The Dissertation Committee for Jackson Chit Liang Certifies that this is the approved version of the following dissertation:

Content Representation in the Human Medial Temporal Lobe

Committee:

Alison R. Preston, Supervisor

Laura L. Colgin

Michael R. Drew

Russell A. Poldrack

David M. Schnyer

Content Representation in the Human Medial Temporal Lobe

by

Jackson Chit Liang, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2015

Acknowledgments

I'd like to thank Dr. Alison Preston for her constant mentorship over the last six years. I'd also like to thank everyone in the Preston lab for their friendship and their insights. I would like to thank my thesis committee, whose feedback shaped the trajectory of my projects and made my graduate career a smooth one. Thanks to my parents, Anna and William, and my brother, Victor, for always cheering me on. And deepest thanks to my dear wife Annie, for her unfailing love and support.

Content Representation in the Human Medial Temporal Lobe

Jackson Chit Liang, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Alison R. Preston

The transformation of sensory inputs into complex memory representations is fundamental to human experience; yet, little is known about how this crucial process is achieved. When you meet your friend at the new cafe in town, what part of the brain encodes this novel scene into long term memory? What part encoded your friend's favorite t-shirt, so that the sight of it gives you a feeling of familiarity rather than surprise? It is well-established that the medial temporal lobe (MTL) is crucial to both processes, but the MTL is not a single homogeneous region. In fact, it is composed of several anatomically distinct subregions including hippocampus, perirhinal cortex (PRC) and parahippocampal cortex (PHC). However, the computations performed by each subregion to encode individual events is still unclear. The present research tests the central hypothesis that different forms of event content are transformed into memory by distinct subregions within the MTL. A critical barrier in the study of content representation thus far has been its focus on comparing univariate peak activations in a region to different stimulus materials. To go beyond this limited approach, we employed multivariate statistical analyses that takes into account how event content is represented by distributed activity in MTL subregions. First, we examine the content-specific contributions of MTL subregions to episodic encoding and retrieval. Then, we

demonstrate how these distributed representations support memory-based prediction to resolve ambiguities in our environment.

Table of Contents

Chapter 1. General Introduction

MEDIAL TEMPORAL LOBE SUBREGIONAL FUNCTION IN HUMAN EPISODIC MEMORY: INSIGHTS FROM HIGH-RESOLUTION FMRI

The majority of text and figures in this section were prepared for a chapter in *The Cognitive Neuroscience of Human Memory*, which is currently in press:

Liang JC, Preston AR. In press. Medial Temporal Lobe Subregional Function in Human Episodic Memory: Insights from High-resolution fMRI. In Addis DR, Barense M, Duarte A, eds. The Wiley Handbook on the Cognitive Neuroscience of Memory, First Edition. John Wiley & Sons, Ltd. p108-130.

Co-author contributions: A. R. Preston is my P.I. and helped write the chapter.

Abstract

Understanding the role of the medial temporal lobe in episodic memory requires precise knowledge about the processing and representational capacity of individual structures within the region. In this chapter, we discuss how the advent of high-resolution functional magnetic imaging techniques has provides novel and important insights about the function of medial temporal lobe structures in the human brain.

Introduction

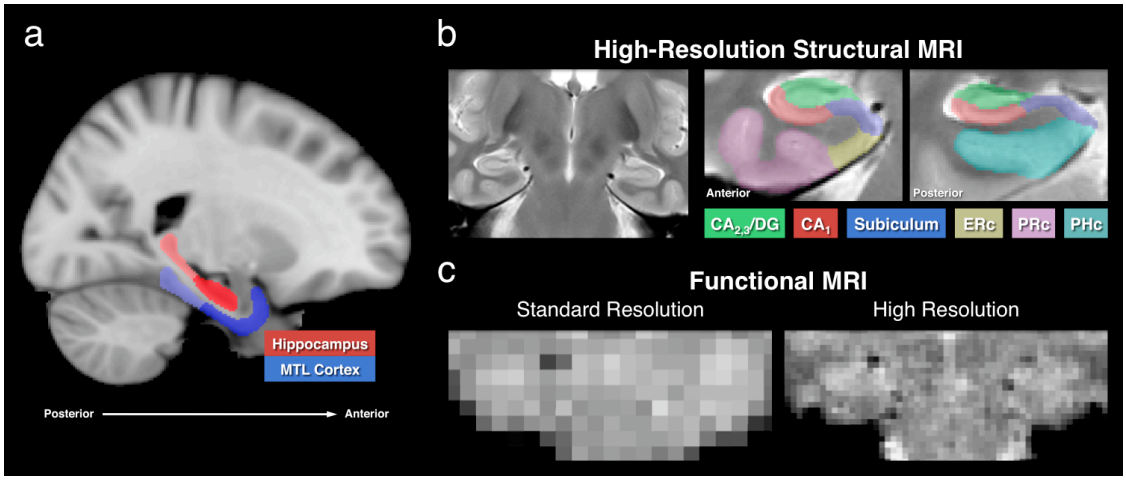
Episodic memory fundamentally shapes human behavior, allowing us to draw upon past experience to inform current decisions and make predictions about upcoming events. Decades of research has documented the critical role of the medial temporal lobe (MTL) in episodic memory (Eichenbaum & Cohen, 2001). While the link between MTL function and episodic memory is beyond debate, a recent focus has centered on characterizing the contributions of specific MTL substructures to episodic memory formation and retrieval.

The MTL (Figure 1.1a) is comprised of a heterogeneous group of structures, each with a unique cellular organization and pattern of anatomical connectivity. The subregions of the MTL include the hippocampus, which itself is comprised of the dentate gyrus (DG), the cornu ammonis (CA) fields, and the subiculum, as well as the surrounding entorhinal (ERc), perirhinal (PRc), and parahippocampal (PHc) cortices (Figure 1.1b). Several theoretical perspectives propose that anatomical differences between MTL subregions give rise to unique functional roles in episodic memory (Davachi, 2006; Diana, Yonelinas, & Ranganath, 2007; McClelland, McNaughton, & O'Reilly, 1995; O'Reilly & Rudy, 2001). However, testing how MTL subregions contribute to human memory function poses a unique challenge for cognitive neuroscience research. Individuals with MTL lesions typically have damage that affects several MTL subregions, spanning both hippocampus and surrounding MTL cortices. Even those individuals with restricted hippocampal lesions have damage to multiple hippocampal subregions. Therefore, while the neuropsychological study of MTL patients

has taught us a great deal about the essential nature of the region for episodic memory, it is limited in its ability to discern the functional roles of individual human MTL subregions. Similarly, because MTL subregions are relatively small, adjacent structures standard approaches to functional magnetic resonance imaging (fMRI) that use voxel dimensions greater than 3 mm cannot resolve signal originating from a particular MTL subfield. Testing the predictions of anatomically based models of MTL function in the human brain thus requires a spatial resolution beyond the limits of neuropsychological study and standard functional neuroimaging methods.

Over the last decade implementation of high-resolution functional magnetic resonance imaging (hr-fMRI) has opened the door for investigation of MTL subregional function in humans. In this chapter, we review the technical aspects of hr-fMRI as applied to the study of the human MTL and discuss two core topics that have dominated research in this area: (1) functional dissociations between hippocampus and surrounding MTL cortices based on episodic memory content and (2) functional distinctions between the components of the hippocampal circuit. We also discuss new multivariate pattern-information analysis techniques, which examine distributed patterns of activation in contrast to average responses pooled across an entire region. Such techniques, when combined with hr-fMRI, have the power to provide new insights into the function of MTL subregions. We end by discussing challenges for hr-fMRI of the human MTL and suggest future directions that could improve our ability to answer questions about the role of this region in episodic memory.

Figure 1.1: High-resolution fMRI of human MTL subregions. (**A**) Sagittal brain slice depicting the location of the hippocampus (in red) and surrounding MTL cortex (in blue). Dark gradation indicates the anterior portions of the hippocampus and MTL cortex, while light gradation depicts the corresponding posterior regions. (**B**) Structural images collected using high-resolution MRI in the coronal plane, perpendicular to the anterior-posterior MTL axis. Anterior and posterior segments show demarcation of anatomical MTL ROIs including hippocampal subfields DG/CA_2/CA_3, CA_1, and subiculum and MTL cortical subregions ERc, PRc and PHc. (**C**) Left panel shows a standard resolution fMRI image acquired using an functional sequence (3.75 x 3.75 x 3.6 mm voxels); right panel shows a high-resolution fMRI image acquired using a GRAPPA-EPI sequence (1.5 mm isotropic voxels).



What is high-resolution when it comes to human MTL imaging?

Standard fMRI methods typically employ inplane resolutions of ≥ 3 x 3 mm (Figure 1.1c). At this spatial resolution, precise identification of distinct hippocampal subfields is not possible, and the ability to differentiate activation arising from the ERc and PRc is also limited. A little over a decade ago, two research groups (Small, Nava, Perera, Delapaz, & Stern, 2000; Small, Wu, et al., 2000; Zeineh, Engel, & Bookheimer, 2000) developed techniques that enabled data acquisition from human MTL with enhanced spatial resolution (< 2 x 2 mm inplane resolution) that when combined with specialized data analysis procedures afford localization of blood oxygen level dependent (BOLD) signals to individual MTL subregions (Figure 1.1c). Such reduced voxels sizes not only improve the ability to distinguish anatomical boundaries between regions, but also reduce partial volume effects that may mask activations of interest (Bellgowan, Bandettini, van Gelderen, Martin, & Bodurka, 2006).

These initial studies, and several that followed them, acquired functional images in the oblique coronal plane (inplane), perpendicular to the long axis of the hippocampus, with a larger voxel dimension in the anterior-posterior direction (thruplane; e.g., 1.6 x 3 x 1.6 mm; Zeineh et al., 2000). By increasing spatial resolution in the coronal plane, these methods maximize the ability to identify key anatomical landmarks (Amaral & Insausti, 1990; Duvernoy, 1998; Insausti, Insausti, Sobreviela, Salinas, & Martinez-Penuela, 1998; Pruessner et al., 2002; Pruessner et al., 2000) that distinguish the boundaries between MTL subregions in the human brain. More recent studies (Bakker, Kirwan, Miller, & Stark, 2008; Hassabis et al., 2009) have advanced acquisition methods further, allowing

for isotropic voxel dimensions at the resolution of 1.5 mm^3. In all cases, hr-fMRI methods enable segmentation of the human hippocampus into the subiculum, CA1, and a combined DG/CA2/CA3 region (these subfields cannot be accurately differentiated even using current hr-fMRI methods; Figure 1.1b). These high-resolution acquisition techniques also afford more accurate segmentation of parahippocampal gyrus into the ERc, PRc, and PHc subregions, and more recently have been used to differentiate medial and lateral regions within ERc (Schultz, Sommer, & Peters, 2012).

To preserve spatial resolution, hr-fMRI studies of MTL function typically forgo or apply only minimal smoothing to minimize blurring of anatomical boundaries between regions. Several hr-fMRI studies have also refrained from conducting voxel-level group analyses due to the inherent challenges of registering small MTL subregions across participants; instead, these studies employ anatomically based region-of-interest (ROI) analyses in the native space of individual participants. In this approach, the functional timeseries is co-registered to an even higher resolution structural image (e.g., 0.4 x 3 x 0.4 mm) at the level of individual participants. Anatomical MTL subregions are then defined on the high-resolution structural image separately for each participant (Figure 1.1b), and task-related activation is extracted from each voxel within a region and averaged across all voxels in a given ROI. While this method avoids the potential issues of across-participant registration, it may also demonstrate reduced detection sensitivity, as voxels that are nonresponsive to the task are included in the averaging. To increase detection sensitivity, some studies first identify task-activated voxels within anatomical regions with a contrast orthogonal to the main question of interest; selective averaging

assessing effects of interest is then performed only on those task-activated voxels. However, both of these ROI methods preclude detection of heterogeneous responses that may be present within individual MTL subregions as an average response is calculated across all selected voxels in a region. If different voxels within a region have distinct response profiles, selective averaging further limits detection sensitivity.

In the past few years, several advances have been made in cross-participant registration techniques (Avants et al., 2011; Ekstrom et al., 2009; Yassa & Stark, 2009) that allow for reliable voxel-level analyses at the group level. These techniques employ fully deformable non-linear registration algorithms to warp each participant's anatomical and functional images to a template image (either a target participant's brain or a study-specific group template) using each participant's anatomically defined MTL subregions as a guide. After cross-participant registration, second-level group analyses can be used to identify activation patterns that are consistent across the group. One previously successful approach to cross-participant analyses relies on computational unfolding of MTL images into two-dimensional flat-maps (Zeineh, Engel, Thompson, & Bookheimer, 2003; Ekstrom et al., 2009). However, the unfolding operation can be prone to error that results in large spatial distortions, leading to inaccurate labeling of subregions after warping. Label-guided alignment approaches result in more accurate correspondence of MTL subregions across subjects and higher statistical sensitivity than standard methods (Yassa & Stark, 2009). Importantly, these methods also permit visualization of the topographic distribution of activation both within and across MTL subfields.

Anatomically-derived theories of MTL subregional function

Before delving into the empirical work using hr-fMRI to study human MTL function, it is important to consider the theoretical frameworks that guide such research. Leading models of MTL function in episodic memory (Davachi, 2006; Diana et al., 2007; Eichenbaum, Yonelinas, & Ranganath, 2007; Knierim, Lee, & Hargreaves, 2006; Manns & Eichenbaum, 2006; McClelland et al., 1995; Norman & O'Reilly, 2003) derive many of their predictions from the anatomical organization of the region, with the putative function of each MTL subregion being linked to the nature of its inputs, outputs, and internal circuitry (Figure 1.2). In the case of MTL cortex, PRc receives predominant input from unimodal visual association areas in ventral temporal cortex, while PHc receives input from posterior visual association areas in parietal cortex as well as auditory and somatosensory information (Jones & Powell, 1970; Suzuki & Amaral, 1994; Van Hoesen & Pandya, 1975; Van Hoesen, Pandya, & Butters, 1975). This pattern of extrinsic connectivity with neocortex suggests episodic memory encoding and retrieval may differentially recruit PRc and PHc depending on the nature of event content, with PRc supporting memory for visual objects and PHc supporting memory for visuospatial information. An influential extension of this view suggests that PRc mediates memory for individual items experienced within single episodes (the "what"), while PHc mediates memory for the context in which those items were experienced (the "where") (Davachi, 2006; Diana et al., 2007; Eichenbaum et al., 2007).

Moreover, PRc and PHc provide the respective inputs to the lateral and medial ERc in the rodent brain (Figure 1.2, dark blue arrows) (Burwell, 2000; Van Hoesen &

Pandya, 1975), suggesting that the segregation of mnemonic content would also be reflected in different regions of the ERc (Knierim et al., 2006; Manns & Eichenbaum, 2006). While PRc and PHc projections remain segregated within ERc, parallel inputs from lateral and medial ERc converge on to the same subsets of DG granule cells and CA3 pyramidal cells in the rodent hippocampus (Figure 1.2, light blue arrows) (Canto et al., 2008). DG in turn projects to CA3 via the mossy fiber pathway (Figure 1.2, purple arrow) (Witter et al., 2000). Projections from CA3 pyramidal cells include collaterals to other CA3 pyramidal cells comprising an extensive system of associational connections within the region (Figure 1.2, brown arrow). The convergence of inputs from lateral and medial ERc as well as CA3 collateral connections potentially distinguishes the putative function of the DG and CA3 from that of MTL cortical regions, with these hippocampal regions playing a domain-general role in episodic memory by binding disparate inputs from PRc and PHc into cohesive memory representations for long-term storage, i.e., binding the "what" happened to the "where" it happened (Diana et al., 2007; Eichenbaum et al., 2007).

While mnemonic processing in MTL cortical regions would be distinguished by their selective responses to specific forms of event content, hippocampal memory traces would reflect the arbitrary relationships among multimodal event elements as well as associations between those elements and the context of their occurrence ("what happened where") (Eichenbaum & Cohen, 2001; Morris et al., 2003). Importantly, the sparse connectivity between DG and CA3 is thought to magnify distinctions between overlapping patterns of cortical input elicited by highly similar events, a process termed

pattern separation (McClelland et al., 1995; O'Reilly & Rudy, 2001). Pattern separation is thought to result in separable memory traces for highly similar events that reduce the likelihood that memories would interfere with one another. CA3 circuitry is also hypothesized to support reactivation of stored memories from partial cues through recurrent excitation, a processed termed pattern completion (McClelland et al., 1995; O'Reilly & Rudy, 2001). While DG and CA3 likely make distinct contributions to pattern separation and pattern completion, hr-fMRI methods to date have not reliably distinguished either between these subfields or adjacent CA2 region. Thus, the majority of hr-fMRI studies commonly define a single region that encompasses all these structures, which is typically referred to as DG/CA2/CA3.

Farther along in the hippocampal circuit, CA3 provides a major input to CA1 (Figure 1.2, green arrow), which also receives direct input from ERc (Figure 1.2, orange arrows). Notably, whereas the projections from lateral and medial ERc converge on the same cells in DG and CA3, they target distinct groups of cells in CA1 and subiculum (Witter, Van Hoesen, & Amaral, 1989; Canto et al., 2008). The differences in ERc connectivity between hippocampal subfields suggests that while different forms of event content may evoke similar response patterns in DG and CA3, the responses of CA1 and subiculum may be heterogeneous with respect to different types of memory content, with different cells mediating memory for different kinds of content. Moreover, the convergence of inputs from CA3 and segregated sensory information from lateral and medial ERc in CA1 suggests that this hippocampal subregion compares memory-based output from CA3 pattern completion mechanisms to incoming sensory information from

ERc to detect deviations between current events and stored memories (Hasselmo & Schnell, 1994; Kumaran & Maguire, 2007; Lisman & Otmakhova, 2001; Vinogradova, 2001). When current experience violates expectations cued from memory, this CA1 comparator mechanism is thought to drive new encoding processes that form a new memory trace or update existing memories to account for new information.

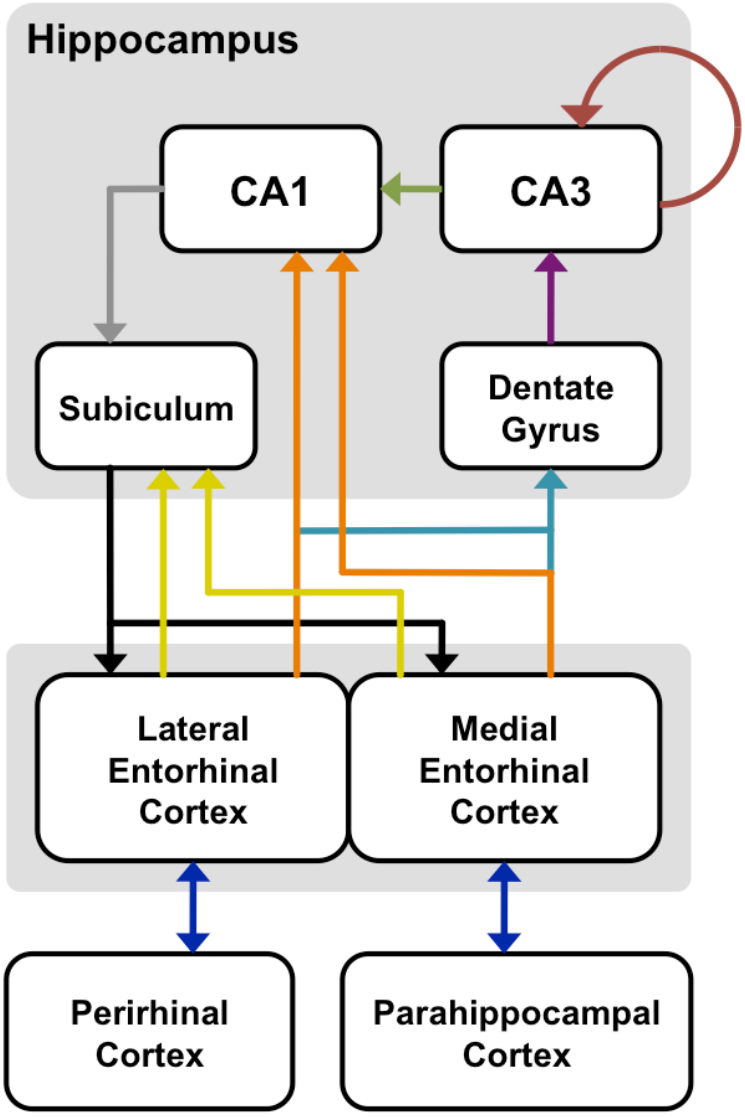
The subiculum, the final structure in the hippocampal circuit, receives highly processed input from CA1 (Figure 1.2, gray arrow) as well as direct inputs from ERc (Figure 1.2, gold arrows), PRc, and PHc. As the output structure of the hippocampus, the role of subiculum may be to distribute highly processed input from the CA fields to the neocortical regions from which the input originated (Kloosterman, Witter, & Van Haeften, 2003). For example, information about reinstated memories resulting from CA3 pattern completion would reach the subiculum via CA1; via back-projections to PRc and PHc (Figure 1.2, black arrows), subiculum could then facilitate reinstatement of the content-specific neocortical patterns active during initial learning. It is important to note that much of what we know about the structure and connectivity of the MTL region is based on the rodent brain, in particular the distinction between medial and lateral ERc, and it remains to be seen whether such distinctions translate to the human brain.

In addition to these hypothesized functional differences between hippocampal subfields, there has been renewed interest in functional differences along the anterior-posterior axis of the hippocampus (Poppenk, Evensmoen, Moscovitch & Nadel, 2013). Animal research has shown that the anatomical connectivity and function of the ventral (anterior in the human) and dorsal (posterior in the human) hippocampus are distinct. In

the rodent brain, the higher density of neuromodulatory inputs to ventral hippocampus relative to the dorsal hippocampus (Gage & Thompson, 1980; Verney et al., 1985) suggests that this region represents the behavioral salience of incoming information to guide memory formation regardless of content type (Fanselow & Dong, 2010; Moser & Moser, 1998). In contrast, animal lesion studies suggest that posterior hippocampus may be selectively involved in spatial learning tasks (Moser, Moser, & Andersen, 1993; Moser, Moser, Forrest, Andersen, & Morris, 1995). Episodic memory representations in the human brain might also reflect such anatomical and functional differences along the anterior-posterior hippocampal axis, with the posterior hippocampus playing a predominant role in mediating memory for information about the spatial context of individual events.

Collectively, these anatomical considerations provide an important theoretical framework motivating the body of studies using hr-fMRI to study human MTL function. In each of the following sections, we consider how hr-fMRI has informed these influential theories of MTL subregional function, beginning with empirical work on content representation in the human MTL.

Figure 1.2: Schematic diagram of connectivity between MTL cortex and hippocampal subfields. See text for detailed description of circuitry. Although not pictured, subiculum also receives direct input from PRc and PHc.



Empirical evidence for content-based dissociations between human MTL subregions

Several neuropsychological (Barense et al., 2005; Barense, Gaffan, & Graham, 2007; Bohbot et al., 1998; Epstein, Deyoe, Press, Rosen, & Kanwisher, 2001; Lee, Buckley, et al., 2005; Lee, Bussey, et al., 2005) and standard-resolution neuroimaging studies in humans (Awipi & Davachi, 2008; Lee, Scahill, & Graham, 2008; Pihlajamaki et al., 2004; Sommer, Rose, Glascher, Wolbers, & Buchel, 2005) have revealed functional differences between PRc and PHc along visual object and visuospatial domains as predicted by anatomically based theories. However, other evidence suggests that processing of specific forms of event content is distributed across subregional boundaries. For instance, PRc responses have been observed during encoding of objects, faces and scenes (Buffalo, Bellgowan, & Martin, 2006) and during binding of items to their specific features (Haskins, Yonelinas, Quamme, & Ranganath, 2008; Staresina & Davachi, 2006, 2008). Similarly, mnemonic responses in PHc have also been demonstrated for multiple forms of events content including spatial and non-spatial contextual information (Aminoff, Gronau, & Bar, 2007; Bar & Aminoff, 2003; Bar, Aminoff, & Ishai, 2008; Litman, Awipi, & Davachi, 2009).

These findings thus suggest two distinct possibilities for the nature of content representation in PRc and PHc: one comprised of well-defined PRc and PHc functional modules exhibiting preferential responding to specific event content, and an alternate possibility, with PRc and PHc processing and representing multiple forms of event content. High-resolution fMRI provides additional empirical leverage to distinguish between these opposing possibilities by enabling more precise delineation of the

boundaries between MTL cortical regions–in particular PRc from ERc–as well as unambiguous discrimination between MTL cortex and hippocampus.

