retrieval), A/P Location (Anterior Third, Middle Third), and Hemisphere (Left, Right)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Uncorrected ROIs</th>
<th>Corrected ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DF</td>
<td>F</td>
</tr>
<tr>
<td>Task</td>
<td>1,11</td>
<td>0.01</td>
</tr>
<tr>
<td>A/P location</td>
<td>1,11</td>
<td>7.73</td>
</tr>
<tr>
<td>Hemisphere</td>
<td>1,11</td>
<td>0.37</td>
</tr>
<tr>
<td>Task × A/P location</td>
<td>1,11</td>
<td>2.1</td>
</tr>
<tr>
<td>Task × hemisphere</td>
<td>1,11</td>
<td>0.17</td>
</tr>
<tr>
<td>A/P location × hemisphere</td>
<td>1,11</td>
<td>0.01</td>
</tr>
<tr>
<td>Task × A/P location × hemisphere</td>
<td>1,11</td>
<td>0.15</td>
</tr>
</tbody>
</table>

ANOVA, analysis of variance; A/P, anterior-posterior; ROI, region of interest.

*The ANOVA was run with uncorrected ROIs and then repeated with ROIs corrected for susceptibility-induced signal loss.

*significant at $P < 0.05$. 
study. One possible explanation for this difference is that the Schacter study used visual stimuli instead of verbal stimuli. Another possibility is that the limited spatial resolution of PET may make fMRI better suited for more precise, anatomical dissections of regional hippocampal activation. Our fMRI study and the Schacter PET study constitute two significant challenges to the HIPER model and demonstrate the importance of a priori testing of hypotheses generated a posteriori from a meta-analysis.

Our data are less conclusive with respect to the Moser and Moser model. We were clearly able to detect activation in the most anterior aspect of the hippocampus during each task, suggesting that this part of the hippocampus is integral to episodic encoding and retrieval of verbal material. However, visual inspection of the voxel-by-voxel T maps (Figs. 2, 3) did suggest a qualitative difference in the degree of activation between the anterior and middle third of the hippocampus. This impression was not borne out by the ROI analysis when all three thirds (anterior, middle and posterior) were compared by ANOVA; however a post-hoc ANOVA comparing the anterior to the middle third indicated marginally significant differences with less activation in the anterior third. There are a few caveats to this finding. First, this result may have been subject to the bias of multiple comparisons, because we compared two factors of an ANOVA directly despite no difference in the initial three-way comparison. We justified this post-hoc comparison based on visual inspection of the voxel-by-voxel analysis, which suggested a difference in the amount of activation between the anterior and middle thirds. Second, the size of the difference between activation in the anterior and middle thirds was modest (16% vs 21%, \( P = 0.018, \eta^2 = 0.89 \)). Finally, this difference was reduced to a statistical trend when we corrected our ROI analysis for susceptibility artifact. If our data support the Moser and Moser model, they suggest the difference between the anteriormost part of the hippocampus and the posterior aspect is a matter of degree rather than kind. Moser and Moser considered this as a possibility in stating “We do not know whether the differentiation of the hippocampus along the longitudinal axis is graded or discontinuous” (p. 613). They favored a discontinuous model of hippocampal differentiation but our data are more supportive of a graded model. We also hasten to point out that the bulk of the Moser and Moser review pertains to visuospatial memory and we cannot generalize beyond verbal memory from our results.

One finding we did not anticipate arose from the voxel-by-voxel analysis, which showed almost no activation in the left hippocampus during retrieval in the block design analysis. This asymmetry was not detected in any of the ANOVA performed in our ROI analysis where we would have expected a task × hemisphere interaction (Tables 1–3). The discrepancy between the voxel-by-voxel and ROI analyses was likely attributable to the fact that the ROI analysis examines activation of individual voxels on a subject-by-subject basis without regard to whether each subject of a group activated exactly the same voxel as other subjects in the group. In contrast, the voxel-by-voxel analysis is more stringent in that it highlights voxels that were activated in common across all or most of the subjects. Thus, the most likely explanation is that in each subject the left hippocampus was as involved in retrieval of verbal material as the right hippocampus, but that across subjects there happened to be greater anatomic variability in location of activated voxels in the left hippocampus. This type of individual variability in activation emphasizes the importance of using an ROI analysis to complement the voxel-by-voxel analysis (Montaldi et al., 1998).

A potential confounding factor in using a block design to investigate retrieval processes involves the inclusion of new words in the retrieval blocks to maintain some trial-to-trial variability. In an event-related fMRI study, Buckner et al. (2001) recently provided evidence of incidental encoding of new words during a recognition test, raising the concern that block designs of retrieval are subject to an incidental encoding confound. This issue was explicitly addressed by Stark and Squire (2000) in a study examining retrieval of nameable objects compared with retrieval of words. In that study, novel words (used as foils in an old-new recognition paradigm) resulted in hippocampal deactivation below baseline. Novel objects showed less hippocampal deactivation than words (but did not show hippocampal activation) and thus tended to blunt hippocampal activation in the old versus new contrast. The investigators concluded that novelty and incidental encoding effects are diminished with verbal stimuli owing to the “high preexperimental familiarity” of words (Stark and Squire, 2000). A study conducted by Kelley et al. (1998) showed a similar discrepancy between the incidental encoding effects of novel words compared with novel objects during passive viewing. In this regard, it is worth noting that the paper by Buckner et al. (2001) showed differential activation during incidental encoding in a number of regions but not in the hippocampus or MTL. Thus, while hippocampal activation to incidental encoding during a retrieval block remains a theoretical concern, it has not been demonstrated in either of the two fMRI studies that have addressed the possibility (Buckner et al, 2001; Stark and Squire 2000). We concur, therefore, with Stark and Squire (2000) that the relative gain in retrieval activation offered by the block design compared with an event-related design (Friston et al., 1999b) offsets the potential confound of incidental encoding, particularly as pertains to verbal stimuli and the hippocampus.