Similarly, by delineating activation patterns arising from individual hippocampal subregions, hr-fMRI may resolve conflicting views of hippocampal function that alternately suggest processing in this region is either content-general (Awipi & Davachi, 2008; Davachi, 2006; Diana et al., 2007; Knierim et al., 2006; Manns & Eichenbaum, 2006; Staresina & Davachi, 2008) or specialized for spatial memory (Bird & Burgess, 2008; Kumaran & Maguire, 2005; Taylor, Henson, & Graham, 2007). One intriguing possibility suggested by the anatomical data is that distinct hippocampal regions, either individual subregions or different regions along the anterior-posterior hippocampal axis, may show dissociable response patterns with respect to representation of different forms of event content.

In an initial hr-fMRI study examining content-sensitivity in MTL regions (Preston et al., 2010), participants performed an incidental target detection task during the presentation of trial-unique, novel face and scene stimuli intermixed with highly familiar faces and scenes. Consistent with its proposed role in visuospatial processing, PHc responses were greater for novel scene trials relative to novel face trials. Moreover, greater activation in PHc scene-selective voxels was associated with enhanced subsequent scene memory. In contrast, PRc showed a pattern of novelty-based responding that was similar for faces and scenes. Moreover, the magnitude of novelty-based responses in face-sensitive and scene-sensitive voxels in PRc and subiculum correlated with later memory performance for each respective form of event content.

While these findings are consistent with a content-specific role for PHc in episodic encoding, they suggest that mnemonic processes in PRc and subiculum are generalized across different forms of event content (see also, Dudukovic, Preston, Archie, Glover, & Wagner, 2010).

Notably, exploration of content-sensitive responses in ERc was limited in these initial hr-fMRI reports, with minimal task-related activation observed in either study. Animal work suggests, however, that ERc plays a key role in episodic encoding and retrieval, with the lateral ERc mediating memory for object-related information and the medial ERc mediating spatial memory (Knierim et al., 2006; Manns & Eichenbaum, 2006). Recent hr-fMRI work has examined these hypothesized dissociations in content representation between the lateral and medial ERc, finding enhanced modulation of lateral ERc activation during face retrieval in contrast to enhanced medial ERc activation during the retrieval of spatial information.

New perspectives on content representation in the MTL have arisen from the application of multivariate pattern-information analyses to hr-fMRI data. Standard univariate fMRI analyses compare the mean response of a group of contiguous voxels across experimental conditions to isolate individual voxels or regions that show a statistically significant response to the experimental conditions of interest. To increase statistical sensitivity, univariate approaches may include spatial averaging across multiple voxels (e.g., a mean response to faces and a mean response to scenes within a specific anatomical region-of-interest, as illustrated by the bar chart in Figure 1.3a). Although this approach reduces noise inherent in all fMRI acquisitions, it also reduces sensitivity by

blurring out fine-grained spatial patterns that might discriminate between experimental conditions (Kriegeskorte & Bandetti, 2007). Instead, multivariate pattern-information approaches enhance detection sensitivity by looking at the contribution of multiple voxels, treating the pattern of response across all voxels within a region as a combinatorial code related to distinct mental operations (e.g., encoding faces vs. encoding scenes).

One such technique is multivoxel pattern analysis (MVPA; Haynes & Rees, 2006; Norman et al., 2006). Whereas univariate approaches use multiple regression to predict the activity of individual voxels based on the experimental condition, classification-based MVPA uses multiple regression to predict the experimental condition based on the activity of multiple voxels. In this approach, a machine-learning algorithm called a neural classifier is trained to distinguish brain patterns based on condition (e.g. whether the participant is encoding a face or a scene) using a subset of data. The trained classifier is then tested on previously unseen data (Figure 1.3b). Only if the experimental conditions are represented by distinct spatial patterns will final classifier predictions be accurate.

A related approach is representational similarity analysis (RSA; Kriegeskorte & Bandettini, 2007; Kriegeskorte et al., 2008) which does not use neural classifiers but rather examines the similarity structure (through correlation) between the multivoxel patterns elicited by experimental conditions (Figure 1.3c). RSA assumes that if two stimuli (e.g., two different faces) are represented similarly in the brain, there should be enhanced similarity between the multivoxel patterns evoked by the two stimuli (i.e., a

higher correlation value). Conversely, distinct mental representations would be reflected by dissimilar multivoxel patterns. Representational similarity matrices can be visualized using multidimensional scaling (MDS), where stimuli evoking similar voxel patterns are plotted closer together in representational space, while stimuli evoking dissimilar voxel patterns are plotted further apart (Figure 1.3d).

To date, two hr-fMRI studies have employed MVPA methods to examine content-based differences in MTL subregional representation. In the first study (Diana et al., 2008), participants viewed images in blocks comprised of different forms visual content (objects, scenes, faces, toys, and abstract shapes). While patterns of activation in PRc and hippocampus did not contain sufficient information to classify the different content types, PHc demonstrated accurate classification performance across all stimulus types including visual objects and faces. However, univariate analyses of PHc responses revealed a selective response to scenes. These findings highlight that univariate and multivariate approaches to fMRI data analyses index different aspects of the neural code, and further indicate that the scene-selectivity of PHc responses observed in previous studies, both standard- and high-resolution, do not convey the full nature of content representation in PHc.

A second hr-fMRI study (Liang et al., 2012) extended this work by demonstrating robust coding of many forms of event content in both PRc and PHc using both MVPA and RSA. In this study, participants incidentally encoded visual (faces, scenes, visual words) and auditory (environmental sounds, spoken words) stimuli. As in prior research, the univariate response in PRc was maximal for faces, while PHc showed a scene-

selective pattern of response. However, these PRc and PHc effects were accompanied by greater consistency between multivoxel patterns evoked by faces and scenes in both regions, and in the case of PHc, auditory stimuli as well. Importantly, the distinct representation of face, scene, and auditory content in PHc was most prominent at the most posterior aspect, counter to the prediction from standard-resolution univariate analyses (e.g., Staresina et al. 2011, Epstein and Kanwisher, 1998) that this posterior aspect should be the most scene-selective region of PHc. Moreover, the distinctive representation of faces and auditory content was observed in PHc despite the absence of an above-baseline response for these stimuli in the univariate analyses, further emphasizing the enhanced sensitivity of multivariate methods to representational content in MTL regions.

Liang et al. (2012) also observed different patterns of content-based responding across the anterior-posterior axis of the hippocampus. Mean univariate responses in anterior hippocampus were above baseline for all content classes. However, the spatial pattern of response in this region did not discriminate between different forms of event content. In contrast, posterior hippocampus showed a distributed coding of scene content that was distinct from other forms of content. These findings are consistent with the anatomically based view that anterior and posterior hippocampus serve different functions with respect to episodic memory. Anterior hippocampal signals may convey the behavioral salience of stimuli (e.g., that a stimulus is novel or rewarding) regardless of the perceptual form of the stimulus, while posterior hippocampus may play a predominant role in representing spatial memory content (see also, Hassabis et al., 2009).

Another means of assessing MTL content representation is to assess how specific regions within the MTL communicate with one another—a method known as functional connectivity. Recent hr-fMRI studies indexing connectivity within the MTL circuit provide convergent evidence for functional differences along the anterior-posterior hippocampal axis, particularly in CA1 and subiculum (Libby, Ekstrom, Ragland, & Ranganath, 2012). This work revealed that anterior regions of CA1 and subiculum show predominant connectivity with PRc regions, while posterior CA1 and subiculum show greater connectivity to PHc. Such differences in anterior-posterior hippocampal connectivity with cortex were notably absent in the DG/CA2/CA3. These findings provide the first evidence that, in the human brain, PRc and PHc communicate with distinct regions of CA1 and subiculum, whereas PRc and PHc communicate with DG/CA2/CA3 in a similar manner. In particular, these connectivity findings suggest that distinct coding of spatial content in posterior hippocampus observed using MVPA approaches (Hassabis et al., 2009; Liang et al., 2012) may primarily reflect CA1 and subiculum representations derived from PHc input.

One issue surrounding high-resolution MVPA studies of MTL content representation (Diana et al., 2008; Hassabis et al., 2009; Liang et al., 2012) is the lack of a direct link between distributed patterns of activation and behavioral measures of memory performance. It is therefore unclear whether the distributed representations of event content observed in PRc, PHc, and posterior hippocampus observed in these studies are related to successful encoding of specific types of stimuli. Recent hr-fMRI evidence suggests that distributed hippocampal activation patterns distinguish individual episodic

memories during vivid recall (Chadwick, Hassabis, & Maguire, 2011; Chadwick, Hassabis, Weiskopf, & Maguire, 2010). Moreover, patterns of hippocampal activation elicited by individual complex scenes have been used to decode participants' choice behavior in a perceptual decision making task (Bonnici, Kumaran, Chadwick, Weiskopf, Hassabis, & Maguire, 2012), further linking distributed hippocampal representations to behavior. Future hr-fMRI work will be necessary to determine whether similar relationships between distributed MTL representations and memory performance are true for a diversity of event content beyond the spatial domain.

Figure 1.3: Univariate and multivariate approaches for fMRI analysis. While standard univariate analyses (**A**) average across multiple voxel, multivariate approaches consider the contribution of multiple voxels, examining the pattern of response across all voxels within a region. Multivariate approaches include multivoxel pattern analysis (**B**) (continued next page).

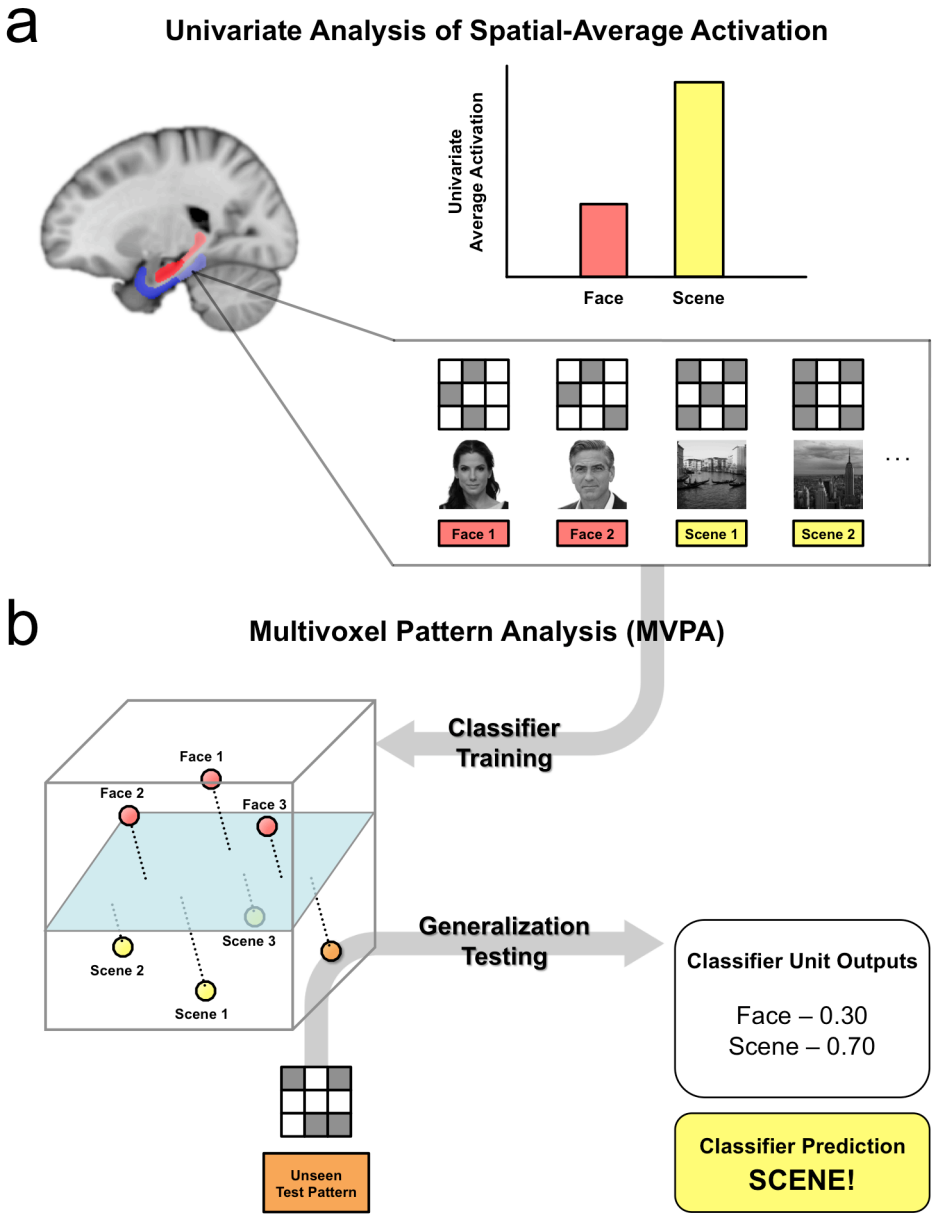
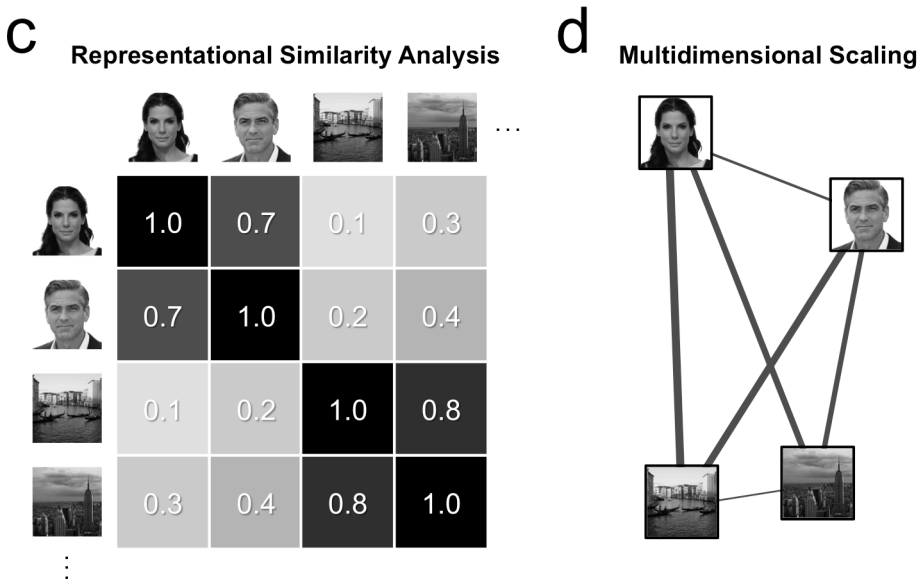


Figure 1.3 (continued): Multivariate approaches also include representational similarity analysis (**C**), which makes use of multidimensional scaling (**D**).



Differentiation of function between hippocampal subfields

In addition to providing key insights into the nature of content representation in the hippocampus and MTL cortex, hr-fMRI studies have played in important role in delineating the specific processes and computations that are supported by individual hippocampal subregions. To date, studies of hippocampal subregional function have focused on three core topics: (1) the differential role of hippocampal subregions in encoding and retrieval processing, (2) hippocampal subregional computations that support pattern separation and pattern completion, and (3) the proposed comparator function of the CA1 field of the hippocampus.

Hippocampal subregional contributions to episodic encoding and retrieval

One of the earliest hr-fMRI studies of the MTL demonstrated a dissociation of encoding and retrieval operations between hippocampal subfields (Zeineh et al., 2003). While DG/CA2/CA3 was engaged during encoding of face-name pairs, subiculum was engaged during retrieval of learned associations (Figure 1.4a). Similarly, several follow up studies found that DG/CA2/CA3 encoding responses are greater for remembered relative to forgotten events (Eldridge, Engel, Zeineh, Bookheimer, & Knowlton, 2005; Suthana, Ekstrom, Moshirvaziri, Knowlton, & Bookheimer, 2011; Suthana et al., 2009) even when memory is tested after a long delay (Carr, Viskontas, Engel, & Knowlton, 2010). In contrast, responses in CA1 and subiculum were associated with success effects at the time of retrieval (Eldridge et al., 2005; Viskontas, Carr, Engel, & Knowlton, 2009). Based on these findings, the authors hypothesized that the input structures of the

hippocampus, DG/CA2/CA3, are predominantly engaged during new event encoding, whereas the output structures of the hippocampus, the CA1 and subiculum, subserve the successful retrieval of memories (see Olsen et al., 2009, for supporting evidence from a delayed-match-to-sample paradigm).

However, in contrast to these studies, anatomical models of hippocampal function emphasize that individual subfields play important roles during both encoding and retrieval and may transiently switch between states (Colgin et al., 2009; Hasselmo & Schnell, 1994; Hasselmo, Schnell, & Barkai, 1995; Meeter, Murre, & Talamini, 2004). Indeed, other hr-fMRI studies have shown encoding and retrieval processes that are localized to multiple hippocampal subfields. During incidental encoding, subiculum activation is modulated by the novelty of presented item (Bakker et al., 2008), with the degree of novelty-related modulation predicting later memory (Preston et al., 2010). Furthermore, a hr-fMRI study examining the effect of reward on encoding responses in MTL subregions found that encoding activation was related to later memory in all hippocampal subfields (Figure 1.4b; Wolosin, Zeithamova, & Preston, 2012). Similarly, several studies have shown retrieval success effects throughout the hippocampal circuit (Chen, Olsen, Preston, Glover, & Wagner, 2011; Suzuki, Johnson, & Rugg, 2011). Together, these recent experiments illustrate that encoding and retrieval processes are not restricted to specific hippocampal subfields as suggested by earlier studies. However, further work is required to determine whether subfields might perform specific aspects of encoding and retrieval, such as encoding of the environment during spatial navigation

(Suthana et al., 2009) or of salience cues encountered during novel events (Wolosin et al., 2012).

Pattern separation and completion

Another central focus of hr-fMRI studies in humans has sought to characterize hippocampal subregional responses elicited by highly overlapping perceptual inputs to determine their putative roles in pattern separation and pattern completion. Convergent electrophysiological research in rodents has shown that DG responses exhibit the greatest differentiation between highly overlapping input patterns, indicating this region's key role in pattern separation (Leutgeb, Leutgeb, Moser, & Moser, 2007). In turn, the role of CA3 and CA1 in pattern separation and pattern completion is thought to vary based on the degree of overlap between inputs representing past and present experiences, with the CA3 responding in a non-linear manner to pattern overlap and the CA1 responding in a linear fashion (Guzowski, Knierim, & Moser, 2004; Lee, Rao, & Knierim, 2004; Leutgeb, Leutgeb, Treves, Moser, & Moser, 2004; Vazdarjanova & Guzowski, 2004). For example, a low degree of overlap between input patterns leads to a novel pattern of response in CA3 (i.e., pattern separation), whereas higher degrees of overlap between input patterns elicits reinstatement of a previously established CA3 response (i.e., pattern completion).

In the first hr-fMRI study to demonstrate pattern separation and pattern completion biases in human hippocampal subfields (Bakker et al. 2008), participants viewed a sequential presentation of visual objects that contained novel objects seen for

the first time, identical repetitions of previously presented objects, and novel lure items that were perceptually similar to previously presented objects. This approach relies on an effect known as repetition suppression, in which MTL regions show a reduced BOLD response to previously viewed stimuli when they are later shown again. The authors hypothesized that regions biased toward pattern completion would automatically reinstate the representation of a previously viewed object when presented with its corresponding perceptual lure, and thus show similar responses to both repeated and lure objects. In contrast, regions biased toward pattern separation would differentiate the lures from previously presented, highly similar objects and treat them as novel, thus showing greater activation for both novel and lure trials relative to repeated objects. The results showed a pattern of activation in CA1 and subiculum consistent with pattern completion, whose response was reduced for both repeated and lure items. In contrast, responses in DG/CA2/CA3 showed a pattern separation bias, successfully differentiating lure trials from similar, familiar objects (for related hr-fMRI findings isolated to the entire hippocampal region see Johnson, Muftuler, & Rugg, 2008).

To address predictions from rodent models that the balance between pattern separation and completion in different hippocampal subregions may depend on the overlap between present input and past experience, a subsequent experiment used two types of perceptual lures that had either high or low degrees of perceptual similarity to previously presented objects (Lacy, Yassa, Stark, Muftuler, & Stark, 2011). Consistent with prior findings, responses in DG/CA2/CA3 demonstrated a pattern separation bias, with the level of bias being similar for high and low similarity lures. In contrast, CA1

responses showed a graded response that depended on the degree of similarity between lures and familiar objects. These data converge with rodent research to suggest that human DG/CA2/CA3 shows a non-linear response to overlapping patterns, while human CA1 responses are more linear in nature. However, future experiments that provide more quantitative manipulations of stimulus similarity across several levels of similarity will be required before making strong claims regarding the nature of pattern separation and completion biases in the human hippocampus. Notably, when the same stimuli and presentation procedures were combined with an intentional task focus, which required participants to identify each object as novel, repeated, or lure, dissociations between hippocampal subfields were not apparent (Kirwan & Stark, 2007). These divergent findings suggest that mnemonic demands have a major influence on processing in the hippocampus (see also Dudukovic & Wagner, 2007; Duncan, Curtis, & Davachi, 2009; Kumaran & Maguire, 2009), with task goals impacting the bias to form new memory representations versus retrieving existing ones. Future hr-fMRI studies are needed that directly address how goal states influence the computational properties of hippocampal subregions and, in particular, the tradeoff between pattern separation and pattern completion.

One could also argue that paradigms manipulating visual similarity between individual objects as a means to study hippocampal pattern separation and completion biases do not assess the true nature of overlapping episodic memories. Using videos depicting real-world actions performed by individuals in different contexts, Chadwick et al. (2010) showed that the distributed pattern of hippocampal activation evoked during

vivid recall distinguished between individual memories. However, because the episodes portrayed in the videos did not share common features, this study did not directly address how the hippocampus codes highly overlapping episodic memories. In a second study (Chadwick et al., 2011), the videos were constructed from two realistic action sequences filmed on a "green screen" background that were superimposed on the same two spatial contexts, resulting in four video clips with highly overlapping features. Hippocampal activation patterns were distinct during recollection of each individual video, providing evidence for pattern-separated hippocampal representations depicting highly overlapping episodic information.

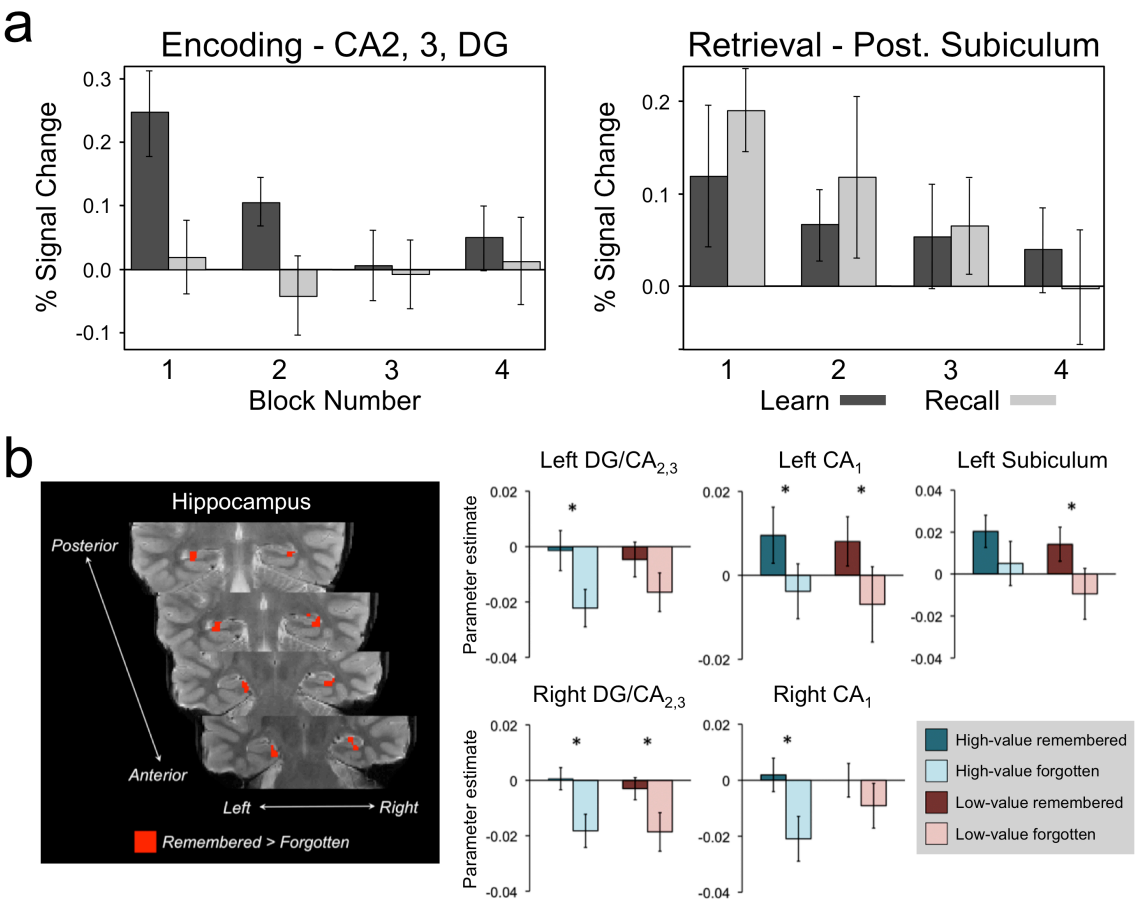
While these studies did not attempt to differentiate individual hippocampal subfields, the combination of these naturalistic stimuli, multivariate analysis methods, and quantitative manipulations of event similarity would be well suited to address critical questions of pattern separation and completion biases in individual MTL subregions. To date, only one published report has used multivariate classification to test for hippocampal subregional differences in pattern separation and completion biases, finding evidence for pattern completion in both CA1 and CA3 (Bonnici, Chadwick, Kumaran, Hassabis, Weiskopf, & Maguire, 2012). Future high-resolution studies combining both univariate and multivariate methods will be necessary to determine how these results line up with prior work associating CA3 predominately with pattern separation.

Hippocampus as a comparator

Several mnemonic processes, including pattern separation and pattern completion, require a comparison of the similarity between new events and existing memory representations. This comparator function is thought to elicit encoding processes when present events deviate from predictions derived from reinstated memory representations and has been hypothesized to rely on the CA1 subfield of the hippocampus (Hasselmo & Schnell, 1994; Kumaran & Maguire, 2007; Lisman & Otmakhova, 2001; Vinogradova, 2001). Two recent hr-fMRI studies (Chen et al., 2011; Duncan, Ketz, Inati, & Davachi, 2012) tested this hypothesis by examining hippocampal subfield responses to memory probes that matched or did not match previously studied events. In one experiment (Chen et al., 2011), participants studied associations between faces and houses prior to fMRI scanning. During the scanned retrieval phase, one member of studied face-house pairs was presented at the beginning of the trial. During a delay period, participants were instructed to recall the stimulus paired with the cue image. At the end of the trial, participants judged whether a probe image was the correct paired associate (a match) or a familiar image from another studied face-house pair (a mismatch). For correctly judged probe items, CA1 showed greater activation for mismatch probes compared to match probe items, consistent with a comparator signal that detects deviations from cued expectations. However, this pattern of CA1 response was only observed for house probes that were preceded by face cues, suggesting that CA1 may specifically serve as a comparator in the spatial domain.

In a second study, participants studied 3-dimensional room layouts prior to hr-fMRI scanning (Duncan et al., 2012). During a scanned recognition phase, participants viewed studied rooms that contained changes in layout and/or pieces of furniture. Within the hippocampus, only CA1 responses demonstrated sensitivity to changes in studied images, with a graded pattern of response based on the number of changes. This graded pattern of CA1 response was observed irrespective of the dimension of change participants were instructed to pay attention only to ("layout" or "furniture"), suggesting that the putative CA1 comparator response is automatic and does not depend on extrinsic task goals. While these studies provide compelling evidence that CA1 serves mnemonic comparator, future work is needed to determine how such automatic CA1 responses relate to successful encoding of new episodic information.

Figure 1.4: Encoding and retrieval effects in hippocampal subfields. (**A**) Encoding- and retrieval-related activation in hippocampus for face-name paired associates. Percent signal change is plotted for each of four alternating encoding and recall blocks in DG/CA_2/CA_3 (left) and posterior subiculum (right). Adapted from Zeineh et al., 2003. (**B**) Encoding activation in hippocampus during paired associate encoding under conditions of high and low reward. Voxels within the hippocampus demonstrating subsequent memory effects are displayed in red (left). Bar graphs (right) depict encoding activation in DG/CA_2/CA_3, CA_1, and subiculum for high-value remembered trials (dark blue), high-value forgotten trials (light blue), low-value remembered trials (red), and low-value forgotten pairs (pink). Adapted from Wolosin et al., 2012.



Limitations and future directions for high-resolution fMRI of human MTL

While hr-fMRI has advantages over standard approaches to brain imaging, it does have limitations. Notably, many of the studies reviewed here employ coronal acquisitions with a large thruplane resolution (3 mm or more). Such acquisition parameters maximize resolution in the inplane direction, in which distinguishing anatomical landmarks are most evident, while minimizing repetition time. However, because the anatomical landmarks that define individual MTL subregions shift gradually along the anterior-posterior MTL axis, large thruplane dimensions may prevent precise localization of activation when voxels include signal from multiple subfields. For this reason, some hr-fMRI studies opt not to make strong claims about individual subfields and prefer to treat the hippocampus as an entire region (e.g., Suzuki et al., 2011).

Current high-resolution methods are also limited in their ability to resolve CA2, CA3, and DG as separate regions, with virtually all current hr-fMRI studies treating these as a single region despite their dramatic differences in connectivity and structure. Recently, one hr-fMRI study reported functional differences between DG and other hippocampal subfields (Bonnici, Chadwick, et al., 2012), made possible through structural acquisitions with higher thruplane resolution than previous studies (0.5 mm). While this procedure undoubtedly benefits from the use anatomical landmarks visible only in the sagittal plane, some caution is warranted. First, the inplane resolution is somewhat lower than what has been reported in other hr-fMRI studies (0.53 versus 0.43mm), thus losing detail in the plane most commonly used to segment hippocampal subfields. Secondly, the resolution of the underlying functional data (1.5 mm^3) is no

different from prior studies and thus the ability to distinguish signal arising from DG and each CA fields in the functional data remains unchanged. For this reason, strong claims about dissociable responses in DG and CA3, for example, likely cannot be made based on such data.

Finally, the effort to test theories concerning small substructures is hampered by signal dropout and geometric distortion in functional acquisitions that are not present in structural images. Such distortion and dropout is particularly evident in anterior MTL regions. Distortion increases the likelihood that signal will be errantly displaced from a hippocampal subfield to one of its neighbors and while such distortion can be corrected using field maps, the low spatial resolution of standard field maps precludes correction of detailed structures like the hippocampal subfields. Efforts are currently underway to integrate higher resolution field maps into imaging analyses, but these maps require additional acquisition time and may provide limited benefits.

Recent technical developments, such as human 7T imaging and multiband parallel imaging techniques (Moeller et al., 2010), may further enhance spatial resolution, allowing us to move beyond current limitations. In particular, multiband parallel imaging techniques dramatically increase the number of slices that can be collected at a single time point. Increasing spatial resolution beyond 1.5 mm^3 will enable finer distinctions between hippocampal subfields or subregions of ERc, while also providing more detailed patterns of activation that can be leveraged using multivariate analysis techniques. Lowering the sampling rate to 1 second (or less) will also provide richer datasets for functional connectivity analyses. Finally, these new acquisition techniques permit

increased coverage beyond the MTL, while maintaining high spatial and temporal resolution, thus permitting novel investigations of how MTL subregions interact with memory centers in the frontal and parietal cortices.

Concluding remarks

High-resolution fMRI is an essential technique for evaluating theories of MTL function that had previously only been tested in animals. The combination of hr-fMRI and multivariate pattern-information analysis techniques, in particular, have substantially advanced our understanding of how memory is represented in MTL subregions, allow for greater parity and convergence with animal studies. High-resolution fMRI techniques also have increasing translational relevance. Many of the paradigms described in this review are being applied to clinical populations with neurological and psychiatric disorders associated with memory impairment (e.g., Gaisler-Salomon, Schobel, Small, & Rayport, 2009; Schobel et al., 2009; Suthana et al., 2009; Yassa et al., 2010), providing further insight into the relationship between pathological changes to MTL subregions and disease processes that affect memory. New technical developments are likely to advance the field further, affording greater leverage to characterize the critical MTL computations and representations that underlie episodic memory.

Chapter 2. Content Representation During Episodic Encoding

CONTENT REPRESENTATION IN THE HUMAN MEDIAL TEMPORAL LOBE

The majority of text and figures in this section have been published in 2013 in *Cerebral Cortex*:

Liang JC, Wagner AD, Preston AR. 2013. Content Representation in the Human Medial Temporal Lobe. Cereb Cortex. 23:80-96.

Co-author contributions: A. R. Preston is my P.I. and helped write the article. A. D. Wagner is a collaborator at Stanford University and provided valuable feedback at every stage of the project.

Abstract

Current theories of medial temporal lobe (MTL) function focus on event content as an important organizational principle that differentiates MTL subregions. Perirhinal and parahippocampal cortices may play content-specific roles in memory, whereas hippocampal processing is alternately hypothesized to be content-specific or content-general. Despite anatomical evidence for content-specific MTL pathways, empirical data for content-based MTL subregional dissociations are mixed. Here, we combined fMRI with multiple statistical approaches to characterize MTL subregional responses to different classes of novel event content (faces, scenes, spoken words, sounds, visual words). Univariate analyses revealed that responses to novel faces and scenes were

distributed across the anterior-posterior axis of MTL cortex, with face responses distributed more anteriorly than scene responses. Moreover, multivariate pattern analyses of perirhinal and parahippocampal data revealed spatially organized representational codes for multiple content classes, including non-preferred visual and auditory stimuli. In contrast, anterior hippocampal responses were content-general, with less accurate overall pattern classification relative to MTL cortex. Finally, posterior hippocampal activation patterns consistently discriminated scenes more accurately than other forms of content. Collectively, our findings indicate differential contributions of MTL subregions to event representation via a distributed code along the anterior-posterior axis of MTL that depends on the nature of event content.

Introduction

The medial temporal lobe (MTL) plays an essential role in episodic memory (Gabrieli 1998; Eichenbaum and Cohen 2001; Squire et al. 2004; Preston and Wagner 2007); yet, it remains an open question how MTL subregions differentially subserve episodic memory. Anatomical evidence suggests that event content might be an important organizing principle for differentiating MTL subregional function, with the nature of to-be-remembered information influencing MTL subregional engagement.

Predominant inputs from ventral visual areas to perirhinal cortex (PRc) and dorsal visual areas to parahippocampal cortex (PHc) suggest that these regions may differentially support memory for visual objects and visuospatial information respectively (Suzuki and Amaral 1994; Suzuki 2009). While several neuropsychological (Bohbot et al. 1998;

Epstein et al. 2001; Barense et al. 2005; Lee, Buckley et al. 2005; Lee, Bussey et al. 2005; Barense et al. 2007) and neuroimaging studies in humans (Pihlajamaki et al. 2004; Sommer et al. 2005; Awipi and Davachi 2008; Lee et al. 2008) have revealed functional differences between PRc and PHc along visual object and visuospatial domains, other evidence suggests distributed processing of event content across subregional boundaries. In particular, encoding responses have been observed for scenes, faces, and objects in human PRc (Buffalo et al. 2006; Dudukovic et al. 2010; Preston et al. 2010) and PHc (Bar and Aminoff 2003; Aminoff et al. 2007; Bar et al. 2008; Litman et al. 2009). Existing evidence thus suggests two distinct possibilities for the nature of content representation in PRc and PHc: one comprised of well-defined PRc and PHc functional modules that exhibit specialized and preferential responding to specific event content, and an alternate possibility where PRc and PHc represent multiple forms of event content. Recent studies have attempted to reconcile these conflicting accounts by demonstrating content-based representational gradients along the anterior-posterior axis of MTL cortex (Litman et al. 2009; Staresina et al. 2011). Encoding responses specific to visual object information have been observed in the anterior extent of PRc, while posterior regions of PHc show encoding responses specific to visuospatial information (Staresina et al. 2011). Interestingly, however, a transitional zone between anterior PRc and posterior PHc contributed to encoding of both visual object and visuospatial information. These findings suggest that discrete functional boundaries may not exist within MTL cortex; rather different forms of event content would evoke a graded pattern of response along the MTL cortical axis, with content-specific responses being more

likely in the anterior and posterior extents of MTL cortex. Notably, the particular distribution of such representational gradients may differ greatly depending on the nature of the event content (Litman et al. 2009).

Neuroimaging research on content representation in the MTL has almost exclusively employed standard univariate measures of response preferences that consider content-sensitivity as a function of the maximal response within a specific region. However, such univariate statistical techniques overlook the possibility that weaker, non-maximal responses represent important information about event content (Haxby et al. 2001; Norman et al. 2006; Harrison and Tong 2009; Serences et al. 2009). Unlike standard univariate analyses, multivariate analysis of neuroimaging data examines the entire pattern of response within a region of interest and is not necessarily limited to responses within a region that reach a predefined statistical threshold (Norman et al. 2006; Poldrack 2006; Kriegeskorte et al. 2008). These methods have proved a powerful tool for understanding the nature of representational codes for different forms of perceptual content in higher-order visual centers in the brain (Haxby et al. 2001; Kriegeskorte and Bandettini 2007; Kriegeskorte et al. 2007; Macevoy and Epstein 2009, 2011). For example, patterns of response in ventral visual regions that project to the MTL discriminate between multiple categories of visual stimuli (including houses, faces, and objects) even in regions that respond maximally to only one category of stimuli, suggesting widely distributed and overlapping representational codes for visual content in this region (Haxby et al. 2001).

Given evidence for a distributed coding of event content in content-selective visual regions, it may follow that representational coding in MTL cortical regions that receive direct input from these regions may also be distributed. In support of this view, recent evidence has shown that patterns of activation within PHc discriminate between non-preferred classes of content, including faces and objects even in the most posterior aspects of the region (Diana et al. 2008). This finding suggests that representational gradients along the anterior-posterior axis of MTL cortex do not sufficiently describe the distribution of content representation in this region. Thus, the precise nature of representational codes for different forms of event content in MTL cortex remains an important open question.

Evidence for the nature of content representation in the hippocampus is similarly mixed. Selective hippocampal damage impairs memory for visuospatial information while sparing memory for non-spatial information (Cipolotti et al. 2006; Bird et al. 2007; Taylor et al. 2007; Bird et al. 2008), suggesting a content-specific hippocampal role in spatial memory (Kumaran and Maguire 2005; Bird and Burgess 2008). Alternatively, the hippocampus may contribute to memory in a domain-general manner given the convergence of neocortical inputs onto hippocampal subfields (Davachi 2006; Knierim et al. 2006; Manns and Eichenbaum 2006; Diana et al. 2007). In support of this view, neuroimaging evidence has revealed hippocampal activation that is generalized across event content (Prince et al. 2005; Awipi and Davachi 2008; Staresina and Davachi 2008; Preston et al. 2010).

The application of multivariate statistical techniques to understand content coding in hippocampus has been limited to a single report to date (Diana et al. 2008; for discussion of related findings, see Rissman and Wagner In press). In Diana et al.'s study, hippocampal activation patterns demonstrated poor discrimination of scene and visual object content, suggesting that hippocampal representations are not sensitive to the modality of event content. However, that study examined the pattern of response across the entire hippocampal region. As in MTL cortex, one possibility is that different regions along the anterior-posterior axis of the hippocampus might demonstrate distinct representational codes for specific forms of event content. Animal research has shown that the anatomical connectivity and function of the ventral (anterior in the human) and dorsal (posterior in the human) hippocampus are distinct (Swanson and Cowan 1977), with the dorsal hippocampus being particularly implicated in spatial learning tasks (Moser M. B. and Moser 1998). Representational codes in the human brain might also reflect such anatomical and functional differences along the anterior-posterior hippocampal axis, with distinct spatial codes being most prevalent in the posterior hippocampus.

To provide an in-depth characterization of content representation in human MTL, we combined high-resolution fMRI (hr-fMRI) with both univariate and multivariate statistical approaches. Univariate analyses assessed responses to different classes of novel event content within anatomically defined MTL subregions. Importantly, by utilizing auditory (spoken words and sounds) and visual (faces, scenes, visual words) content, the current study aimed to broaden our knowledge of content representation in the human

MTL beyond the visual domain. As a complement to these univariate approaches, multivariate pattern classifiers trained on data from MTL subregions assessed whether distributed activity in each subregion discriminated between distinct content classes, including "non-preferred" content. Representational similarity analysis (RSA) (Kriegeskorte and Bandettini 2007; Kriegeskorte et al. 2008) further characterized the representational distance between exemplars from the same content class and between exemplars from different content classes to determine whether MTL subregions maintain distinctive codes for specific forms of information content. Given existing evidence for gradations in content-sensitivity that cross anatomical boundaries, we examined univariate and multivariate responses within individual anatomically defined MTL subregions, as well as the distribution of novelty responses along the anterior-posterior axis of the hippocampus and parahippocampal gyrus to test for content-based representational gradients.

By combining multiple statistical approaches with hr-fMRI, the present study aimed to provide a more precise characterization of content-representation in human hippocampus and MTL cortex than afforded by previous research. In particular, univariate and multivariate methods each index different aspects of the neural code. The use of both analysis methods in the current study provides a means to directly compare findings derived from these different approaches to present a comprehensive picture of representational coding in MTL subregions.

Materials and Methods

Participants

Twenty-five healthy, right-handed volunteers participated after giving informed consent in accordance with a protocol approved by the Stanford Institutional Review Board. Participants received $20/hr for their involvement. Data from 19 participants were included in the analyses (age 18-23 yrs, mean = 20.4 ± 1.7 yrs; 7 females), with data from six participants being excluded due to failure to respond on more than 20% of trials (three participants), scanner spiking during functional runs (one participant), and excessive motion (two participants).

Behavioral procedures

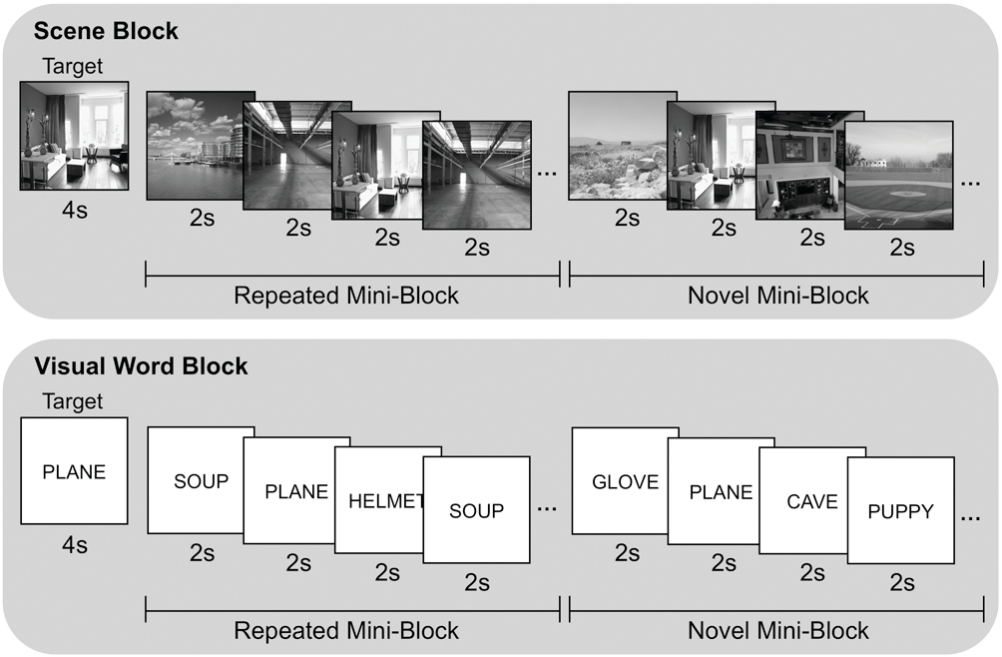
During functional scanning, participants performed a target detection task with five classes of stimuli: grayscale images of scenes, grayscale images of faces, visually presented words referencing common objects (white text on a black background; Arial 48 point), spoken words referencing common objects, and environmental sounds (e.g., jet engine, door creaking, water gurgling). During scanning, stimuli were generated using PsyScope (Cohen et al. 1993) on an Apple Macintosh computer and back-projected via a magnet-compatible projector onto a screen that could be viewed through a mirror mounted above the participant's head. Participants responded with an optical button pad held in their right hand.

During eight blocked-design functional runs, participants viewed or heard novel and repeated stimuli from each of the five stimulus classes while performing a target

detection task (Figure 2.1). Each run consisted of five stimulus-class blocks, one of each of the five stimulus classes, along with baseline blocks. At the start of each stimulus-class block, a cuing stimulus appeared for 4 s that represented the target for that block. Following this target cuing, two repeated and two novel mini-blocks of the stimulus class were presented in random order. During novel mini-blocks, participants were presented with eight stimuli (one target and seven trial-unique, novel stimuli) in a random order; each stimulus was presented for 2 s and participants indicated with a yes/no key press whether the stimulus was the target. Repeated mini-blocks also consisted of eight stimuli, including one target, with presentation and response procedures identical to novel mini-blocks. However, for repeated mini-blocks, the seven non-target stimuli consisted of two repeated stimuli that were used throughout the entire experiment. Participants viewed the two repeated stimuli from each class 20 times each prior to scanning.

Novel and repeated mini-blocks lasted 16 s each; thus, each stimulus-class block had a duration of 68 s (4-s target, 2 × 16-s novel mini-blocks, 2 × 16-s repeated mini-blocks). Across the entire experiment, participants performed the target detection task for 16 novel and 16 repeated blocks from each stimulus class. The presentation order of the stimulus-class blocks within each functional run was determined by one of three random orders, counterbalanced across participants. One 16-s baseline task block occurred at the beginning and end of each functional run. During baseline blocks, participants performed an arrow detection task; on each of eight trials, an arrow was presented for 2 s, and participants indicated by key press whether the arrow pointed to the left or right.

Figure 2.1: During functional scanning, participants performed target detection on novel and repeated stimuli from five classes: faces, scenes, sounds, spoken words, and visual words. At the beginning of a stimulus class block, a target stimulus would appear followed by two novel and two repeated mini-blocks in random order.



fMRI acquisition

Imaging data were acquired on a 3.0 T Signa whole-body MRI system (GE Medical Systems, Milwaukee, WI, USA) with a single-channel, custom-made transmit/receive head coil. Head movement was minimized using a 'bite bar' and additional foam padding. Prior to functional imaging, high-resolution, T2-weighted, flow-compensated spin-echo structural images (TR = 3000 ms; TE = 68 ms; 0.43 × 0.43 mm in-plane resolution) were acquired in 22 3-mm thick oblique coronal slices oriented perpendicular to the main axis of the hippocampus allowing for visualization of hippocampal subfields and MTL cortices. These high-resolution imaging parameters optimized coverage across the entire length of MTL but precluded collection of whole brain imaging data.

Functional images were acquired using a high-resolution T2*-sensitive gradient echo spiral in/out pulse sequence (Glover and Law 2001) with the same slice locations as the structural images (TR = 4000 ms; TE = 34 ms; flip angle = 90°; FOV = 22 cm; 1.7 × 1.7 × 3.0 mm resolution). Prior to functional scanning, a high-order shimming procedure, based on spiral acquisitions, was utilized to reduce B0 heterogeneity (Kim et al. 2002). Critically, spiral in/out methods are optimized to increase SNR and BOLD contrast-to-noise ratio in uniform brain regions while reducing signal loss in regions compromised by susceptibility-induced field gradients (SFG) (Glover and Law 2001), including the anterior MTL. Compared to other imaging techniques (Glover and Lai 1998), spiral in/out methods result in less signal dropout and greater task-related activation in MTL

(Preston et al. 2004), allowing targeting of structures that have previously proven difficult to image due to SFG.

A total of 768 functional volumes were acquired for each participant over eight scanning runs. To obtain a field map for correction of magnetic field heterogeneity, the first time frame of the functional timeseries was collected with an echo time 2 ms longer than all subsequent frames. For each slice, the map was calculated from the phase of the first two time frames and applied as a first order correction during reconstruction of the functional images. In this way, blurring and geometric distortion were minimized on a per-slice basis. In addition, correction for off-resonance due to breathing was applied on a per-time-frame basis using phase navigation (Pfeuffer et al. 2002). This initial volume was then discarded as well as the following two volumes of each scan (a total of 12 s) to allow for T1 stabilization.

Preprocessing of fMRI data

Data were preprocessed using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK) and custom Matlab routines. An artifact repair algorithm (http://cibsr.stanford.edu/tools/ArtRepair/ArtRepair.htm) was first implemented to detect and remove noise from individual functional volumes using linear interpolation of the immediately preceding and following volumes in the timeseries. Functional images were then corrected to account for the differences in slice acquisition times by interpolating the voxel time series using sinc interpolation and resampling the time series using the center slice as a reference point. Functional volumes were then realigned to the first volume in

the time series to correct for motion. A mean T2*-weighted volume was computed during realignment, and the T2-weighted anatomical volume was coregistered to this mean functional volume. Functional volumes were high-pass filtered to remove low frequency drift (longer than 128 s) before being converted to percentage signal change in preparation for univariate statistical analyses, or z-scored in preparation for multivoxel pattern analysis.

Univariate fMRI analyses

Voxel-based statistical analyses were conducted at the individual participant level according to a general linear model (Worsley and Friston 1995). A statistical model was calculated with regressors for novel and repeated mini-blocks for each stimulus class. In this model, each mini-block was treated as a boxcar, which was convolved with a canonical hemodynamic response function.