A second potential confound in our encoding task is that the experimental epochs involve both an intentional encoding component and a semantic component (required for the man-made versus not man-made judgment). To facilitate a comparison between the encoding and retrieval tasks, we decided to use an identical baseline across the tasks. The baseline epochs did not require semantic processing; thus, subtracting it from our encoding epochs did not account for the semantic processing. Of the several event-related studies comparing subsequently remembered versus forgotten words, one has shown (Otten et al., 2001) and one has suggested (Kirchhoff et al., 2000) that successful encoding, distinct from semantic processing, activates the hippocampus. The converse—that semantic processing distinct from encoding, activates the hippocampus—has not been demonstrated. The question of whether semantic processing alone activates the hippocampus could be addressed in an event-related study in which the semantic processing versus nonsensory processing contrast was restricted to words that were subsequently forgotten. To our knowledge, such an analysis has not been reported. Based on the literature and given that our analysis was limited to the hippocampus, we hold that the hippocampal activation in our encoding task is mainly a reflection of encoding, though we cannot dismiss the possibility...
that semantic processing also contributed to hippocampal activation.

Our data strongly suggest that susceptibility artifact has been a potential confound in fmRI studies of regional hippocampal activation. We attempted to minimize this artifact by using a specially designed shim technique (Spielman et al., 1998) and a gradient echo spiral pulse sequence rather than a traditional echo planar sequence (G. Glover, personal communication). Given the relative paucity of fMRI studies showing activation in the anterior hippocampus, the presence of any activation in this region suggests we have succeeded in reducing some artifactual signal loss. Nonetheless, the averaged resting T2* voxel intensity in our anterior hippocampal ROIs was significantly less than in our posterior hippocampal ROIs. More importantly, voxel intensity decreases were substantial enough to impair the detectability of BOLD signal changes. We found that a greater percentage of voxels in the anterior hippocampal ROIs fell below the ITDA where BOLD effects could not be detected. After adjusting ROIs for signal loss, the apparent functional distinction between the anterior and middle third of the hippocampus was reduced to a nonsignificant trend. Finally, it should be noted that this adjustment only accounts for signal loss severe enough to place voxels below the ITDA. Lipschutz et al. (2001) showed that sensitivity to BOLD effects is proportional to signal intensity, so that even if, for example, an anterior hippocampal voxel survives the ITDA, it would prove more difficult to activate than a posterior hippocampal voxel with a higher baseline T2* signal. Thus, it is possible that the residual difference, after adjusting for voxels below the ITDA, between the anterior third of the hippocampus and the more posterior regions is still due to susceptibility artifact.

We suspect that many previous fMRI studies could not accurately assess activation differences between hippocampal subregions owing to the effects of susceptibility artifact. Other authors have mentioned the potential role of susceptibility artifact in their hippocampal studies, but to our knowledge none have attempted to quantify it. In their review of the fMRI literature, Schacter and Wagner (1999) displayed data from their own lab (p.21, Fig. 3) showing the anterior MTL to have a signal-to-noise ratio (SNR) of 60 compared with >90 in the posterior MTL. They suggested that this apparently substantial difference was not due to susceptibility artifact because regions such as the temporal pole that “demonstrate considerable susceptibility-induced signal loss” had SNRs closer to 20. On the contrary, we suggest that their data complement the findings of Ojemann et al. (1997), who described a gradient of susceptibility artifact moving away from the skull base and sinuses. It follows that the temporal pole would have the most artifact, the anterior MTL a moderate amount, and the posterior MTL the least, a pattern that matches the SNRs reported by Schacter and Wagner (1999). A laudable but incomplete attempt at accounting for susceptibility artifact was described in a review by Stern and Hasselmo (1999). Stern’s landmark 1996 experiment, which demonstrated posterior hippocampal activation to picture encoding, was performed on a 1.5-T scanner (Stern et al., 1996). This experiment was repeated by the same laboratory in 1998 while attempting to increase SNR by adding within-subject signal averaging and scanning with a 3-T magnet (Kirchhoff et al., 1998). Although the primary finding of posterior hippocampal activation was replicated, this study further obscured the central issue of susceptibility artifact which increases rather than decreases with increasing magnetic field strength (Abduljalil and Robitaille, 1999).

In conclusion, our data suggest that the hippocampus functions as a rather homogeneous unit during encoding and retrieval of episodic verbal memories. The relatively decreased anterior activation, apparent in our voxel-by-voxel analysis, was less impressive in the ROI analyses where we were able to account for some of the effects of susceptibility artifact. At the most, our data show a small difference in degree, not kind, between activation in the anterior-most aspect of the hippocampus compared with the posterior aspects. Our study does not rule out the possibility that encoding and retrieval of visuospatial material preferentially activate the posterior hippocampus. This possibility is consistent with the Moser and Moser model, the fMRI study by Stern et al. (1996), and the PET study by Schacter et al. (1999), all of which focused on visual memory and showed predominantly posterior hippocampal activation. Such a distinction between verbal and visual episodic memory merits further investigation. Finally, regardless of the model to be tested, we propose that future fMRI studies of the hippocampus should include some estimate of susceptibility artifact. Recent work by Devlin et al. (2000) and Cordes et al. (2000) has shown that susceptibility artifact is not limited to the temporal pole or lateral temporal regions and our study confirms that it can significantly impact regional analyses of hippocampal function.

REFERENCES

engages the posterior hippocampus: a parametrically analyzed functional magnetic resonance imaging study. J Neurosci 18:1841–1847.