To implement group-level analyses, we used a non-linear diffeomorphic transformation method (Vercauteren et al. 2009) implemented in the software package MedINRIA (version 1.8.0, ASCLEPIOS Research Team, France). Specifically, each participant's anatomically defined MTL regions-of-interest (ROIs) were aligned with those of a representative "target" subject using a diffeomorphic deformation algorithm that implements a biologically plausible transformation respecting the boundaries dictated by the ROIs. Anatomically defined ROIs were demarcated on the T2-weighted, high-resolution in-plane structural images for each individual participant, using techniques adapted for analysis and visualization of MTL subregions (Pruessner et al.

2000; Zeineh et al. 2000; Pruessner et al. 2002; Zeineh et al. 2003; Olsen et al. 2009; Preston et al. 2010). A single participant's structural image was then chosen as the target, and all other participants' images were warped into a common space in a manner that maintained the between-region boundaries. To maximize the accuracy of registration within local regions and minimize distortion, separate registrations were performed for left hippocampus, right hippocampus, left MTL cortex, and right MTL cortex. Compared to standard whole-brain normalization techniques, this ROI-alignment or "ROI-AL-Demons" approach results in more accurate correspondence of MTL subregions across participants and higher statistical sensitivity (e.g., Kirwan and Stark 2007; Yassa and Stark 2009).

The transformation matrix generated from the anatomical data for each region was then applied to modestly smoothed (3 mm FWHM) beta images derived from the first-level individual participant analysis modeling novel and repeated stimuli for each content class. To assess how novelty-based MTL responses vary as a function of information content, two anatomically based ROI approaches were implemented. For the first analysis, parameter estimates for novel and repeated blocks for each of the five stimulus classes were extracted from five anatomically defined ROIs: PRc, PHc, entorhinal cortex (ERc), anterior hippocampus, and posterior hippocampus. Group-level repeated-measures ANOVA was used to test for differences in activation between novel blocks for each of the stimulus classes in each of the ROIs. Subsequent pairwise comparisons between content classes further characterized the stimulus sensitivity in each region.

In the present study, ERc did not demonstrate significant task-based modulation for any condition. Given the putative role of the ERc in the relay of sensory information to the hippocampus (Knierim et al. 2006; Manns and Eichenbaum 2006), the lack of task-based modulation despite the diversity of stimulus content may be somewhat surprising. To address the possibility that signal dropout in ERc might account for these null findings, we calculated the signal-to-noise ratio (SNR) observed during the baseline task within each anatomical ROI. Pairwise comparisons between ROIs revealed that posterior MTL regions exhibited higher SNR relative to anterior regions (all $p < 0.01$); posterior hippocampus had the highest SNR (mean = 9.17, SE = 0.30), followed by PHc (7.06 ± 0.30), anterior hippocampus (5.82 ± 0.21), and finally ERc (3.03 ± 0.21) and PRc (2.80 ± 0.18). Notably, SNR within ERc and PRc did not significantly differ ($p > 0.2$); yet, the present findings reveal above-baseline responding to multiple experimental conditions in PRc. Thus, signal dropout in anterior MTL remains a possible but inconclusive explanation for our lack of findings in ERc.

Because of the lack of task-based modulation of ERc, we focused our subsequent analyses of MTL cortical activation on the PRc and PHc ROIs. Region × content interactions, comparing PRc with PHc and anterior with posterior hippocampus, examined whether content sensitivity differed across the anterior-posterior axis of MTL cortex and hippocampus, respectively. A parallel set of analyses assessed differences between novel and repeated blocks for each class of content. Where appropriate, alpha-level adjustment was calculated using a Huynh-Feldt correction for non-sphericity.

A second anatomical ROI approach examined the distribution of novelty-based responses across the anterior-posterior axis of MTL cortex and hippocampus. To perform this analysis, the length of MTL cortex was divided into 11 anatomical ROIs defined using the representative "target" participant as the model. The placement of the ROIs along the anterior-posterior axis was selected to maintain the anatomical boundary between PRc and PHc. Each ROI was 4.5-mm long; however, due to the hemispheric asymmetry in length of the parahippocampal gyrus in the model subject, the anterior-most ROI of PRc was only 3-mm long in the left hemisphere and 6-mm long in the right hemisphere. Similarly, the longitudinal axis of the hippocampus was divided into nine ROIs based on the model participant, and the placement of the ROIs was selected to maintain the anatomical boundaries between the hippocampal head and body and between the hippocampal body and tail. Each ROI was 4.5-mm long; but again, due to the particular anatomy of the model subject, the posterior-most ROI in the hippocampal tail was 3-mm long in the right hemisphere and 6-mm long in the left hemisphere. Repeated-measures ANOVA assessed novelty-based activation (measured as both the response to novel stimulus blocks relative to baseline and the difference between novel and repeated blocks) as a function of content and anterior-posterior position along the axis of each structure. For both MTL cortex and hippocampus, one participant was excluded from this analysis because the slice prescription did not include the anterior-most aspect of the MTL region. For all analyses, hemisphere (left, right) was included as a within-subjects factor; however, because the effect of hemisphere did not interact significantly with any effect of interest (all $p > 0.1$), it is not considered in the Results. Moreover, the lack of

any observable effect of hemisphere suggests that the size discrepancy between the model participant's left and right MTL ROIs had no significant impact on the observed pattern of results.

Multivariate pattern analysis of fMRI data

In addition to the preceding univariate statistical analyses, we used multivoxel pattern analysis (MVPA) to determine the sensitivity of MTL subregions (anterior hippocampus, posterior hippocampus, PRc, and PHc) to different forms of event content. Pattern classification analyses were implemented using the Princeton MVPA toolbox and custom code for MATLAB. MVPA were performed at the individual participant level using the functional timeseries in native space. Classification was performed for each anatomical ROI region separately, and included all voxels within each ROI.

MVPA classification was performed by first creating a regressor matrix to label each timeseries image according to the experimental condition to which it belonged (e.g., novel faces, novel scenes, novel visual words, etc.). Classification was restricted to novel stimulus blocks, and there were an equal number of timepoints in each condition in the analysis (64 timepoints per condition). For each anatomical ROI, we assessed how accurately the classifier could discriminate between the stimulus classes. Classification performance for each ROI for each participant was assessed using an 8-fold cross-validation procedure that implemented a regularized logistic regression algorithm (Bishop 2006; Rissman et al. 2010) to train the classifier. Data from seven scanning runs were used for classifier training, and the remaining run was used as test data to assess the

generalization performance of the trained classifier. This process was iteratively repeated eight times, one for each of the possible configurations of training and testing runs. Ridge penalties were applied to each cross-validation procedure to provide L2 regularization. The penalties were selected based on performance during classification over a broad range of penalties, followed by a penalty optimization routine that conducted a narrower search for the penalty term that maximized classification accuracy (Rissman et al. 2010). Classifications performed for the purpose of L2 penalty selection were applied only to training data to avoid peeking at test data. The final cross-validated classification was performed once the optimal penalties were selected. The classification performances across the iterative training were then averaged to obtain the final pattern classification performance for each ROI for each participant.

To more closely examine the underlying activation patterns driving MVPA classification performance, we constructed confusion matrices indicating how often the MVPA classifier categorized voxel patterns correctly and how often it confused the voxel patterns with each other class of content. The goal of this analysis was to determine the distribution of classification errors for each class of stimuli (i.e., if a stimulus block was not correctly categorized, what stimulus class did the classifier identify it as). To do so, we constructed confusion matrices for each ROI from each participant and averaged them across the group. We then normalized each row of a given confusion matrix (representing one stimulus class) by dividing each cell of the matrix by the proportion of correctly classified test patterns for that stimulus category. This normalization procedure yielded values along the matrix diagonal equal to 1, and the resulting off-diagonal values indicate

confusability relative to the correct class of content. For example, stimulus classes that were highly confusable with the correct stimulus class would also yield values close to 1. To determine whether the level of confusability between stimulus classes was significantly different from chance, we scrambled the MVPA regressor matrix for each ROI for each participant so that each image of the timeseries was given a random condition label. Using Monte Carlo simulation (1000 iterations), we then created a null distribution of classification performance for each stimulus class based on the randomly labeled data as well as a null distribution of classifier confusion matrices. Classifier confusion values that lay outside of the confidence intervals based on the null distributions were determined to be significant. The alpha-level of the confidence intervals was chosen based on Bonferroni correction for each of 80 statistical tests of significance performed across all anatomical ROIs ($\alpha = 10^{-3}$).

Representational similarity analysis of fMRI data

To more precisely characterize the underlying representational structure for each form of stimulus content within MTL subregions, we examined responses to individual blocks of novel context using representational similarity analysis (RSA) (Kriegeskorte and Bandettini 2007; Kriegeskorte et al. 2008). We compared the patterns evoked by individual stimuli within and across content classes by considering the voxelwise responses observed for each novel mini-block viewed by the participants. Each novel mini-block contained the same configuration of eight stimuli across participants (though the mini-blocks were seen in different orders across participants). Here, we considered

each mini-block to represent an "exemplar" of a content class (Kriegeskorte et al. 2008) and constructed a separate general linear model with individual regressors for every mini-block of novel content. We first performed this analysis within the anatomically defined PHc, PRc, posterior hippocampus, and anterior hippocampus ROIs. To understand how representational structure changes as a function of position along the anterior-posterior axis of MTL, we also performed this analysis within each anterior-posterior segment of MTL cortex and hippocampus.

Representational dissimilarity matrices (RDM) were constructed for each MTL subregion for each individual participant. Each cell in the RDM indicates the Pearson linear correlation distance (1 – r) between voxelwise parameter estimates for any given pair of novel mini-blocks. Individual participants RDMs were averaged across the group. To better visualize the dissimilarities between stimulus class exemplars, we applied metric multi-dimensional scaling (MDS) to the group-averaged RDMs, which resulted in a two-dimensional characterization of the representational space of each region (Edelman 1998; Kriegeskorte et al. 2008). Metric MDS minimizes Kruskal's normalized STRESS1 criterion to represent each stimulus class exemplar as a point in, here, 2-dimensional space so that the rank order of linear distances between points matches the rank order of dissimilarities between exemplars in each RDM.

Based on these 2-dimensional representations of the RDMs, we calculated the mean within-class linear distance for each form of stimulus content as well as the mean cross-class linear distances for each pair of content classes. This analysis allowed us to determine whether exemplars from the same class of stimulus content (e.g., face mini-

block A versus face mini-block B) were clustered together in the representational structure of a given MTL subregion and whether the representation of those exemplars was distinct from exemplars from other contents classes (e.g., the distance between face mini-block A versus scene mini-block A). Monte Carlo simulation was used to assess whether within-class and cross-class linear distances were significantly different from the distances expected by chance. For each of 1000 iterations, the exemplar labels for each row and column of individual participant RDMs was randomly scrambled. These scrambled RDMs were averaged across participants and transformed using metric MDS to obtain null distributions of within-class and cross-class linear distances. Linear distances that lay outside of confidence intervals based on the null distributions were determined to be significant. The alpha-level of the confidence intervals was chosen based on Bonferroni correction for each of 15 statistical tests of significance performed within all anatomical ROIs ($\alpha = 10^{-2}$).

Results

Behavioral performance

Percent correct performance on the target detection task averaged 97.4 (SE = 0.41) for spoken words, 97.5 (0.68) for faces, 98.1 (0.34) for scenes, 96.5 (0.48) for sounds, and 98.7 (0.23) for visual words. A repeated-measures ANOVA revealed an effect of block type (novel, repeated: $F(1,18) = 7.25$, $p = 0.02$), an effect of stimulus content (spoken words, faces, scenes, sounds, visual words: $F(4,72) = 3.99$, $p = 0.01$), but no interaction ($F < 1.0$). Performance for novel blocks (97.9, 0.21) was superior to

performance for repeated blocks (97.4, 0.35). Pairwise comparisons revealed superior performance for visual word blocks relative to spoken word, face, and sound blocks (all $t > 2.20$, $p < 0.05$), as well as superior performance during scene blocks relative to spoken word and sound blocks (all $t > 2.10$, $p < 0.05$).

Analyses of reaction times (RTs) revealed effects of novelty ($F(1,18) = 29.58$, $p < 0.001$), stimulus content ($F(4,72) = 150.32$, $p < 0.001$), and an interaction between novelty and content ($F(4,72) = 3.12$, $p = 0.04$). RTs for repeated blocks (670 ms, SE = 20 ms) were faster than those for novel blocks (699 ms, 23). Significant differences in RTs were observed between all stimulus classes (all $t > 2.65$, $p < 0.05$), with the fastest RTs for visual word blocks (512 ms, 19), followed by scene (570 ms, 21), face (603 ms, 26), spoken word (832 ms, 23), and sound (895 ms, 34) blocks. The novelty × content interaction revealed that RTs decreased from repeated to novel blocks for all stimulus classes (all $t > 2.40$, $p < 0.05$), except sound blocks that demonstrated no RT difference between repeated and novel blocks ($t(18) = 1.69$). Performance on the baseline arrows task averaged 96.6% correct (SE = 0.75%).

Content-sensitivity within anatomically defined MTL ROIs

We first assessed whether activation during novel stimulus blocks varied based on content using a standard univariate analysis approach employed in several prior studies examining content-specific responding in MTL. Parameter estimates for novel blocks from each anatomically defined MTL ROI (ERc, PRc, PHc, anterior hippocampus, posterior hippocampus) were subjected to repeated-measures ANOVA for an effect of

content. Within MTL cortex, significant task-based modulation was observed only in PHc and PRc; we did not observe significant modulation of ERc activation for any condition or stimulus class (all $F < 1$), and therefore, we did not consider this region in any further analyses.

PHc activation during novel stimulus blocks demonstrated a significant main effect of content ($F(4,72) = 22.82$, $p < 0.001$). Among the five stimulus classes, only novel scenes elicited a significant response above baseline ($t(18) = 5.98$, $p < 0.001$; Figure 2.2a). Pairwise comparisons revealed that PHc activation for novel scenes was greater than activation for novel stimuli of all other stimulus classes (all $t > 5.95$, $p < 0.001$). Similar effects were observed for a parallel analysis assessing differences in PHc activation between novel and repeated stimuli for each class of content (Figure 2.2b). The difference in activation for novel relative to repeated stimuli demonstrated a significant effect of content ($F(4,72) = 7.39$, $p < 0.001$), with the novel – repeated difference being significant only for scenes ($t(18) = 5.92$, $p < 0.001$).

In PRc, activation during novel stimulus blocks was not different from baseline (all $t < 1.1$), and did not vary based on information content ($F(4,72) = 2.01$, $p = 0.12$; Figure 2.2a). When considering the difference in activation between novel and repeated blocks, however, a significant effect of content was observed in PRc ($F(4,72) = 4.03$, $p < 0.01$; Figure 2.2b), with the novel – repeated difference being significant only for faces ($t(18) = 3.84$, $p = 0.001$). Pairwise comparisons revealed that the novel – repeated difference in activation was greater for faces than for visual words, spoken words, and sounds (all $t > 3.0$, $p < 0.05$), with a trend for a difference from scenes ($t(18) = 1.89$, $p =$

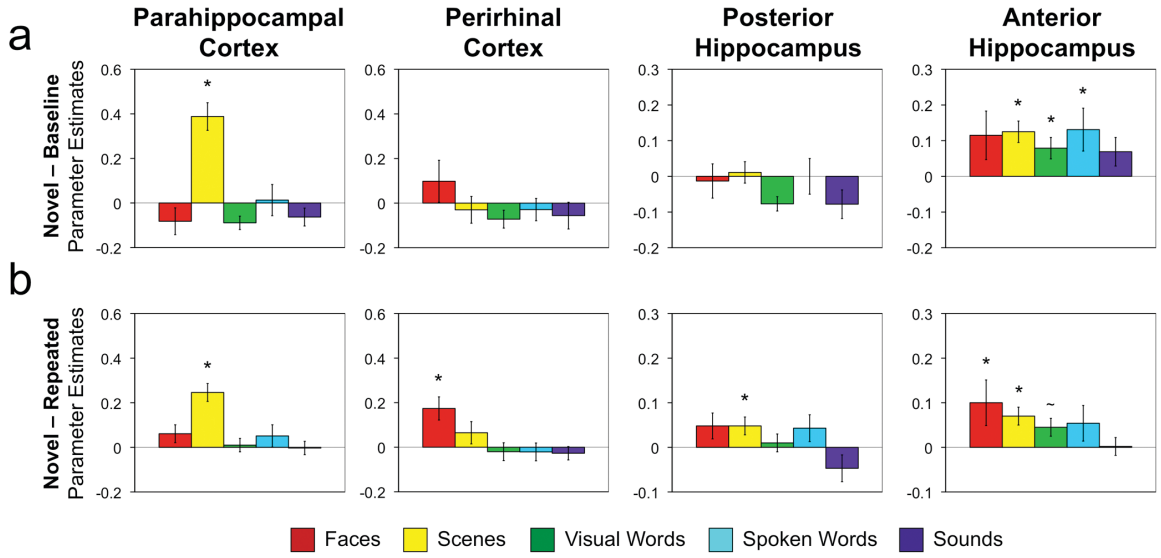
0.08). Finally, the apparent difference in content sensitivity in PHc and PRc was confirmed by a significant region × content interaction, both when considering responses to novel stimuli in isolation ($F(4,72) = 24.96$, $p < 0.001$) and when considering differences between novel and repeated stimuli ($F(4,72) = 4.90$, $p = 0.005$).

Within hippocampus, novelty-based activation was observed primarily in the anterior extent. Specifically, in anterior hippocampus, activation during novel stimulus blocks did not differ based on information content ($F < 1.0$; Figure 2.2a), and was significantly above baseline for scenes ($t(18) = 4.32$, $p < 0.001$), spoken words ($t(18) = 2.36$, $p < 0.05$), and visual words ($t(18) = 2.47$, $p < 0.05$). When comparing anterior hippocampus activation for novel relative to repeated stimuli, again there were no significant differences across stimulus content ($F(4,72) = 1.16$, $p = 0.33$; Figure 2.2b), with significant effects observed for face ($t(18) = 2.06$, $p = 0.05$) and scene ($t(18) = 3.58$, $p < 0.01$) stimuli.

By contrast, posterior hippocampal activation during novel stimulus blocks did not differ from baseline for any class of stimuli (all $t < 0.5$; Figure 2.2a). While there was a significant difference in posterior hippocampal activation when comparing novel relative to repeated scenes ($t(18) = 2.16$, $p = 0.04$), there was only a trend for an effect of content ($F(4,72) = 2.71$, $p = 0.06$; Figure 2.2b) and no pairwise comparison between content classes reached significance (all $t < 1.5$). The apparent difference in novelty-based responding in anterior and posterior hippocampus was supported by a main effect of region when considering responses to novel stimuli in isolation ($F(4,72) = 39.22$, $p < 0.001$), and when comparing differences between novel and repeated stimuli ($F(4,72) =$

20.13, $p < 0.001$); however, because this finding was not accompanied by a region × content interaction, interpretative caution is warranted. Finally, anterior hippocampus demonstrated a different pattern of content-sensitivity relative to MTL cortical regions, as reflected in a significant region × content interaction for novel stimuli ($F(4,72) = 38.59$, $p < 0.001$) and for the difference between novel and repeated stimuli ($F(4,72) = 7.54$, $p < 0.001$) when compared to activation in PHc, and trends for region × content interactions when compared to PRc activation (novel: $F(4,72) = 2.08$, $p = 0.10$; novel – repeated: $F(4,72) = 2.26$, $p = 0.09$).

Figure 2.2: Response to novel event content in anatomically defined MTL ROIs (PHc, PRc, posterior hippocampus, and anterior hippocampus). A, Parameter estimates representing activation during novel content blocks relative to baseline. Error bars represent standard error of the mean. Asterisks indicate significant differences from baseline ($p < 0.05$); tilde indicates a trend for difference ($p < 0.10$). B, Difference in parameter estimates between novel and repeated content blocks. Error bars represent standard error of the mean. Asterisks indicate significant differences between novel and repeated blocks ($p < 0.05$); tilde indicates a trend for difference ($p < 0.10$).



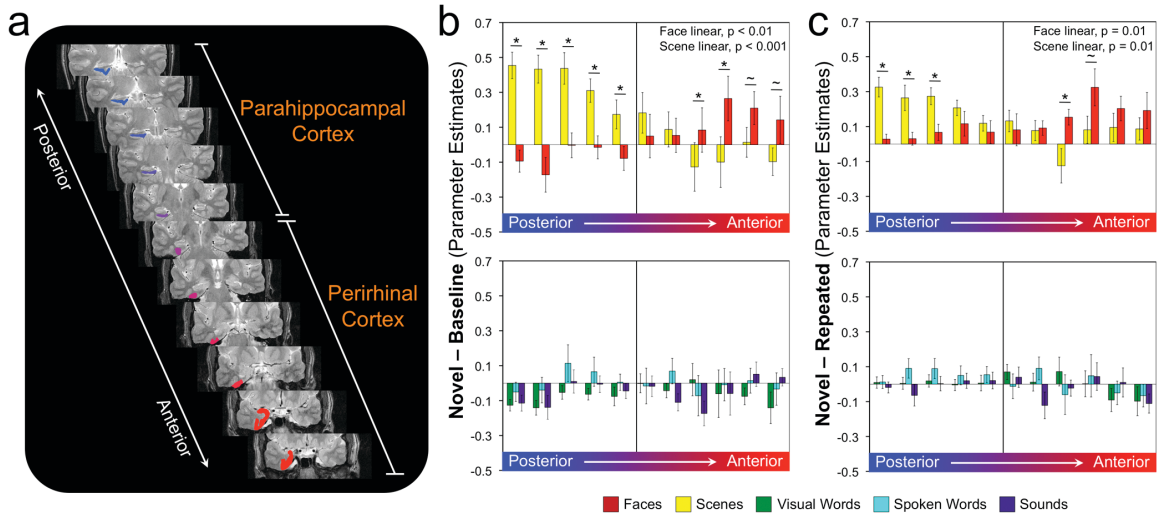
Distribution of content-sensitivity across PRc and PHc

The preceding results assume that content-sensitivity is uniform within anatomically defined MTL subregions. It is possible, however, that content-sensitivity does not adhere to discrete anatomical boundaries, but rather is distributed across anatomical subregions. This possibility would further suggest that content-sensitivity within anatomical subregions should be heterogeneous. To address this hypothesis, we examined content-sensitive novelty responses in MTL cortex and hippocampus as a function of position along the anterior-posterior axis of each structure (Figures 2.3-4). (For a similar analysis performed within PRc and PHc individually see Supplementary Results.)

Within MTL cortex, activation for novel stimuli demonstrated a significant main effect of content ($F(4,68) = 4.89$, $p < 0.005$) and an interaction between anterior-posterior position and content ($F(40,680) = 6.15$, $p < 0.001$; Figure 2.3b). The main effect of content was reflected by greater activation for novel scenes relative to spoken words, visual words, and sounds (all $t > 2.9$, $p < 0.005$). Moreover, responses to novel scenes demonstrated a significant linear trend along the anterior-posterior axis ($F(1,17) = 34.50$, $p < 0.001$), with maximal activation in the posterior MTL cortex and decreasing as one moves anteriorly. The opposite linear trend was observed for novel faces ($F(1,17) = 10.07$, $p < 0.01$), with maximal activation in anterior regions and decreasing as one moves posteriorly. No other class of content demonstrated significant linear trends along the anterior-posterior axis of MTL cortex (all $F < 1.6$).

When considering the difference in activation between novel and repeated blocks, a similar distribution was observed across MTL cortex, where there was a significant main effect of content ($F(4,68) = 4.10$, $p < 0.01$) as well as a significant interaction between anterior-posterior position and content ($F(40,680) = 2.80$, $p < 0.05$; Figure 2.3c). The effect of content in this case was reflected by greater difference between novel and repeated stimuli for scenes and faces relative to all other forms of stimulus content (all $t > 1.7$, $p < 0.05$). The interaction between position and content was reflected by a decreasing scene response from posterior to anterior ($F(1,17) = 8.36$, $p = 0.01$), and an increasing face response from posterior to anterior ($F(1,17) = 8.42$, $p = 0.01$). No other class of content demonstrated significant linear trends (all $F < 2.1$). Notably, these observed functional gradients in MTL cortex were not the result of individual differences in the anterior-posterior boundary between PRc and PHc across individuals (see Supplementary Results).

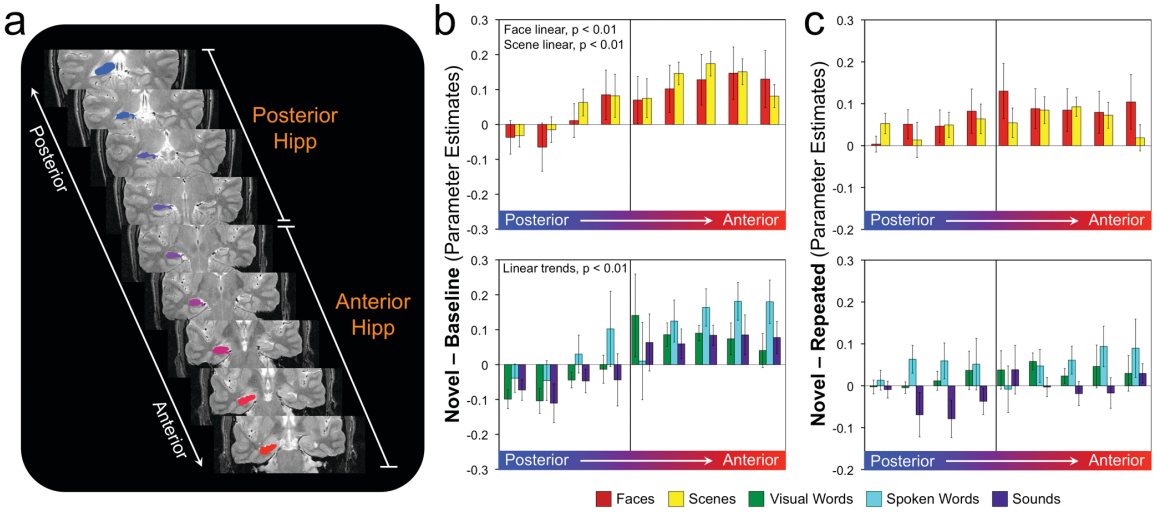
Figure 2.3: Responses to novel event content along the anterior-posterior axis of the parahippocampal gyrus. A, Coronal slices through parahippocampal gyrus with anatomical ROIs represented as color-coded regions in the right hemisphere. B, Top, Parameter estimates for novel faces and scene blocks relative to baseline in each of the anatomically defined ROIs along the anterior-posterior axis of MTL cortex. Bottom, Parameter estimates for novel visual word, sound, and spoken word blocks relative to baseline. C, Top, Parameter estimates for novel – repeated face and scene blocks along the anterior-posterior axis of MTL cortex. Bottom, Novel – repeated parameter estimates for visual words, sounds, and spoken words. Error bars represent standard error of the mean. Asterisks indicate significant pairwise differences ($p < 0.05$); tilde indicates a trend for difference ($p < 0.10$).



Distribution of content-sensitivity across anterior and posterior hippocampus

We performed similar analyses examining activation for novel stimuli along the anterior-posterior axis of the hippocampus (Figure 2.4). Within hippocampus, we observed a main effect of anterior-posterior position ($F(8,136) = 6.88$, $p < 0.001$), but did not observe an effect of content ($F(4,68) = 1.01$, $p = 0.39$) or a content × position interaction ($F(32,544) = 1.02$, $p = 0.40$). Significant linear trends were observed for all content classes (all $F(1,17) > 11.81$, $p < 0.01$), with activation for novel stimuli increasing from posterior to anterior hippocampus (Figure 2.4b). When considering the difference in activation for novel and repeated stimuli (Figure 2.4c), only a trend for an effect of position ($F(8,136) = 2.67$, $p = 0.06$) was observed, reflected greater novel – repeated differences in the four anterior-most hippocampal positions compared to the three most posterior positions (all $t > 2.4$). These differences were not reflected in a significant linear trend for any class of content (all $F(1,17) < 2.35$). For similar analyses performed at the level of individual participants see Supplementary Results.

Figure 2.4: Responses to novel event content along the anterior-posterior axis of the hippocampus. A, Coronal slices through hippocampus with anatomical ROIs represented as color-coded regions in the right hemisphere. B, Top, Parameter estimates for novel face and scene blocks relative to baseline in each of the anatomically defined ROIs along the anterior-posterior axis of hippocampus. Bottom, Parameter estimates for novel visual word, sound, and spoken word blocks relative to baseline. C, Top, Parameter estimates for novel – repeated face and scene blocks along the anterior-posterior axis of hippocampus. Bottom, Novel – repeated parameter estimates for visual words, sounds, and spoken words. Error bars represent standard error of the mean. Asterisks indicate significant pairwise differences ($p < 0.05$); tilde indicates a trend for difference ($p < 0.10$).



Multivariate pattern classification in MTL subregions

Using multivoxel pattern analysis (MVPA), we examined whether each MTL subregion carries sufficient information about a specific class of content to distinguish it from other categories of information, providing an additional measure of content-sensitivity distinct from standard univariate measures. For each region—PHc, PRc, anterior hippocampus, and posterior hippocampus—we trained a classifier to differentiate between novel stimulus blocks from each of the five content classes and tested classification accuracy using a cross-validation procedure. Overall classification accuracy (Figure 2.5, gray bars) was significantly above chance (20%) using data from each MTL subregion (all $t(18) > 4.67$, $p < 0.001$).

We also determined the number of participants whose overall classification performance lay significantly outside of an assumed binomial distribution of performance given a theoretical 20% chance-level accuracy. Overall performance in the top 5% of the binomial distribution was considered above chance. The binomial test revealed that the number of participants with above chance classification performance was greater in PRc ($n = 18$) and PHc ($n = 18$) than in anterior ($n = 13$) and posterior hippocampus ($n = 10$). Superior classification performance in MTL cortical regions relative to hippocampus was further revealed by repeated measures ANOVA assessing the difference in classification accuracy across regions. A significant main effect of region was observed when comparing classification accuracy for anterior hippocampus with PRc ($F(1,18) = 42.84$, $p < 0.001$) and PHc ($F(1,18) = 90.53$, $p < 0.001$), and when comparing classification

accuracy for posterior hippocampus with PRc ($F(1,18) = 23.85$, $p < 0.001$) and PHc ($F(1,18) = 50.68$, $p < 0.001$).

We also considered individual classification accuracies for each class of information content to determine whether certain classes of content evoked more consistent and meaningful patterns of activation within MTL subregions than others, and whether classification of individual classes of content differed by region (Figure 2.5). While classification accuracy in PHc was significantly above chance for all classes of stimulus content (all $t(18) > 2.91$, $p < 0.01$), repeated measures ANOVA revealed a significant main effect of content ($F(4,72) = 53.23$, $p < 0.001$), with classification accuracy for novel scenes being greater than all other classes of content (all $t(18) > 7.51$, $p < 0.001$) and classification accuracy for novel faces being greater than that for visual words, spoken words, and sounds (all $t(18) > 3.71$, $p < 0.001$).

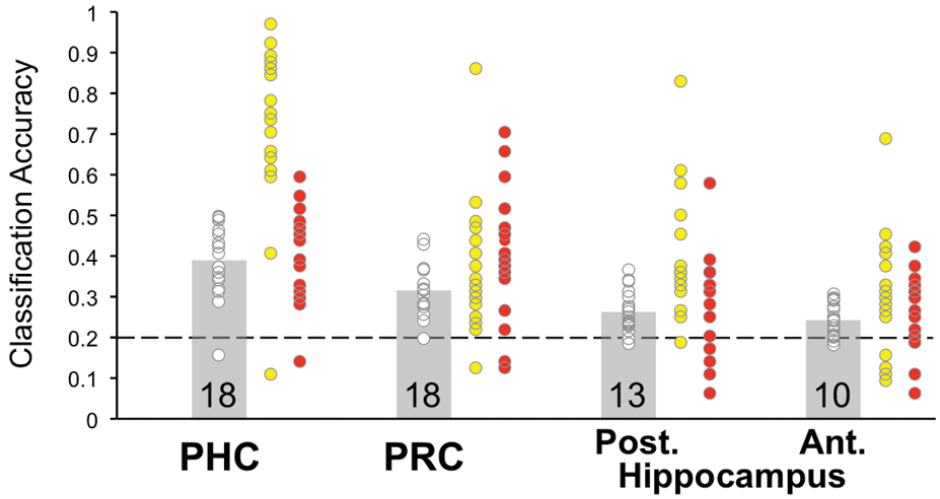
In PRc, classification accuracy exceeded chance for all stimulus classes (all $t > 2.90$, $p < 0.05$) except spoken words ($t(18) = 1.81$, $p = 0.09$). A significant main effect of content on classification accuracy was also observed in PRc ($F(4,72) = 6.31$, $p < 0.001$), with greater classification accuracy for novel faces and scenes relative to visual words and spoken words (all $t > 2.60$, $p < 0.05$), and greater accuracy for novel sounds relative to visual words ($t(18) = 2.29$, $p = 0.04$). When considering classification accuracies for individual classes of content across PHc and PRc, a significant region × content interaction was observed ($F(4,72) = 25.07$, $p < 0.001$).

In posterior hippocampus, repeated measures ANOVA revealed a main effect of content on classification accuracies ($F(4,72) = 6.74$, $p < 0.001$), with only the

classification of novel scenes being significantly above chance ($t(18) = 5.20$, $p < 0.001$) and being significantly better than classification of every other class of content (all $t(18) > 2.88$, $p < 0.05$). In contrast, classification accuracies in anterior hippocampus were above chance for novel faces, scenes, and sounds (all $t > 2.56$, $p < 0.05$). A main effect of content on classification accuracies was further observed in anterior hippocampus ($F(4,72) = 2.91$, $p < 0.05$), with lower classification accuracy for visual words compared to all other stimulus classes (all $t > 2.51$, $p < 0.05$). When comparing classification accuracies for individual classes of content across anterior and posterior hippocampus, we observed trends for a main effect of region ($F(1,18) = 3.97$, $p = 0.06$) and a region × content interaction ($F(4,72) = 2.49$, $p = 0.08$), suggesting modest differences in the representation of novel information content across the long-axis of the hippocampus.

We also investigated the possibility that higher classification performance in MTL cortical subregions was driven primarily by their ability to discriminate preferred content identified in the univariate analyses (i.e., novel faces in PRc and novel scenes in PHc). Three follow-up analyses interrogated subregional pattern classification performance: one analysis omitting novel faces from classification, one omitting novel scenes, and one omitting both novel faces and scenes from classification training and testing. Importantly, classification performance in PHc and PRc remained greater than that of hippocampal subregions despite the omission of preferred subregional content (Supplementary Figures S2.1-3). For additional details on these analyses see Supplementary Results.

Figure 2.5: MVPA classification accuracy in anatomically defined MTL subregions. Top, Overall classification accuracy across the five classes of event content in each anatomical region. Gray bars indicate the overall mean classification accuracy across participants. Chance classification performance is indicated by the dashed line. White circles represent overall classification accuracy for individual participants. Numbers indicate the number of individual participants with above chance classification accuracy. Yellow circles indicate individual participant accuracies for scenes, and red circles individual participant accuracies for faces. Bottom, Classification accuracies for each content class expressed as the proportion of hits across participants. Significant classification accuracy is indicated by bold/italics.



	PHC	PRC	Post. Hippocampus	Ant. Hippocampus
Faces	***0.40***	***0.41***	0.25	***0.25***
Scenes	***0.72***	***0.36***	***0.39***	***0.29***
Visual Words	***0.28***	***0.25***	0.21	0.18
Spoken Words	***0.27***	0.24	0.25	0.24
Sounds	***0.28***	***0.32***	0.23	***0.25***

Multivariate pattern confusion in MTL subregions

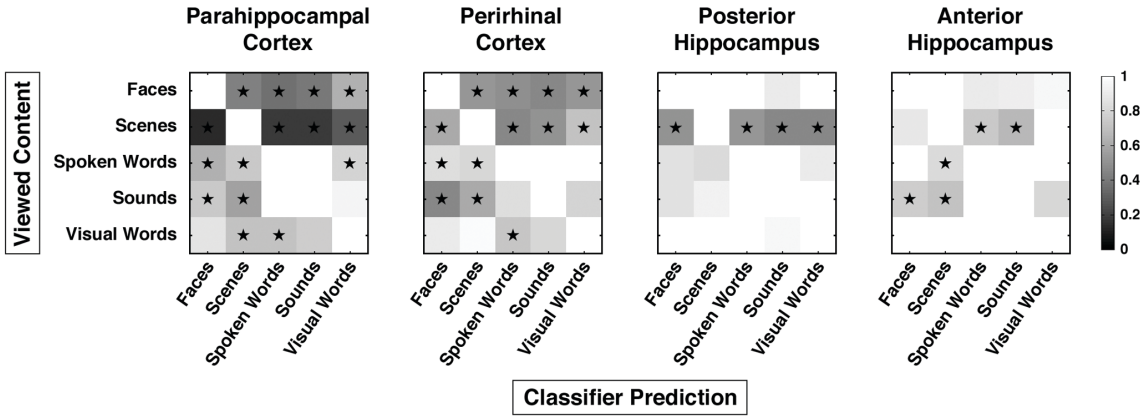
The preceding results suggest a substantial difference between MTL cortical subregions and hippocampus in their ability to classify different forms of stimulus content. Classification accuracy alone, however, provides only a limited view of representational coding in MTL subregions. We also examined MVPA classifier confusion matrices to determine how often the classifier confused different forms of stimulus content, which provided a measure of the similarity between voxel patterns evoked by different forms of event content. Specifically, we constructed MVPA confusion matrices for each anatomical region indicating how often voxel patterns for each stimulus class were classified correctly, and if incorrectly classified, what form of content a given voxel pattern was labeled as (Figure 2.6).

In PHc, only voxel patterns evoked by novel faces and scenes were distinct from patterns evoked by other classes of content, as indicated by lower cross-content confusion values than would be expected by chance (all $p < 10^{-3}$). By comparison, voxel patterns evoked by spoken words were significantly dissimilar from those evoked by faces, scenes, and visual words (all $p < 10^{-3}$) but not sounds. Similarly, voxel patterns evoked by novel sounds were dissimilar from those evoked by faces and scenes (all $p < 10^{-3}$) but not spoken words, consistent with an overlapping representation of spoken words and sounds in PHc that is distinct from novel face and scene visual content. Voxel patterns evoked by novel visual words were dissimilar from those of spoken words and scenes (all $p < 10^{-3}$).

Similar to PHc, voxel patterns evoked by novel faces and scenes in PRc were distinct from those evoked by other classes of content (all $p < 10^{-3}$). Voxel patterns evoked by novel spoken words were distinct from those of faces and scenes (all $p < 10^{-3}$) but not sounds or visual words. The same pattern was observed for responses to novel sounds, which were distinct from those evoked by faces and scenes (all $p < 10^{-3}$) but not spoken or visual works. Voxel patterns evoked by novel visual words were distinct from those of spoken words ($p < 10^{-3}$) but not faces, scenes, or sounds. Together, evidence from the classifier confusion matrices indicates that PHc and PRc both contain representationally distinct codes for novel faces and scenes, while the representation of different forms of auditory content is highly overlapping.

In anterior hippocampus, voxel patterns evoked by visual and auditory content were somewhat distinct from each other, with voxel patterns evoked by novel spoken words and novel sounds differing from those evoked by scenes (all $p < 10^{-3}$), and faces in the case of novel sounds ($p < 10^{-3}$). Voxel patterns evoked by different forms of visual event content in anterior hippocampus did not significantly differ based on the criterion threshold, reflecting less distinctiveness between the representation of visual content in this region. In contrast, the voxel patterns evoked by novel scenes in posterior hippocampus differed from all other classes of event content (all $p < 10^{-3}$), indicating a distinct representation of scene information in this region. The voxel patterns evoked by novel faces, spoken words, sounds, and visual words in posterior hippocampus did not significantly differ from one another based on the criterion threshold.

Figure 2.6: MVPA classifier confusion matrices in anatomically defined MTL subregions. Each row displays classifier performance on the test patterns drawn from each of the five content classes. For a given content class, the cells in each row indicate the proportion of trials that those test patterns were classified as each of the five content classes normalized to the proportion of correctly classified test patterns for that stimulus class. Therefore, values along the diagonal are always equal to 1. Grayscale intensity along each row indicates confusability relative to the correct class of content. Test patterns that were highly confusable with the correctly classified content would yield values close to 1 (off-diagonal white squares). Stars indicate when classifier confusion values lay outside of the confidence intervals derived from null distributions of classification performance based on Monte Carlo simulation. The alpha-level of the confidence intervals was chosen based on Bonferroni correction for each of the statistical tests performed across all anatomical ROIs ($\alpha = 10^{-3}$).



Representational similarity analysis in MTL subregions

The preceding classifier confusion analysis characterizes when voxelwise patterns evoked by each form of event content are different from other content classes. This analysis, however, does not directly assess whether such differences arise solely from the distinct representation of exemplars from different content classes (i.e., low cross-class similarity), or whether such differences also result from highly similar representations of exemplars within a given stimulus class (i.e., high within-class similarity). To further interrogate the pattern of results observed in our MVPA analyses, we employed RSA (Kriegeskorte and Bandettini 2007; Kriegeskorte et al. 2008) to measure the representational distance between evoked responses for each content class exemplar and all other exemplars in the experiment. The correlation distances between exemplars were visualized in two-dimensions using MDS (Edelman 1998; Kriegeskorte et al. 2008). This characterization of the data enabled us to measure the voxelwise pattern similarity between any two exemplars as their linear distance in the two-dimensional space. Thus, RSA not only provides a means to directly compare the representational similarity between exemplars from different forms of content, but also provides a means to directly measure the representational similarity of exemplars within a class.

Moreover, RSA extends upon the MVPA classifier confusion analysis by measuring within- and cross-class representational similarity not only within individual MTL subregions but also as a function of the anterior-posterior position along MTL cortex and hippocampus. Importantly, the univariate analyses identified gradients of content-sensitive responding in the MTL. It is possible that the multivoxel patterns most

important for representing any given form of event content might also be distributed in a non-uniform manner within or across MTL subregions. If true, RSA performed across an entire region might fail to find distinct representational codes for different classes of content, while a consideration of the representational codes along the anterior-posterior axis might demonstrate clear distinctions between content classes. By constructing representational similarity matrices for each anatomical ROI segment of MTL cortex and hippocampus, we sought to more precisely identify positions along the anterior-posterior MTL axis where voxel patterns evoked by different forms of event content are distinct. This analysis approach yielded a rich set of data, and here, we have focused our reporting on the set of findings that help elucidate the representational codes underlying our MVPA findings. The full results of the representational similarity analyses are available from the authors upon request.

Distinct face and scene representations in PHc and PRc

First, we considered face and scene representation within MTL cortical subregions, which were revealed to be distinct from other forms of content using MVPA. RSA revealed smaller within-class linear distances between novel face exemplars and between novel scene exemplars in PHc than would be expected by chance (Figure 2.7a; all $p < 10^{-2}$), indicating a highly clustered within-class representational structure for both forms of event content within this region. The distinct representation of scene content in PHc was further supported by significantly larger cross-class linear distances between novel scene exemplars and novel face, spoken word, and sound exemplars (all $p < 10^{-2}$).

In contrast, within-class linear distances between novel face exemplars and between novel scene exemplars in PRc did not reach significance (all $p > 10^{-2}$), despite evidence from MVPA suggesting highly distinct face and scene voxel patterns in the region. Furthermore, we observed significantly larger cross-class distance only between voxel patterns evoked by novel scenes and those evoked by spoken words in PRc ($p < 10^{-2}$). When examining face and scene representation as a function of anterior-posterior position along the axis of MTL cortex, we found that the distinctiveness of face and scene representations was localized primarily in the most posterior positions (Figure 2.8a). In the three posterior-most ROIs (corresponding to the posterior aspect of PHc), we observed within-class distances between individual face and individual scene exemplars that were smaller than expected by chance (all $p < 10^{-2}$), indicating a distinct representational code for both face and scene content classes in posterior MTL cortex. The distinct representation of scene content continued anteriorly, with significantly smaller within-class distances for individual scene exemplars in the five most posterior ROIs in MTL cortex (all $p < 10^{-2}$). Moreover, in the posterior-most ROI, we observed significantly larger cross-class distances between voxel patterns evoked by faces and scenes exemplars and those evoked by each other form of content except visual words (all $p < 10^{-2}$). These significantly larger cross-class distances disappeared one-by-one as we moved anteriorly (Figure 2.8a), being absent by the middle slice in MTL cortex corresponding to posterior PRc. Finally, we observed significant within-class clustering of face exemplars in the anterior-most ROI corresponding to PRc ($p < 10^{-2}$), although this

effect was absent in every other anterior MTL cortical ROI; we did not observe significant within-class clustering of scene exemplars in any anterior MTL cortical ROIs.

Figure 2.7a: Neural pattern distances between novel content exemplars visualized by multi-dimensional scaling in MTL cortex. Each content class exemplar (i.e., a novel mini-block) is represented by a colored dot in the panels for each MTL subregion. Dots placed close together in the two-dimensional space indicate that those two exemplars were associated with a similar pattern of activation. Dots placed farther apart indicate that those two exemplars were associated with more distinct activation patterns. The tables below each plot indicated the mean within-class linear distance for each content class and the mean cross-class linear distance between each pair of novel content. Bolded values indicated when linear distances lay outside of confidence intervals derived from null distributions of within-class and cross-class linear distances based on Monte Carlo simulation. The alpha-level of the confidence intervals was chosen based on Bonferroni correction for each of the statistical tests performed for all anatomical ROIs ($\alpha = 10^{-2}$). Crosses indicate when linear distances were significantly smaller than expected by chance and reflect greater similarity in the activation patterns evoked by content class exemplars. Asterisks indicate when linear distances were significantly larger than expected by chance and reflect more distinct representation of individual exemplars.

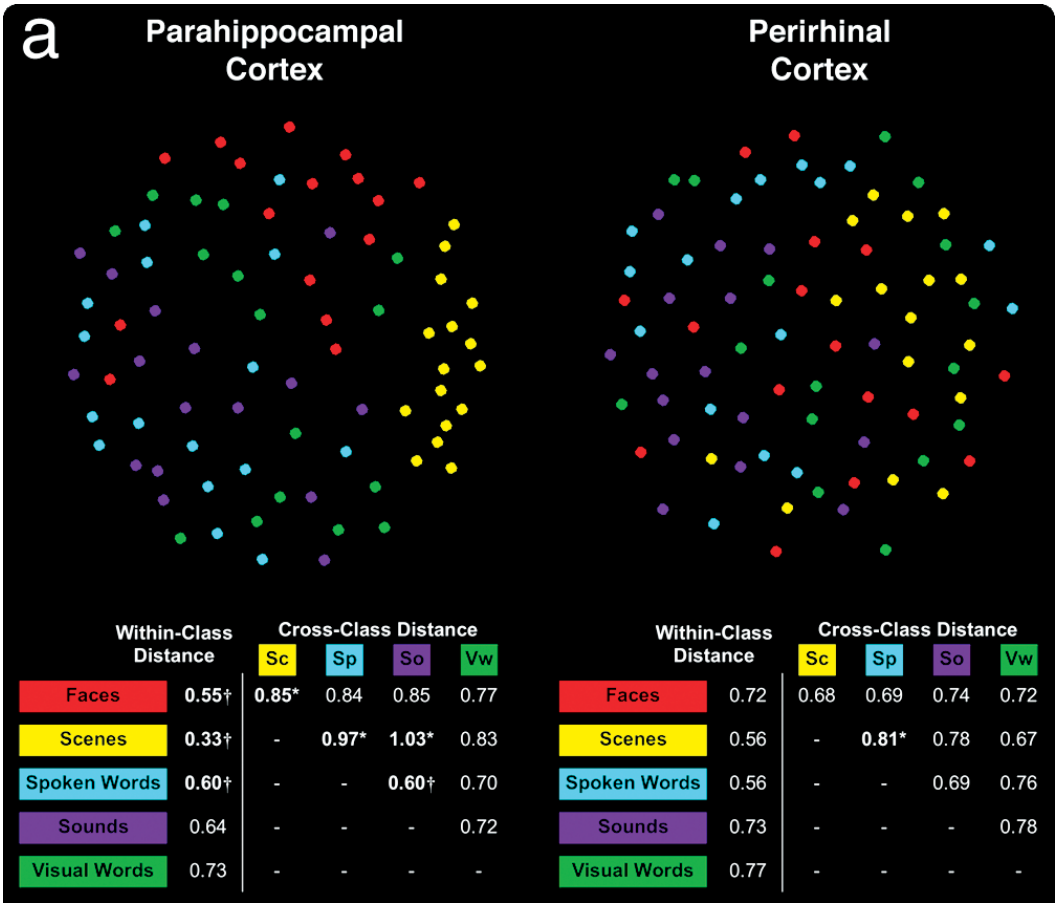
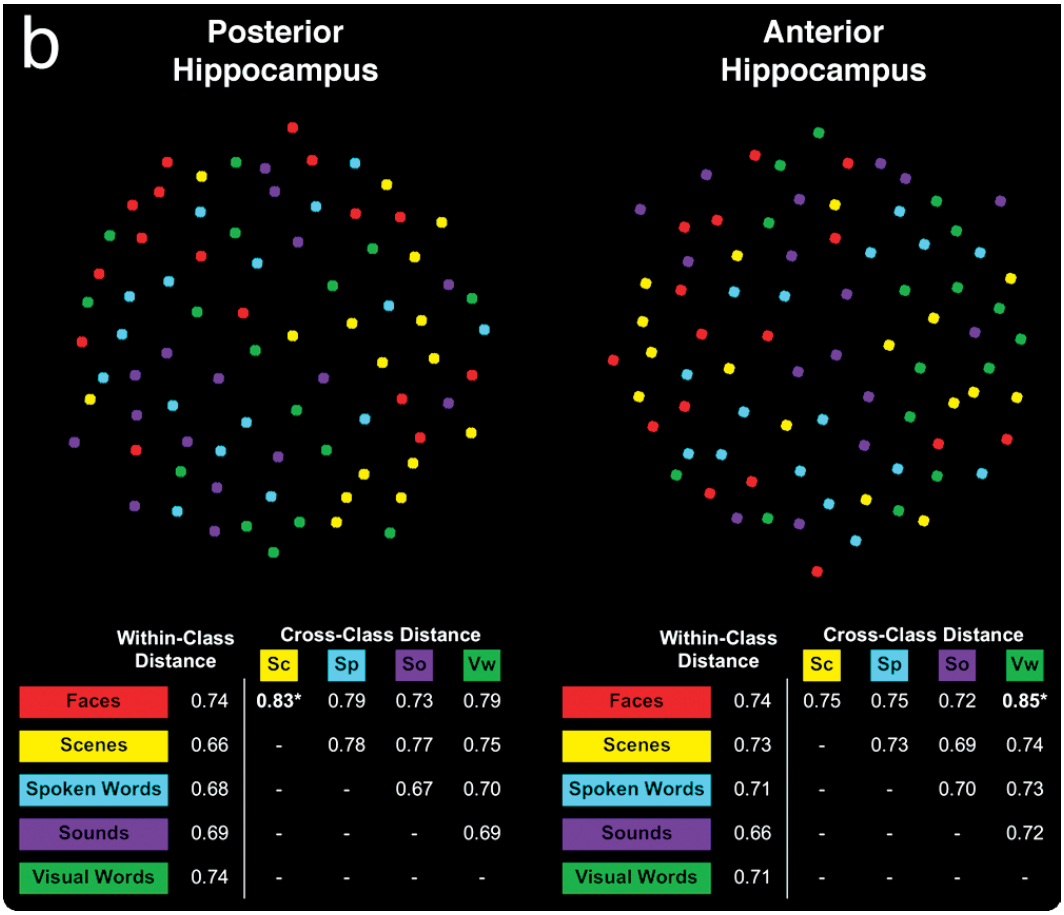


Figure 2.7b: Neural pattern distances between novel content exemplars visualized by multi-dimensional scaling in hippocampus. Each content class exemplar (i.e., a novel mini-block) is represented by a colored dot in the panels for each MTL subregion. Dots placed close together in the two-dimensional space indicate that those two exemplars were associated with a similar pattern of activation. Dots placed farther apart indicate that those two exemplars were associated with more distinct activation patterns. The tables below each plot indicated the mean within-class linear distance for each content class and the mean cross-class linear distance between each pair of novel content. Bolded values indicated when linear distances lay outside of confidence intervals derived from null distributions of within-class and cross-class linear distances based on Monte Carlo simulation. The alpha-level of the confidence intervals was chosen based on Bonferroni correction for each of the statistical tests performed for all anatomical ROIs ($\alpha = 10^{-2}$). Crosses indicate when linear distances were significantly smaller than expected by chance and reflect greater similarity in the activation patterns evoked by content class exemplars. Asterisks indicate when linear distances were significantly larger than expected by chance and reflect more distinct representation of individual exemplars.



	Within-Class Distance	Cross-Class Distance			
		Sc	Sp	So	Vw
Faces	0.74	**0.83***	0.79	0.73	0.79
Scenes	0.66	-	0.78	0.77	0.75
Spoken Words	0.68	-	-	0.67	0.70
Sounds	0.69	-	-	-	0.69
Visual Words	0.74	-	-	-	-

	Within-Class Distance	Cross-Class Distance			
		Sc	Sp	So	Vw
Faces	0.74	0.75	0.75	0.72	**0.85***
Scenes	0.73	-	0.73	0.69	0.74
Spoken Words	0.71	-	-	0.70	0.73
Sounds	0.66	-	-	-	0.72
Visual Words	0.71	-	-	-	-

Scene representation in hippocampus

MVPA of hippocampal responses revealed accurate discrimination of voxel patterns evoked by scenes from those evoked by each other class of content. Moreover, the classifier confusion analysis revealed that voxelwise responses to scenes was distinct from other forms of content in posterior hippocampus, which otherwise demonstrated high confusion between all other forms of content.

Using RSA to investigate voxel patterns evoked in the entire posterior hippocampal region, we did not find evidence for a distinct representation of scene content as within-class distance between individual scene exemplars did not reach our criterion threshold (Figure 2.7b; $p > 10^{-2}$). Moreover, significant cross-class linear distances were only observed between voxel patterns evoked by scenes and those evoked by faces in posterior hippocampus ($p < 10^{-2}$) but not other forms of event content (all $p > 10^{-2}$). Although we found few effects of representational distance to explain the distinctiveness of scenes in our MVPA analysis when examining the posterior hippocampus as a whole, we considered whether such distinct coding of scene content might be found in specific locations along the hippocampal axis (Figure 2.8b). Indeed, significantly smaller within-class distances were present between individual scene exemplars in the second and third posterior-most ROIs of hippocampus (all $p < 10^{-2}$). In both ROIs, this effect was accompanied by a significantly larger cross-class distance between scene and faces exemplars (all $p < 10^{-2}$), while a significantly larger cross-class distance between scenes and visual words was additionally observed in the second posterior-most ROI ($p < 10^{-2}$). Such representationally distinct coding of scene content

was not observed in the anterior-most ROIs of hippocampus, nor were there significant within-class linear distances for any other class of stimuli in any portion of hippocampus (all $p > 10^{-2}$). Together, these observations in hippocampus show that the distinctive representation of scene content is explained primarily by the presence of a consistent spatial code in the posterior extent of this region.

Representation of auditory content in MTL

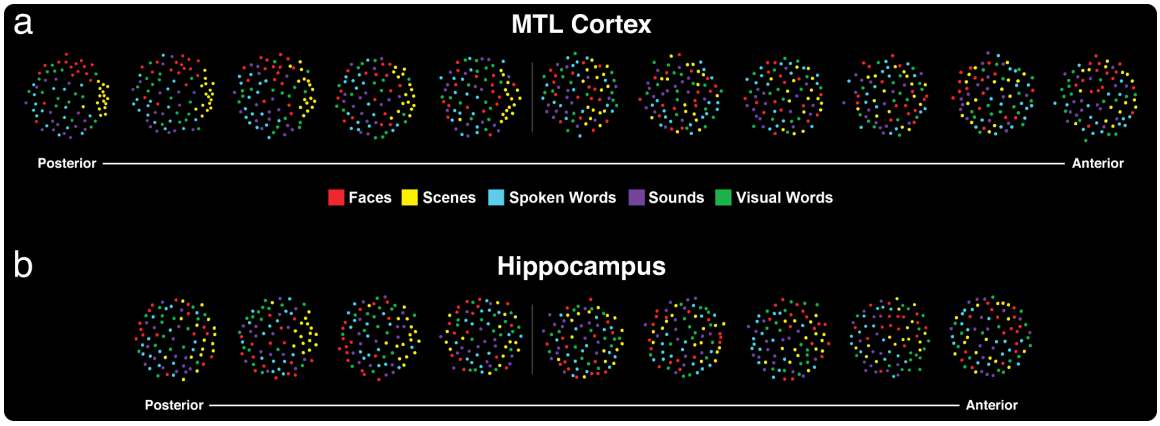
MVPA analysis revealed that voxel patterns evoked by spoken words and sounds in PHc and PRc are distinct from patterns evoked by visual forms of content, but not from one another. One possibility is that, being the only forms of auditory stimuli presented to the participants, spoken words and sounds might be distinguished from other visual content based on sensory modality. However, this finding does not necessarily entail that these forms of auditory content share a common representational structure in PHc and PRc. To directly address how auditory content is represented in MTL cortex, we compared voxel patterns evoked by individual spoken word and sound exemplars using RSA.

In PHc, within-class linear distances were significantly smaller than would be expected by chance for spoken word (Figure 2.7a; $p < 10^{-2}$) but not sound exemplars. However, there was significant within-class clustering of individual sound exemplars in the third posterior-most ROI in MTL cortex (Figure 2.8a; $p < 10^{-2}$). We also found that the voxel patterns evoked by spoken words and sounds showed cross-class distances that were significantly smaller than chance ($p < 10^{-2}$), indicating a highly overlapping

representation of these content forms in PHc. The overlapping representation of auditory content was evident in the four posterior-most ROIs of PHc, with significantly smaller cross-class distances between spoken word and sound exemplars than would be expected by chance (all $p < 10^{-2}$). Moreover, we found that auditory content was distinct from scene content in the four posterior-most ROIs and from face content in the two posterior-most ROIs, as revealed by significantly larger cross-class distances between both forms of auditory content and face and scene visual content (all $p < 10^{-2}$). In contrast, voxel patterns evoked by spoken words and sounds in PRc did not demonstrate significant within-class representational similarity, nor did they demonstrate significant cross-class clustering with one another (all $p > 10^{-2}$), suggesting that these two forms of content do not share a common representational structure in PRc. This pattern of results was true both when RSA was performed for PRc as a whole and when it was performed on the anterior-most MTL cortical ROI corresponding to PRc. Together these findings suggest that representations of auditory content are highly overlapping throughout PHc, and are increasingly distinguished from visual content in the posterior extremity of MTL cortex. When we examined the voxel patterns evoked by spoken words and sounds across the anterior-posterior hippocampal axis, we found significantly smaller cross-class distances between spoken word exemplars and sound exemplars in the second and third anterior-most ROIs and in the posterior-most ROI (Figure 2.8b; all $p < 10^{-2}$). The MVPA confusion matrices had previously indicated a high overall degree of classifier confusion in hippocampus, but did not identify the precise nature of poor performance for any particular class of content. Here, the use of RSA within segmented hippocampal ROIs

revealed that different forms of auditory content were highly confusable because they evoked similar distributed patterns of response. However, unlike PHc, this effect was not accompanied by consistently larger cross-class distance between face and scene visual content (all $p > 10^{-2}$).

Figure 2.8: Neural pattern distances between novel content exemplars visualized by multi-dimensional scaling along the anterior-posterior axis of A, MTL cortex and B, hippocampus. Each content class exemplar (i.e., a novel mini-block) is represented by a colored dot in the panels for each MTL subregion. Dots placed close together in the two-dimensional space indicate that those two exemplars were associated with a similar pattern of activation. Dots placed farther apart indicate that those two exemplars were associated with more distinct activation patterns. Results tables for each plot are available from the authors upon request.



Discussion

Whether MTL subregions make distinct contributions to episodic memory remains a topic of considerable debate. In the present study, we combined hr-fMRI (Carr et al. 2010) with both univariate and multivariate statistical measures to investigate whether event content differentiates the function of hippocampus and MTL cortical subregions. First, our findings revealed a distributed code for event content in PRc and PHc that crosses anatomical boundaries, despite significant differences in responding to novel versus repeated items for only one stimulus class in each region (novel faces in PRc and novel scenes in PHc). In particular, multivariate analysis of responses to novel content showed that PRc and PHc both contain distinct representational codes for faces and scenes. Second, we observed a dissociation in content representation along the anterior-posterior axis of the hippocampus. Anterior hippocampus demonstrated peak amplitude responses that were content-general; moreover, the spatial pattern of response in this region did not discriminate between different forms of event content. In contrast, posterior hippocampus did not demonstrate significant peak amplitude responses for novel stimuli from any content class, but did show a distributed coding of scene content that was representationally distinct from other content classes. By taking advantage of the complementary aspects of univariate and multivariate approaches, the present data provide new insights into the nature of representational coding in the MTL.

Content representation in MTL cortex

While many studies have focused on content-based dissociations between PRc and PHc (Pihlajamaki et al. 2004; Lee, Buckley et al. 2005; Sommer et al. 2005; Lee et al. 2008; see also Dudukovic et al. 2010), several recent reports have observed encoding responses for visual object and visuospatial information in human PRc (Buffalo et al. 2006; Litman et al. 2009; Preston et al. 2010) as well as PHc (Bar and Aminoff 2003; Aminoff et al. 2007; Bar et al. 2008; Litman et al. 2009). In the present study, PRc novelty responses were maximal for faces, while PHc demonstrated maximal novelty responses to scenes, consistent with previous reports of content-based dissociations between PRc and PHc. However, when examining the distribution of novelty-based responses across MTL cortex, a response to novel scenes was observed in posterior PRc, indicating that processing of scene information is not unique to PHc. Notably, these representational gradients were evident at the level of individual participants (see Supplementary Results).

These findings complement recent reports that demonstrated greater responses to visual object content in anterior PRc and visuospatial content in posterior PHc, with a mixed response to scene, object, and face content in a transitional zone at the border between PHc and PRc (Litman et al. 2009; Staresina et al. 2011). Such findings have led to the conclusion that discrete functional boundaries do not exist in MTL cortex and the further speculation that selective responses to a single content class are limited to the anterior and posterior extents of MTL cortex. However, as discussed below, our

multivariate findings suggest that distributed representations of event content can be observed at extreme ends of MTL cortex.

MVPA revealed significant differentiation of event content in PRc and PHc, when treated as two separate regions, both across the group and in the majority of participants. Importantly, successful classification was observed even when preferred content (i.e., novel faces and scenes) was removed from classifier training and testing (see Supplementary Results). Further consideration of the classifier confusion matrices showed that PRc and PHc both maintain distinct codes for face and scene content, as those stimuli were significantly differentiated from all other forms of event content. However, as indicated by the present findings and prior reports (Litman et al. 2009; Staresina et al. 2011), clear functional boundaries between PHc and PRc may not exist. These observations of a mixed representation of event content as revealed by MVPA may inadvertently result from the fact that this analysis considered these regions as two distinct areas. Critically, in the present study, we used RSA to examine how patterns of activation represent different forms of event content both within individual anatomically defined PRc and PHc and as a function of anterior-posterior position along the axis of MTL cortex.

In PRc, RSA revealed significant within-class clustering for face content in the anterior-most portion of this region, and while it did not reach our threshold for correction for multiple comparisons, there was also evidence for distinctive scene representations both in PRc as a whole ($p = 0.004$) and in the most posterior aspect of PRc ($p = 0.008$) as revealed by MVPA. Moreover, the MVPA confusion matrices showed

clear distinctions between the representation of face and scene content in PRc. MVPA may have emphasized distinctive face and scene codes by placing greater weight on voxels from the anterior and posterior regions of PRc, making these effects more apparent in the classifier confusion matrices. Our RSA findings are informative, however, in that they converge with our univariate findings in PRc, demonstrating a predominately face-selective response in anterior PRc combined with a scene-sensitive response in the posterior aspect of this region.

When we considered PHc as a whole region in the RSA analysis, we observed significant within-class clustering of multiple forms of content, including faces and scenes. Moreover, face and scene representations were significantly distinct from other stimulus classes. When we considered patterns of activation within individual ROIs along the anterior-posterior extent of PHc, we noted that the distinctive representation of faces and scenes was most prominent in the posterior aspect of the region and gradually became less distinct as one moved to the anterior portion of the region. Notably, the distinctive representation of faces was observed in PHc despite the absence of an above-baseline response for faces in the univariate analysis. Similarly, while univariate analysis showed no evidence for above-baseline responding to auditory content in PHc, RSA revealed a representation of auditory content that was distinct from visual content, again most evident in the posterior extent of PHc. The fact that representational distinctions were observed for multiple content classes in the posterior PHc runs counter to the hypothesis that content coding would be most scene-selective at this extreme end of PHc. Thus, the present data indicate that the distributed representation of event content in MTL

cortex extends beyond a transitional zone at the border between PRc and PHc (Litman et al. 2009; Staresina et al. 2011) and is also evident in posterior PHc.

It is possible that the differences in novelty-based responding observed in MTL cortex result from differences in low-level perceptual features of the stimuli used in the present study rather than differences based on encoding of conceptual information about different categories of stimuli. Because one of our goals was to assess MTL responses to a wide variety of auditory and visual event content, we did not control for perceptual differences between classes of stimuli. However, representational gradients for visual object and visuospatial information are evident in MTL cortex even when perceptual features are equated across content domains (Staresina et al. 2011). Moreover, previous work examining content representation in ventral temporal cortex has shown that patterns of non-maximal responses that discriminate between different forms of event content are not dependent on the low-level characteristics of the stimuli, such as luminance, contrast, and spatial frequency (Haxby et al. 2001). Collectively, these converging findings suggest that distributed coding of event content observed here extends beyond simple differences in the perceptual features of events.

Content representation in anterior hippocampus

Several observations of functional dissociations between anterior and posterior hippocampus are present in the neuroimaging literature (Prince et al. 2005; Strange et al. 2005; Chua et al. 2007; Awipi and Davachi 2008; Poppenk et al. 2010). However, few studies have considered the possible representational basis for such dissociations. The

present findings indicate that dissociations between anterior and posterior hippocampus may result from differences in content-based representational coding between these two regions.

A prevailing view of MTL function proposes that hippocampus plays a domain-general role in episodic memory by binding content-specific inputs from MTL cortex into integrated memory representations (Davachi 2006; Manns and Eichenbaum 2006; Diana et al. 2007). Consistent with this view, domain-general encoding and retrieval responses have been observed in hippocampus relative to content-specific processing in MTL cortex (Awipi and Davachi 2008; Staresina and Davachi 2008; Diana et al. 2010). Human electrophysiological evidence also suggests an invariant representation of perceptual information in hippocampal neurons relative to MTL cortex (Quian Quiroga et al. 2009). Our findings indicate that such domain-general memory functions may be specific to the anterior hippocampus. In the present study, we observed generalized responses to novel event content that were limited to the most anterior region of hippocampus. Unlike PRc and PHc, anterior hippocampal responses were observed for all forms of novel content, reflecting domain-general engagement of this region during the presentation of novel stimuli. Multivariate analyses further demonstrated that the representational code in anterior hippocampus does not differentiate between content classes. Specifically, distributed activation patterns in anterior hippocampus afforded reduced discrimination and demonstrated more confusability between content classes than exhibited by PHc and PRc. Together, the univariate and multivariate findings indicate that anterior hippocampus is engaged by many different forms of content and that the spatial patterns

of response evoked by different forms of content are not distinct, consistent with a domain-general representational code.

Domain-general coding, however, could take many forms. Some theories have proposed that hippocampal representations are abstract, reflecting arbitrary relationships between different sensory inputs, and do not contain sufficient information to discriminate between distinct forms of sensory content (Eichenbaum and Cohen 2001; Morris et al. 2003). An alternate possibility suggests that some hippocampal neurons would have direct visual object inputs, others direct visuospatial inputs, and yet others direct auditory inputs; by linking the activity of neurons that code related content (e.g., a person's face, voice, and written name), content-specific hippocampal neurons could demonstrate domain-general responses that code abstract concepts and be cued from multiple sensory modalities (Quian Quiroga et al. 2009).

One further possibility is that representational codes in anterior hippocampus convey important information about the salience or significance of specific stimuli (e.g., a stimulus is novel or associated with an extrinsic reward) that would be applicable to stimuli from many content classes. Notably, in the rodent brain, the density of dopaminergic, noradrenergic, and serotonergic inputs is greater in ventral (anterior in the human) hippocampus relative to the dorsal (posterior in the human) hippocampus (Gage and Thompson 1980; Verney et al. 1985). Based on the high-density of neuromodulatory inputs in anterior hippocampus, it is possible that this region is sensitive to motivational states (Moser M. B. and Moser 1998; Fanselow and Dong 2010) that might indicate the behavioral salience of incoming information to guide memory formation. Novelty in the

current study may serve as an important indicator of salience (Lisman and Grace 2005; Wittmann et al. 2007), and thus preferentially lead to domain-general maximal responding in anterior hippocampal regions sensitive to this motivational modulation. Such a generalized salience code would not necessarily be expected to further differentiate the content class of particular stimuli. While the current data cannot differentiate these alternate accounts of domain-general coding, our findings do indicate that anterior hippocampus maintains a less spatially organized coding of event content that is distinct from the content representations in both MTL cortex and posterior hippocampus.

Content representation in posterior hippocampus

Neuropsychological observations have led some to posit that hippocampus differentially mediates spatial memory (Bird and Burgess 2008). For example, some patients with selective hippocampal lesions demonstrate impaired recognition (Cipolotti et al. 2006; Bird et al. 2007; Bird et al. 2008) and visual discrimination of visuospatial information (Lee, Buckley et al. 2005; Lee, Bussey et al. 2005), with preserved performance for faces. Animal research suggests that such spatial memory impairments result primarily from damage to the dorsal (posterior in the human) hippocampus (Moser M. B. and Moser 1998). Lesions to the dorsal, but not ventral, hippocampus in the rodent severely impair memory formation in maze learning tasks, with the magnitude of the impairment being proportional to the size of the dorsal hippocampal lesion (Moser E. et al. 1993; Moser M. B. et al. 1995). Moreover, while place cells that demonstrate spatially

restricted firing patterns are present in both dorsal and ventral hippocampus, the proportion of such cells is lower in the ventral hippocampus, and place fields in ventral place cells are larger and less selective than dorsal hippocampal place fields (Jung et al. 1994).

Here, we demonstrate that in the human brain distinct representational coding of spatial information is primarily observed in the posterior hippocampus. While posterior hippocampus showed poor overall classification accuracy relative to PRc and PHc in our MVPA analysis, classification accuracy for scenes was significantly above chance. Further consideration of the classifier confusion matrices indicates that the classifier readily identified posterior hippocampal activation patterns for scenes in the presence of a high level of confusability between all other forms of content. Moreover, the difference between posterior hippocampal and MTL cortical classification accuracy was most apparent when scenes were removed from classifier training and testing (Supplementary Figure S2.2). When doing so, significant classification accuracy in posterior hippocampal regions was apparent in less than half of participants, whereas classification in PRc and PHc was significant in the majority of participants.

Perhaps most compellingly, our RSA findings provide a clear indication that the most posterior aspect of hippocampus maintains a coherent spatial code for scenes that is distinct from other forms of content. These findings revealed a high-degree of representational clustering of scene content in the posterior-most aspect of the hippocampus that was not observed in any portion of anterior hippocampus. These findings of a distinct representational code for scenes in posterior hippocampus are in

notable contrast to a previous report documenting poor content discrimination in hippocampus using MVPA (Diana et al. 2008). One primary difference between the present finding and this prior research is the consideration of anterior and posterior hippocampus as separate regions in the current study, which proved critical to our ability to resolve the distinctive representational codes maintained by these regions.

More generally, the multivariate techniques utilized in the present study were especially critical to our ability to determine the content-sensitivity of posterior hippocampus. To date, fMRI research in humans has made almost exclusive use of univariate statistical approaches to examine content coding in the hippocampus. Here, we did not observe significant peak amplitude responses in posterior hippocampus relative to baseline for any novel content class, including scenes, which would have limited our conclusions regarding content coding in posterior hippocampus. The differences between the univariate and multivariate findings in posterior hippocampus again highlight the power of combining different analysis approaches to understand the nature of representational coding in MTL subregions.

Collectively, our findings of a distinct representation of scene content in posterior hippocampus and domain-general responsiveness in anterior hippocampus suggest that the hippocampus consists of at least two functional modules whose functions may combine to support memory. This dissociation between the representational properties of anterior and posterior hippocampus may, to some degree, resolve conflicting findings from the literature that have shown both domain-general and scene-selective functional properties in hippocampus.

MTL representations of auditory content

An additional novel aspect of the current study is the inclusion of auditory information. Research on episodic memory has made predominate use of visual content, such as visual words, faces, objects, and scenes, and very little is known about the neurobiological substrates of memory for auditory events. Direct auditory inputs to PRc and PHc are meager relative to visual inputs, and it is possible that most auditory information reaches PRc and PHc through indirect connections with other structures (Munoz-Lopez et al. 2010). In the present study, we did not observe significant peak amplitude responding for either form of auditory content in PRc or PHc. Moreover, while PRc and PHc demonstrated the ability to classify some auditory content, classification performance for auditory stimuli was far below classification accuracies for scenes and faces. However, RSA revealed both overlapping representation of different forms of auditory content and discrimination of auditory and visual content in posterior PHc, suggesting a representation of auditory content in this region that is distinct from visual content. In the primate brain, PHc, but not PRc, receives limited input from unimodal auditory association cortex in the superior temporal gyrus (Suzuki and Amaral 1994), which may contribute to a more distinctive representation of auditory content in PHc than in PRc.

Alternatively, it is possible that another route for auditory information exists within the MTL that does not include connections to PRc and PHc. Entorhinal cortex receives direct auditory input from superior temporal gyrus (Amaral et al. 1983; Insausti

and Amaral 2008), through which auditory information could reach the hippocampus. Interestingly, univariate analyses revealed responses to novel auditory stimuli only in anterior hippocampus, raising the possibility that memories for auditory information are processed via different pathways than visual content within MTL. Human electrophysiological data provide additional evidence for this possibility, as neurons in hippocampus and entorhinal cortex, but not PHc, demonstrate responses to auditory stimuli (Quian Quiroga et al. 2009). Our findings emphasize the need for future research to consider potential differences in MTL pathways for visual and auditory memories.

Relationship between novelty responses and episodic encoding

While the present study cannot directly link content-based novelty responses to successful episodic encoding, a considerable body of research has demonstrated the relationship between novelty responses and successful memory formation (e.g., Kirchhoff et al. 2000; Ranganath and Rainer 2003; Fernandez and Tendolkar 2006; Dudukovic et al. 2010). Notably, in a previous study employing a similar incidental target detection task, we found that the magnitude of novelty responding in MTL cortex and hippocampus predicted subsequent memory outcome (Preston et al. 2010), providing some indication that novelty effects observed in the current study reflect episodic encoding. Importantly, the use of incidental novelty encoding paradigms in this and prior research suggests that MTL encoding occurs automatically, regardless of the particular goals of the task. Moreover, when task goals are held constant, as they are in the current study, we observe functional gradients in MTL cortex and hippocampus that differ based on the nature of

event content and can resolve specifics about event content from the distributed pattern of data. Recently, such distributed representations of face and scene content in prefrontal and temporal lobe structures during word-image encoding have been linked to successful memory formation (Kuhl et al. 2011). This finding suggests that the multivoxel representations of event content observed in the present study may play an important role in episodic encoding, and future hr-fMRI studies will help determine how distributed content codes impact memory performance.

Conclusions

While several leading theories focus on content as an important organizational principle for MTL function, the present data highlight the widely distributed and overlapping nature of content representation within the MTL. Moreover, the findings highlight the necessity of using multiple analysis approaches to characterize the representational capacity of MTL subregions. In particular, multivariate techniques may afford greater sensitivity to the nature of MTL subregional representation by taking into account the entire pattern of data within a region, not just those voxels that are maximally responsive to a predefined contrast.

SUPPLEMENTARY RESULTS

Distribution of content-sensitivity within anatomical MTL subregions

It is possible that anterior-posterior differences in content-sensitivity within MTL cortex observed in our univariate analysis of representational gradients were driven

primarily by greater overall activation for novel scenes in PHc compared to PRc, and a greater activation for novel faces in PRc compared to PHc, rather than a systematic change in content representation along the anterior-posterior extent of individual subregions. To investigate this possibility, we performed the same content × position ANOVAs within PHc and within PRc. For novel stimuli, a significant interaction between position and content was observed in PHc ($F(16,288) = 5.92$, $p < 0.001$) and there was a trend for an interaction in PRc ($F(20,340) = 1.82$, $p = 0.06$). Moreover, when we restricted our comparisons to novel faces and scenes, a significant content × position interaction was observed in both PHc ($F(4,72) = 10.98$, $p < 0.001$) and PRc ($F(5,85) = 4.98$, $p < 0.005$). In PHc, each anterior-posterior position demonstrated a significantly greater response to novel scenes relative to novel faces (all $t > 2.5$), and there was a decrease in activation to novel scenes from posterior to anterior positions as reflected by a significant linear trend ($F(1,18) = 8.82$, $p < 0.01$). In PRc, a significant linear trend was observed for novel scenes ($F(1,17) = 4.69$, $p = 0.05$), but not for novel faces ($F(1,17) = 1.00$, $p = 0.33$). Moreover, the content × position interaction observed in this region was driven by the fact that only the four anterior-most positions within PRc showed significant differences, or trends for differences, in activation for novel faces relative to novel scenes (all $t > 2$).

When considering the difference in activation for novel relative to repeated stimuli across the anterior-posterior axis of PHc and PRc individually, there was a significant content × position interaction in PHc ($F(16,288) = 2.44$, $p < 0.05$), reflecting a significant linear decrease in the novel – repeated difference for scenes from posterior to

anterior PHc. The content × position interaction was also reflected by the fact that the scene response differed from the face response in only the posterior-most positions in PHc (all $t > 2.6$, $p < 0.05$). A significant content × position interaction in PHc was also observed when the analysis was restricted to face and scene content ($F(4,72) = 3.09$, $p < 0.05$). When the same position × content ANOVAs were conducted within each hippocampal region, we observed neither a main effect of position nor an interaction between content and position in either anterior or posterior hippocampus (all $F < 2.37$).

To ensure that the observed gradients in novelty-based responding were not solely the result of the moderate spatial smoothing applied to the data, we repeated all of the univariate gradient analyses (both within and across MTL subregions) using non-smoothed data. Critically, the pattern of results when using non-smoothed data was identical to the pattern we observed for the smoothed data, suggesting that the observed gradients in novelty responding were not affected by the relatively small spatial smoothing kernel applied to the data.

Individual participant analysis of the distribution of content-sensitivity in MTL subregions

It is possible that the observed gradients in response across MTL cortex result from relatively small differences in the localization of the anterior-posterior boundary between PRc and PHc across individuals. In this case, when activation maps are normalized and averaged across the group, slight differences in localization of this boundary could give the impression of a linear gradient where none exists. To rule out

localization differences as a source of our gradient effects, we assessed content-sensitive novelty responses as a function of anterior-posterior position for each individual participant. A binomial test was used to determine whether the number of participants showing significant effects exceeded the number that would be expected purely by chance.

Importantly, 11 of 19 participants demonstrated a significant content × position interaction in MTL cortex when comparing responses to novel scenes and novel faces, with two additional participants showing a trend for an interaction. These results converge with the group-based statistics and provide evidence for a distribution of content-sensitivity novelty effects in MTL cortex at the level of individual participants. Results of the binomial test revealed that this number of participants exceeded the number that would be expected to show such an effect purely by chance ($p < 0.01$). Moreover, significant linear trends were also observed for novel scene and novel face stimuli at the level of individual participants. Twelve participants showed significant linear trends for scene stimuli that decreased from posterior MTL cortex to anterior regions (binomial $p < 0.01$); seven participants showed significant linear trends for face stimuli that decreased from anterior to posterior MTL cortex (binomial $p < 0.01$).

We performed a similar analysis assessing individual participant responses along the anterior-posterior axis of hippocampus. Within hippocampus, 17 participants demonstrated a significant main effect of anterior-posterior position (binomial $p < 0.001$), with only 2 participants demonstrating a significant content × position interaction (binomial $p = 0.25$). Binomial tests further revealed that the number of participants

showing significant linear decreases in novelty responding from anterior to posterior hippocampus exceeded the number expected by chance for all classes of novel content: 9 participants for auditory words, 10 participants for faces, 8 participants for scenes, 10 participants for sounds, and 11 participants for visual words (all $p < 0.01$). Collectively, the patterns of response at the individual participant level closely resemble the patterns observed from the group analyses, suggesting that the observed distribution of novelty-based responding at the group level were not due to individual differences in the localization of anterior-posterior boundaries within MTL cortex and hippocampus.

Multivariate pattern classification omitting preferred content

One possibility is that the classification accuracies derived from the model including all stimulus classes were driven solely by classification of a region's preferred class of content. However, an alternative possibility suggests that MTL subregional representations of non-preferred content may be sufficiently dissimilar from one another to allow for successful classification. Accordingly, we assessed the accuracy of the classifier when it was trained on novel stimuli excluding first faces (differentially preferred by PRc), then scenes (differentially preferred by PHc), and finally both faces and scenes.

Omission of faces

Importantly, when faces were omitted from the training and test data, overall classification accuracies remained significantly above chance (25%) in all MTL regions

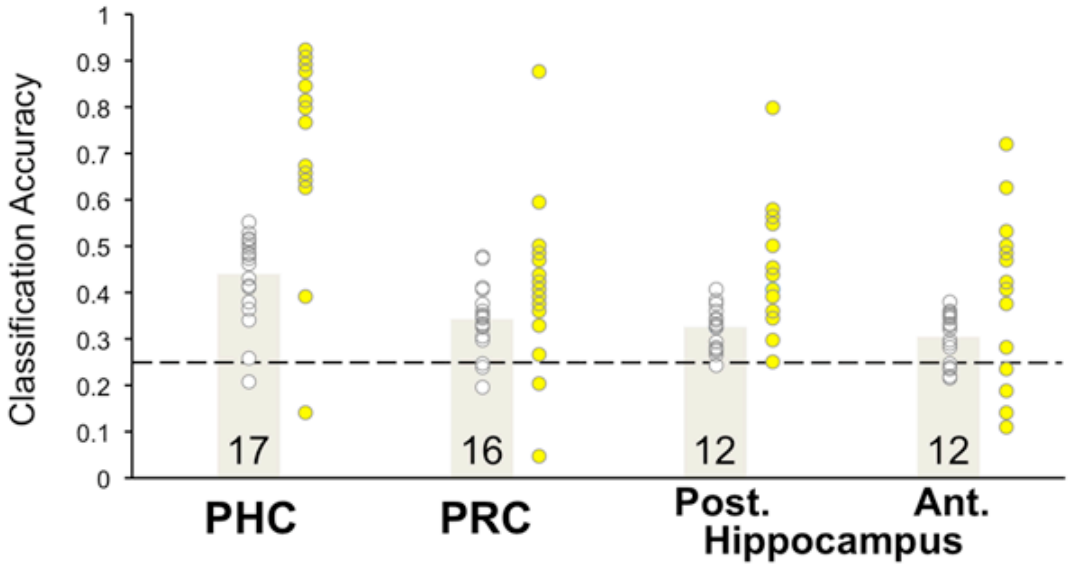
(all $t(18) > 4.52$, $p < 0.001$; Supplementary Figure S2.1 gray bars). Binomial tests revealed that the number of participants demonstrating above chance classification accuracy remained higher in PHc (n = 17) and PRc (n = 16) relative to anterior (n = 12) and posterior hippocampus (n = 12).

Moreover, repeated measures ANOVA showed a significant main effect of region on classifier accuracies when comparing anterior and posterior hippocampus to PHc (all $F > 28.30$, $p < 0.001$) and anterior hippocampus to PRc ($F(1,18) = 5.82$, $p < 0.05$). Specifically, PHc classification accuracy remained above chance for all remaining classes of content (all $t(18) > 2.70$, $p < 0.05$). A main effect of content on classification accuracy was also observed in PHc ($F(3,54) = 64.88$, $p < 0.001$), with novel scene classification being better than that of every other class of content (all $t(18) > 8.03$, $p < 0.001$) and visual word classification being better than spoken word classification ($t(18) = 2.60$, $p < 0.05$). Individual classification accuracies in PRc were similar to performance when all forms of content were included, with above chance classification accuracy observed for all remaining classes of content (all $t > 3.53$, $p < 0.01$) except spoken words ($t(18) = 1.15$, $p = 0.26$). Moreover, a significant main effect of content was observed in PRc ($F(3,54) = 4.71$, $p < 0.01$), with classification accuracy for scenes being better than that for spoken and visual words (all $t > 2.35$, $p < 0.05$). When considering individual classification accuracies excluding faces across MTL cortical regions, a significant region × content interaction was observed between PHc and PRc ($F(3,54) = 24.50$, $p < 0.001$).

Similarly, when faces were excluded from classification, significant main effects of content remained in anterior ($F(3,54) = 5.52$, $p < 0.01$) and posterior hippocampus

($F(3,54) = 10.58$, $p < 0.001$). In posterior hippocampus, classification accuracy was above chance only for scenes ($t(18) = 6.53$, $p < 0.001$), which differed from classification of all other classes of content (all $t > 3.83$, $p < 0.01$). In anterior hippocampus, classification of novel scenes and sounds exceeded chance levels (all $t > 2.58$, $p < 0.05$), and the classification accuracy for scenes was better than that of visual and spoken words (all $t > 2.39$, $p < 0.05$). The pattern of classification accuracies across hippocampal regions did not differ, as neither the main effect of region ($F(1,18) = 2.08$, $p = 0.17$) or the region × content interaction ($F(3,54) = 1.67$, $p = 0.19$) was significant.

Supplementary Figure S2.1



	PHC	PRC	Post. Hippocampus	Ant.
Faces	–	–	–	–
Scenes	***0.73***	***0.42***	***0.44***	***0.39***
Visual Words	***0.39***	***0.31***	0.28	0.24
Spoken Words	***0.32***	0.28	*0.29*	0.28
Sounds	***0.32***	***0.36***	0.29	***0.30***

Omission of scenes

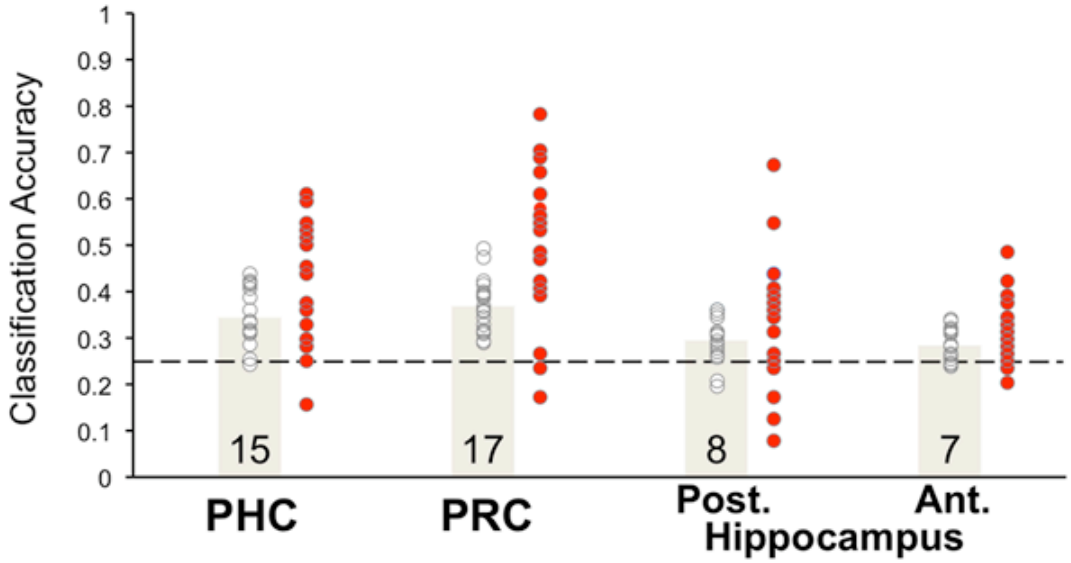
When novel scenes were omitted from classifier training and testing, overall classification accuracies remained above chance in all MTL regions (all $t(18) > 4.10$, $p < 0.001$; Supplementary Figure S2.2, gray bars). The number of participants demonstrating above chance classification accuracy was reduced in anterior (n = 7) and posterior (n = 8) hippocampus, while the number demonstrating significant classification in PHc (n = 15) and PRc (n = 17) was similar to classification when all forms of content were included as well as when faces were excluded from classifier training. The difference between classification accuracy in MTL cortex and hippocampus was further evidenced by a significant effect of region on classification accuracy for each of the comparisons between hippocampal and MTL cortical subregions (all $F(1,18) > 16.36$, $p < 0.001$).

Specifically, despite the omission of scenes from the model, above chance classification accuracy was observed for novel faces, visual words, and spoken words in PHc (all $t(18) > 2.60$, $p < 0.05$). A significant main effect of content ($F(3,54) = 4.99$, $p < 0.01$) was observed in PHc, with superior classification accuracy for novel faces relative to sounds ($t(18) = 4.24$, $p < 0.001$) and visual words ($t(18) = 2.91$, $p < 0.01$). Similarly, the omission of scenes from classifier training and testing did not impact the pattern of response in PRc, with above chance classification observed for all remaining classes of content (all $t(18) > 4.26$, $p < 0.001$) with the exception of spoken words ($t(18) = 1.36$, $p = 0.19$). Classification accuracies in PRc demonstrated a significant main effect of content ($F(3,54) = 9.22$, $p < 0.001$), with classification being better for faces relative to all other classes of content (all $t(18) > 2.89$, $p < 0.01$). The pattern of classification differed

between PHc and PRc, as evidenced by a significant region × content interaction ($F(3,54) = 3.16$, $p < 0.05$).

Similar patterns of classification accuracy were observed in anterior and posterior hippocampus when scenes were omitted from the MVPA analysis. In both anterior and posterior hippocampus, significant classification accuracy was observed only for faces (anterior: $t(18) = 4.69$, $p < 0.001$; posterior: $t(18) = 2.16$, $p < 0.05$). However, the effect of content on classification accuracies for individual stimulus classes was not significant in posterior hippocampus ($F(3,54) = 0.42$, $p = 0.74$), and exhibited only a trend for significance in anterior hippocampus ($F(3,54) = 2.35$, $p = 0.08$).

Supplementary Figure S2.2

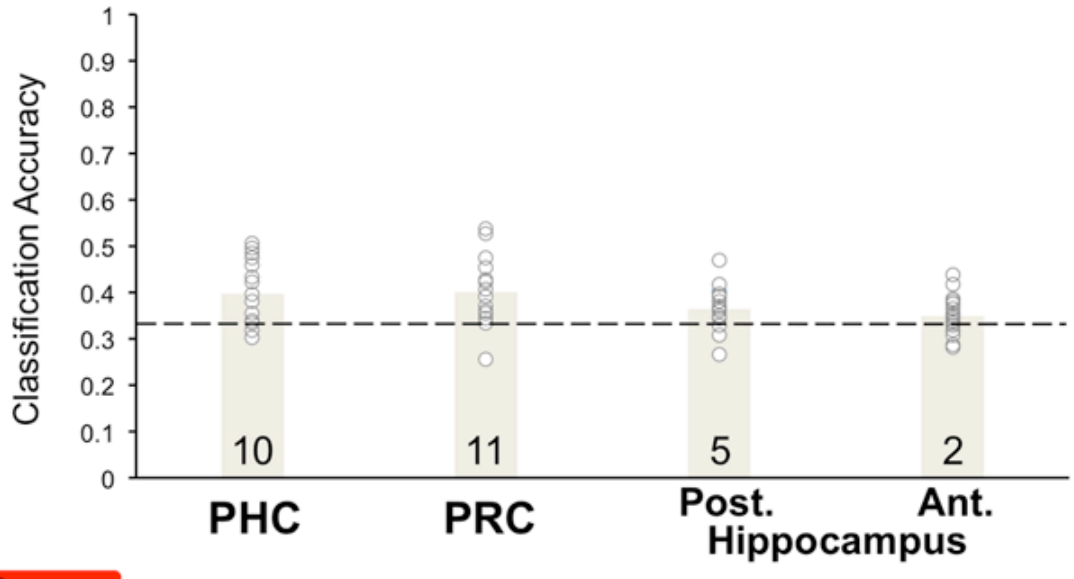


	PHC	PRC	Post. Hippocampus	Ant.
Faces	**0.43**	***0.49***	***0.39***	***0.33***
Scenes	–	–	–	–
Visual Words	***0.33***	***0.35***	0.28	0.25
Spoken Words	***0.33***	0.29	*0.29*	0.27
Sounds	0.29	***0.35***	0.29	0.29

Omission of faces and scenes

Finally, when both faces and scenes were omitted from classifier training and testing, above chance classification accuracies (greater than 33%) were observed in PRc (40%), PHc (40%), and posterior hippocampus (36%; all $t > 3.03$, $p < 0.01$), but not in anterior hippocampus (35%; $t(18) = 1.73$, $p = 0.10$; Supplementary Figure S2.3, gray bars). At the individual participant level, significant classification performance was observed in a greater number of participants in PRc (n = 11) and PHc (n = 10) relative to anterior (n = 2) and posterior hippocampus (n = 5). These results suggest that despite the omission of faces and scenes, MTL cortical subregions maintain their ability to classify non-preferred stimulus content. In PRc, classification accuracy for sounds ($t(18) = 2.91$, $p < 0.01$) and visual words ($t(18) = 4.88$, $p < 0.001$) remained significant. In PHc, classification of visual words was significantly above chance ($t(18) = 5.61$, $p < 0.001$), and a trend for above chance classification was observed for spoken words ($t(18) = 2.03$, $p = 0.06$). By comparison, within hippocampus classification performance was only significant for visual words ($t(18) = 2.26$, $p = 0.04$) in posterior hippocampus.

Supplementary Figure S2.3



	PHC	PRC	Post. Hippocampus	Ant.
Faces	–	–	–	–
Scenes	–	–	–	–
Visual Words	***0.45***	***0.44***	***0.38***	0.34
Spoken Words	*0.39*	0.36	0.36	0.33
Sounds	*0.35*	***0.40***	0.35	0.37

Chapter 3. Content Representation During Episodic Retrieval

REINSTATEMENT OF CONTENT-SPECIFIC EPISODIC DETAILS IN HUMAN MEDIAL TEMPORAL LOBE

The majority of text and figures in this section were prepared as a manuscript for submission to *Neuropsychologia*.

Co-author contributions: A. R. Preston is my P.I. and helped write the article.

Abstract

Remembrance of past experiences critically depends upon hippocampus and surrounding medial temporal lobe cortex (MTLC), with current theories proposing that perirhinal (PRc) and parahippocampal cortex (PHc) play content-specific roles in memory through respective processing of visual object and visuospatial event detail. Neuroimaging studies of content-specific MTLC function in humans have predominantly focused on episodic encoding, with conflicting findings. Standard univariate approaches generally show selective encoding responses in PRc and PHc for visual object and visuospatial content. In contrast, recent studies employing multivariate approaches indicate a distribution of content coding both within and across MTLC regions. However, distributed MTLC patterns have not been directly linked to mnemonic behavior. Here, we used event-related functional MRI and multivoxel pattern analysis (MVPA) to test whether distributed content representations in PRc and PHc formed during encoding are

reinstated during retrieval of specific event details. During a scanned episodic encoding task, participants visualized faces, objects, or scenes in response to presented adjectives. Participants then performed a surprise source recognition task during scanning, in which they indicated whether presented adjectives were studied in the face, object, or scene imagery tasks or were new adjectives. MVPA classifiers were trained to distinguish neural patterns evoked in PRC and PHC by the different imagery tasks during encoding. The trained classifier was then applied to data from the source retrieval task to index the strength of content-specific cortical reinstatement in each MTLC region and its relationship to performance.

Introduction

Episodic memory – memory for individual events – relies upon the engagement of a network of regions within the medial temporal lobe (MTL). The processing of specific forms of information to-be-remembered, however, is thought to be performed by specialized content-sensitive structures within this network. Recent neuroimaging studies have established the crucial involvement of MTL structures including perirhinal cortex (PRc), parahippocampal cortex (PHc), and hippocampus in representing different forms of visual object and visual scene information during episodic encoding (Staresina and Davachi, 2009; Litman et al., 2009; Diana et al., 2010; Liang et al., 2012). A fundamental question remains of the importance of these representations to mnemonic behavior, and in particular their relationship to the representations evoked within these MTL structures during episodic retrieval.

Predominant anatomy-based theories have described a content-based functional specialization along the anterior-posterior axis of MTL cortex, where PRc is critically involved in processing visual objects and faces, while PHc is involved in processing visual scene information. Early neuropsychological and neuroimaging studies have found support for this view with regard to episodic encoding, with more recent papers finding evidence for graded content representations, where the most selective responses for visual object information were found in anterior MTL cortex versus scenes in posterior MTL cortex, with a transitional zone in the middle that was involved in processing both objects and scenes. New studies employing multivariate pattern analyses have found that distributed activity patterns in PRc and PHc observed during episodic encoding could in fact distinguish between multiple forms of content including faces, objects, and scenes (Diana et al., 2010; Liang et al., 2013). While these recent findings argue against the traditional view that representation in MTL cortex was strongly selective, they nonetheless concur that distributed content representations differ substantially along the anterior-posterior axis of MTL cortex, with the most distinctive representations for face and scene information to be found at the anterior and posterior extremes, respectively (Liang et al., 2013).

There has also been a long-standing view that the anterior and posterior subregions of hippocampus play different roles in episodic memory, with rodent models proposing that the former supports memory for general contextual information while the latter supports detailed space representations (e.g., Fanselow and Dong, 2010). In humans, there has been a mix of evidence, with some studies implicating posterior

hippocampus in spatial navigation and content-specific spatial memory (e.g., Suthana et al., 2009, 2011; Bird and Burgess, 2008), while others found hippocampal responses that did not differ based on content (e.g., Awipi and Davachi, 2008; Prince et al., 2005; Preston et al., 2010). Until recently, however, the majority of papers have not considered looking at anterior and posterior hippocampus separately. Now there is evidence for content-general responses to novel stimuli in anterior hippocampus using univariate approaches, while a multivariate approach identified a distributed code for scene information in posterior hippocampus (Liang et al., 2013). However, the distributed content representations in posterior hippocampus or MTL cortex have yet to be linked to behavior, and the predominant focus on episodic encoding does not shed light on what relationship they have to the representations that are engaged during episodic retrieval.

One mechanism by which distributed content representations might contribute to successful memory is episodic reinstatement. According to this view, the network of neurons that fire preferentially during initial learning experience a strengthening of internal connections. The success of later retrieval, therefore, would be reflected by the regeneration of that original network. In the MTL, the concept of reinstatement may further distinguish complementary roles for hippocampus and MTL cortex. Whereas MTL cortical structures would represent different forms of information encountered during episodic encoding, the hippocampus may play a critical role in indexing the specific representations to be reinstated, or utilizing the event information conveyed by MTL cortex to guide behavior.

Several recent papers have focused on examining reinstatement in the human brain. One study by Johnson et al. (2009) provided evidence that a global brain pattern can reflect reactivation of task-related information present during the initial experience. However, the whole brain approach does not allow one to make conclusions about the specific roles of MTL subregions in reinstating different forms of event content. Additional work by Staresina et al. (2012) took a similar strategy and found that reinstatement of scene-encoding patterns in PHc was higher during recollection of specific scenes. While these and other findings (e.g. Ritchey et al., 2012) implicated PHc in a general form of reinstatement, the focus on PHc and scene stimuli does not allow testing of prominent theories of the roles of MTL subregions in representing unique forms of visual content.

The present study leverages powerful pattern classification techniques on a complete encoding and retrieval dataset to examine the content-sensitive contributions different MTL cortical and hippocampal subregions to episodic reinstatement, and reveals novel links between cortical reinstatement and mnemonic behavior.

Materials and Methods

Participants

Twenty-one healthy, right-handed volunteers participated in the experiment after giving informed consent in accordance with a protocol approved by the University of Texas at Austin Institutional Review Board. Participants received $25/hr for their involvement. Data from 15 participants were included in the analyses (age 19-33 yrs,

mean = 23.9 ± 3.9 yrs; 12 females), with data from six total participants being excluded due to technical problems with the scanner (three participants), early termination of the experiment because of discomfort (one participant), and failure to perform the behavioral task (two participants).

Behavioral procedures

During functional scanning, participants performed a mental imagery task in response to visual word cues (black text on white background; Arial 36 point). In separate retrieval scans, participants performed a source recognition task in response to the same visual word cues. Stimuli were generated in MATLAB (The MathWorks, Inc., Natick, MA), using the Psychophysics Toolbox extension (Brainard, 1997), on an Apple MacBook laptop computer and back-projected via a magnet-compatible projector onto a screen that could be viewed through a mirror mounted above the participant's head. Participants responded with an optical button pad held in their right hand.

Incidental encoding. During five slow event-related encoding scans, participants were presented with written cues that designated whether they were to internally visualize a face, object, or scene in response to a simultaneously presented adjective (Figure 3.1 top). At the start of each imagery trial, the written cue and target adjective were presented for 3.6 s. Following presentation of the imagery cue and adjective, a written prompt was presented for 1.4 s during which time participants indicated with a key press whether their mental image was “vivid with strong details”, “vivid but lacking in detail”, “vague/unclear”, or “could not be visualized”.

For face trials, participants were instructed to visualize the face of someone famous or familiar to them who could be described by the adjective. For object trials, participants were instructed to visualize a single non-living object in isolation that could be described by the adjective. For scene trials, participants were instructed to visualize a spatial environment that could be described by the adjective and did not contain single prominent foreground objects. During each encoding run, participants performed ten trials under each of the three imagery conditions (faces, objects, scenes). At no point were participants informed that their memory for the adjectives would later be tested. Thus, any encoding done by the participants was purely incidental.

In order to isolate the neural responses evoked during incidental encoding, each imagery trial was separated from one another by a 6 s baseline task (Figure 3.1 top). During this time participants were presented with three arrows for 2 s and indicated with a key press whether it was pointing to the left or right. A 1 s fixation crosshair indicated the start of the next imagery trial. The presentation order of face, object, and scene imagery trials was generated using a sequencing algorithm to optimize efficiency in event-related designs and to ensure that within any run, trials of each imagery task were equally likely to be followed by a face, object, or scene trial (Dale, 1999). Five presentation orders were generated to be used as counterbalancing groups.

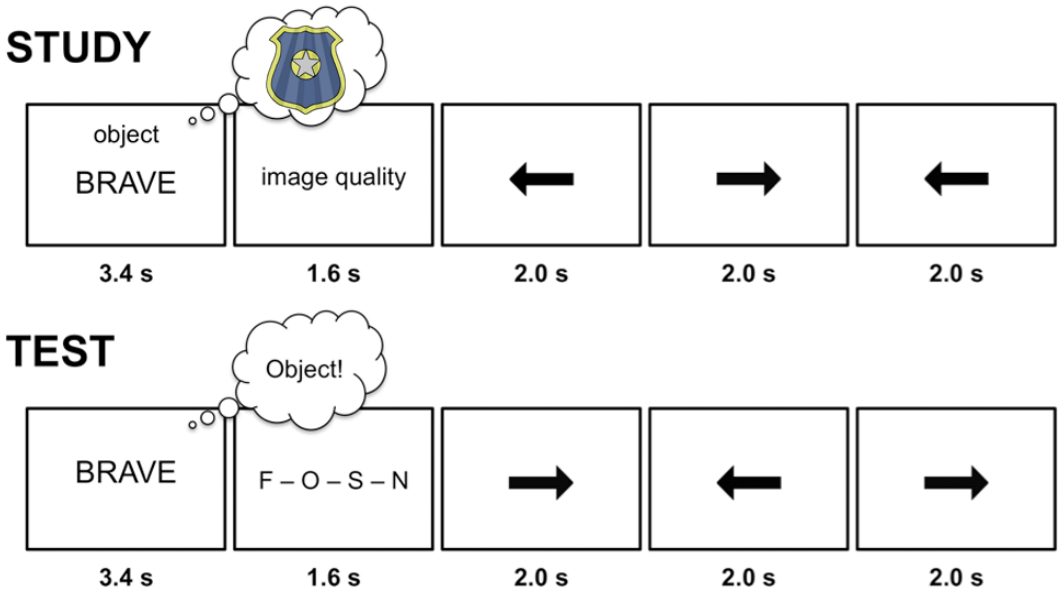
In total, all participants viewed the same set of 200 adjectives across the experiment, 150 of which were randomly assigned to the three imagery conditions. The remaining 50 adjectives were assigned to the novel lure condition, and were later presented to the participants during source retrieval. The adjectives used in this

experiment were selected from an online linguistic database (http://websites.psychology.uwa.edu.au/MRCDatabase/uwa_mrc.htm), based on independent experimenter ratings to find adjectives that were equally and highly imageable across the imagery conditions.

Source retrieval. In five separate retrieval scans, participants were given a surprise cued recall test on the adjectives studied in the encoding runs. Fifty previously unseen adjectives were also presented as novel lures. At the start of each retrieval trial, one of the studied adjectives or lures appeared in the center of the screen for 3.6 s (Figure 3.1 bottom), and participants were instructed to silently recall the mental image that they had created for that adjective during encoding. Following presentation of the adjective, a written prompt was presented (1.4 s), during which time participants indicated with a key press whether the retrieved image associated with the adjective corresponded to a face, object, or scene, or whether the adjective was entirely novel.

During each retrieval scan, participants viewed 30 studied adjectives (10 from each of the imagery conditions) and 10 lure adjectives. The presentation order of face-, object-, and scene-related adjectives and novel lures was generated using the same sequencing algorithm as that used for the encoding runs (Dale, 1999). As with encoding, five distinct trial orders were created using this algorithm that allowed for counterbalanced stimulus presentation across participants. In order to isolate the neural responses evoked during cued recall, each retrieval trial was separated from one another by a 6 s baseline task (Figure 3.1 bottom). A 1 s fixation crosshair indicated the start of the next retrieval trial.

Figure 3.1: Experimental design.



fMRI acquisition

Imaging data were acquired on a 3.0 T Signa whole-body MRI system (GE Medical Systems, Milwaukee, WI, USA) with a single-channel, custom-made transmit/receive head coil. Head movement was minimized using additional foam padding. For each participant, a high-resolution T1-weighted SPGR image (sagittal plane, 1 mm^3 resolution) was acquired. This image was used as the whole brain structural image to which all other participant-specific images were initially coregistered and was also used to calculate non-linear transformation parameters when normalizing each participant's structural and functional images to the MNI template. Prior to the incidental encoding task, a T2-weighted, flow-compensated spin-echo structural image (TR = 3000 ms; TE = 68 ms; 0.47 × 0.47 mm in-plane resolution) was then acquired, with 33 3-mm thick oblique axial slices (0.6 mm gap) oriented parallel to the main axis of the hippocampus and adjusted to maximize coverage of the whole brain; functional volumes from the incidental encoding task used the same slice locations as this T2 image. Prior to the retrieval scans, another T2-weighted structural volume was acquired with identical parameters to the first except for lower in-plane resolution (0.94 x 0.94 mm); functional volumes from the cued recall task used the same slice locations as this T2 image. This procedure enabled accurate and independent spatial coregistration of the encoding and retrieval scans to the high-resolution T1 image collected at the beginning of scanning.

Functional images collected during the encoding and retrieval scans were acquired using a T2*-sensitive gradient echo EPI sequence (TR = 2000 ms; TE = 30 ms; flip angle = 73°; FOV = 24 cm; 3.75 × 3.75 × 3.6 mm resolution, interleaved slice

acquisition). The slice locations of the encoding scans were prescribed to match those of the T2 structural image that preceded them; likewise, the retrieval scans were prescribed to match their respective T2 structural image. Immediately prior to acquisition of the incidental encoding data, a high-order shimming procedure was utilized to reduce B0 heterogeneity. This high-order shimming procedure was also used prior to the retrieval scans. For each participant, a total of 920 functional volumes were acquired over five encoding scans and 1220 volumes were acquired over five retrieval scans. In each functional run, four EPI volumes (a total of 8 s) were collected prior to beginning the first trial of the experiment to allow for T1 stabilization. These initial four volumes were discarded prior to fMRI data analysis.

Preprocessing of fMRI data

Data were preprocessed using SPM5 (Wellcome Department of Imaging Neuroscience, London, UK) and custom Matlab routines. For each participant, the functional volumes from the encoding scans were realigned to the first volume in the timeseries to correct for motion. The first volume of the encoding timeseries was then coregistered to the T2-weighted structural image acquired prior to the encoding scans. The resulting coregistration parameters were then applied to the entire encoding timeseries. These steps were then separately performed for the retrieval scans, so that the entire retrieval timeseries was coregistered with the T2-weighted structural image taken immediately prior to retrieval. Both T2-weighted structural images were then

coregistered with the high-resolution T1-weighted SPGR, and the resulting coregistration parameters were applied to the respective functional timeseries.

To enable group-level analyses, we used the Advanced Normalization Tools (ANTS) to normalize individual participant's brains to the MNI template. Specifically, each participant's high-resolution T1-weighted anatomical volume was normalized to the MNI template based on the shape of the gray- and white-matter boundaries, using non-linear diffeomorphic transformations. The transformation parameters were then applied to all of the participant's structural and functional volumes. Functional volumes were spatially smoothed (5mm FWHM), and then high-pass filtered to remove low frequency drift (longer than 128 s). The resulting functional timeseries volumes were converted to percent signal change in preparation for univariate statistical analyses, or z-scored in preparation for multivoxel pattern analysis.

Identification of MTL Regions-of-Interest

Anatomically defined regions-of-interest (ROIs) were demarcated on the T1-weighted, high-resolution (1 mm^3) standard MNI template, through automatic volumetric segmentation via the FreeSurfer image analysis suite (http://surfer.nmr.mgh.harvard.edu). The resulting ROI masks were then manually edited to ensure correspondence to current guidelines for demarcating the border of human perirhinal cortex (Ding and van Hoesen, 2010). Given previous studies that have raised the possibility that content representation may differ significantly along the anterior-posterior axis of the MTL (e.g. Staresina et al., 2011; Litman et al., 2009; Diana et al., 2010; Liang et al., 2013), we further divided the

hippocampal and MTL cortical masks into three ROIs of equal thickness (anterior, middle, and posterior hippocampus; anterior, middle, and posterior MTL cortex). Anterior MTL cortex most closely corresponds to the perirhinal and entorhinal cortex, while posterior MTL cortex most closely corresponds to parahippocampal cortex with the middle portion of MTL cortex corresponding to a transitional zone that has previously demonstrated encoding activation for both visual object and scene stimuli (Staresina et al., 2011).

Multivariate pattern analysis of fMRI data

The goal of the current study is to assess whether different MTL subregions make unique content-based contributions to episodic reinstatement. To test this we used a multivoxel pattern classifier to measure reinstatement of encoding patterns during cued source retrieval of different forms of imagined content. We explored whether reinstatement was related to memory performance by asking if it 1) tracked participants' correct responses and 2) predicted participants' source errors.

MVPA classification of imagery-based encoding. As an initial step, we assessed whether patterns of activity evoked during encoding could discriminate between the three imagery conditions (i.e., imagined faces, imagined objects, and imagined scenes). First, we calculated a activation pattern for each encoding trial by averaging the three functional volumes corresponding to the peak of the hemodynamic response. We then created a regressor matrix to label each trial's activation pattern according to the imagery condition to which it belonged (i.e., face imagery, object imagery, scene imagery). Every

encoding trial was included regardless of the vividness rating at encoding, so that there were an equal number of timepoints included for each condition (50 per condition).

Pattern classification analyses were implemented using the Princeton MVPA toolbox and custom code for Matlab. Within each anatomical ROI, classifier performance during imagery-based encoding for each participant was calculated using a 5-fold cross-validation procedure that implemented a regularized logistic regression algorithm to train the classifier (Bishop, 2006; Rissman et al., 2010). Data from four encoding runs were used for classifier training, and the remaining run was used as test data to assess the generalization performance of the trained classifier. This process was iteratively repeated five times, one for each of the possible configurations of encoding runs. The classification performances from each fold of the cross-validation procedure were averaged to obtain the final pattern classification performances for every participant in each ROI. Two-tailed Student's t-tests were conducted in each MTL ROI to assess whether classification accuracy across participants were significantly greater than the theoretical chance level of 33%.

Classification of retrieval patterns during source retrieval. Next, we applied the classifier trained on all encoding patterns to data from the cued recall task to determine if content-specific activation patterns were reinstated during successful recall of source details. As was previously done with the encoding dataset, we first calculated an activation pattern for each of the cued recall trials by averaging the three functional volumes corresponding to the peak of its hemodynamic response. We sorted these retrieval activation patterns according to the imagery task they were associated with, and

whether or not the participant successfully recalled the imagery task. Novel trials were not considered, resulting in six conditions: correct source identification of face detail, object detail, and scene detail, and incorrect source identification of face detail, object detail, and scene detail.

For each MTL ROI, a classifier was trained on all encoding activation patterns associated with the three imagery tasks. We then used the retrieval activation patterns as test data to assess the generalization performance of the classifier. Reinstatement for each of the six conditions was quantified by the mean classification accuracy, which was calculated as the proportion of trials whose associated imagery task was correctly predicted by the classifier. To assess whether reinstatement was 1) mediated by participants source memory performance and 2) specific to different forms of imagined content across MTL subregions, we subjected the mean classifier accuracies to repeated-measures ANOVA. Specifically, we conducted 3-way ANOVAs using ROI as a factor with three levels (e.g., anterior, middle, and posterior), imagery task as a factor with three levels (e.g., face, object, and scene), and source memory performance as a factor with two levels (e.g., source correct and source incorrect). This procedure was performed separately across anterior-posterior ROIs in hippocampus and MTL cortex.

To identify the unique content-based differences in reinstatement within MTL ROIs, we conducted two planned comparisons. First, we performed two-tailed Student's t-tests across participants to assess whether reinstatement in each of the six retrieval conditions was significantly different from the theoretical chance level of 33%. Second, we performed one-tailed Student's t-tests across participants to assess whether mean

classification accuracy for source correct trials was significantly greater than source incorrect trials for each form of content.

Classification of participants' source memory errors. In addition to assessing whether reinstatement differed between correct versus incorrect source memory judgments, we asked whether the classifier could predict the participants' actual choices during errors in source judgments. To test the relationship between classifier performance and source memory errors, we analyzed trials in which a participant identified the wrong imagery task (source error) and trials in which they identified a novel lure as being studied in one of the three imagery conditions (false alarms). For these trials we labeled the classifier output for the content chosen by the participant as "selected content" and the combined (averaged) output for the content classes that were not chosen as "alternate options". In each MTL ROI, we calculated the mean classifier output for the selected content and alternate options across all source error and false alarm trials. We performed a two-tailed paired sample t-test across participants to assess whether mean classifier output significantly differed between the selected content and alternate options.

Results

Behavioral performance

Vividness ratings during imagery-based encoding. To ensure that subjects were able to visualize different forms of content during encoding, we first binned trials into vivid imagery trials (e.g., where the participant indicated that imagery was "vivid with strong details" or "vivid but lacking detail") and poor imagery trials (e.g., where the

participant indicated that imagery was “vague/unclear” or “could not be visualized”). Across participants, the proportion of vivid imagery trials was high for all content classes (face mean: 81% SE: 2.4%, object mean: 77% SE: 2.4%, scene mean: 82% SE: 2.9%). Importantly, a repeated measures ANOVA revealed that the vividness rating did not significantly differ by imagery condition ($F(2,28) = 2.656$, $p = 0.09$).

Source recall performance. At retrieval, participants were able to successfully identify the information content that they visualized for studied adjectives (face mean: 71% SD: 3.0%, object mean: 63% SD: 3.7%, scene mean: 64% SD: 3.2%, all $t > 10.45$, all $p < 0.001$). A repeated-measures ANOVA of correct source trials revealed a significant main effect of content ($F(2,28) = 3.394$, $p = 0.048$), that reflected better source identification of faces relative to objects ($t(14) = 2.47$, $p = 0.026$). Correct source identification of faces did not differ from that of scenes, nor did correct source identification of objects differ from that of scenes (all $t < 1.90$, all $p > 0.05$).

Source errors, in which participants indicated the wrong imagery task for studied adjectives, did not significantly differ by content (face mean: 14% SD: 2.1%, object mean: 19% SD: 2.9%, scene mean: 19% SD: 1.9%, $F(2,28) = 2.865$, $p = 0.074$). Source misses, in which participants indicated a novel response for studied adjectives, did not significantly differ by content (face mean: 10% SD: 1.1%, object mean: 12% SD: 1.1%, scene mean: 10% SD: 1.6%, $F(2,28) = 0.530$, $p = 0.594$).

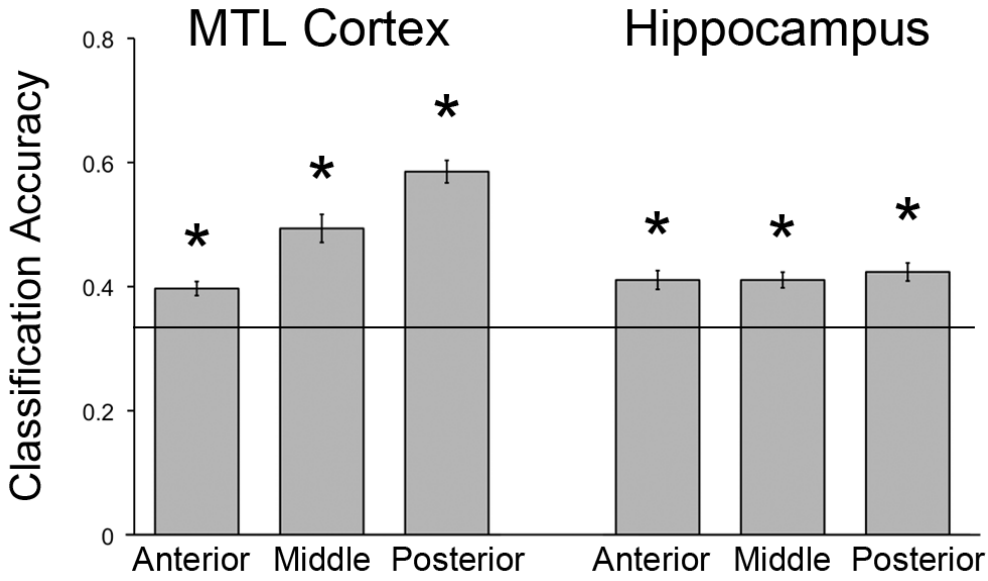
Participants were also able to correctly identify novel words (correct rejection mean: 72% SE: 2.9%, $t = 16.13$, $p < 0.001$). False alarms rates for novel words did not

significantly differ by content (face mean: 8.4% SD: 1.6%, object mean: 6.1% SD: 1.1%, scene mean: 7.3% SD: 1.0%, $F(2,28) = 1.17$, $p = 0.324$).

Visualization of content-specific detail evokes distinct representations in MTL

First, it was critical to confirm that internal visualization of faces, objects, or scenes during encoding evoked distinguishable patterns of activity in MTL subregions before examining content-specific reinstatement during source retrieval. All anatomical MTL ROIs demonstrated classification of the three imagery conditions above the theoretical chance level of 33% (Figure 3.2, all $t > 5.332$, all $p < 0.001$). This indicated that internally generated images of different content types evoked in MTL during the encoding task were distinct from one another.

Figure 3.2: Accuracy of face, object, and scene decoding. The height of the bar represents the mean classification accuracy across participants for each MTL region-of-interest. The error bars represent the standard error. The horizontal line represents chance level accuracy (33%). Asterisks indicate significantly greater than chance level accuracy ($p < 0.05$).



Reinstatement in MTL is related to successful source memory

Successful performance in the cued recall task critically depends on retrieving source details specific to the original encoding event. This is thought to be reflected by reinstatement of brain patterns active during learning. Here, we measured reinstatement by training an MVPA classifier on encoding data in anterior, middle, and posterior hippocampus and MTL cortex and testing it on activation patterns during source retrieval.

Content-sensitive reinstatement in MTL cortex. We first examined whether reinstatement of encoding patterns in anterior, middle, and posterior MTL cortex differed based on imagined content and source memory performance. A repeated-measures ANOVA revealed a significant 3-way interaction between ROI, content, and source memory (Figure 3.3a, $F(4,56) = 3.357$, $p = 0.016$). Follow-up planned comparisons within each individual MTL cortical ROI were conducted to identify whether it contributed to reinstatement during source retrieval in a content-specific manner.

In anterior MTL cortex, we first conducted one-tailed t-tests comparing reinstatement associated with correct source identification versus incorrect source identification. This analysis revealed greater reinstatement during correct versus incorrect source identification of faces ($t(14) = 2.377$, $p = 0.016$), but not objects or scenes (all $t < 1.135$, all $p > 0.05$). Furthermore, two-tailed t-tests showed that the classifier predicted the imagery task at above chance level during correct source identification of faces ($t(14) = 2.496$, $p = 0.028$) but not during incorrect source identification of faces ($t(14) = -1.894$, $p = 0.08$).

In the middle segment of MTL cortex, one-tailed t-tests comparing reinstatement associated with correct source identification versus incorrect source identification revealed greater reinstatement during correct versus incorrect source identification of objects ($t(14) = 2.091$, $p = 0.027$) and scenes ($t(14) = 2.839$, $p = 0.006$), but not faces ($t(14) = -0.156$, $p = 0.561$). Furthermore, two-tailed t-tests showed that the classifier predicted the imagery task at above chance level during correct source identification of objects and scenes (all $t > 1.904$, $p < 0.05$) but not during incorrect source identification of objects and scenes (all $t < -0.990$, all $p > 0.05$).

In posterior MTL cortex, one-tailed t-tests comparing reinstatement associated with correct source identification versus incorrect source identification revealed greater reinstatement during correct versus incorrect source identification of objects ($t(14) = 1.798$, $p = 0.047$) and scenes ($t(14) = 4.70$, $p < 0.001$), but only a trend for faces ($t(14) = 1.573$, $p = 0.069$). Subsequent two-tailed t-tests revealed that the classifier predicted the imagery task at above chance level during correct source identification of scenes ($t(14) = 9.078$, $p < 0.001$) but not objects ($t(14) = 0.063$, $p = 0.95$). By comparison, the classifier did not perform above chance level for incorrect source retrieval of scenes or objects (all $t < 1.724$, all $p > 0.05$).

Content-sensitive reinstatement in hippocampus. We next examined whether reinstatement of encoding patterns in anterior, middle, and posterior hippocampus differed based on imagined content and source memory performance. A repeated-measures ANOVA revealed a significant 3-way interaction between ROI, content, and source memory (Figure 3.3b, $F(4,56) = 2.807$, $p = 0.034$). Follow-up planned

comparisons within each individual hippocampal ROI were conducted to identify whether it contributed to reinstatement during source retrieval in a content-specific manner.

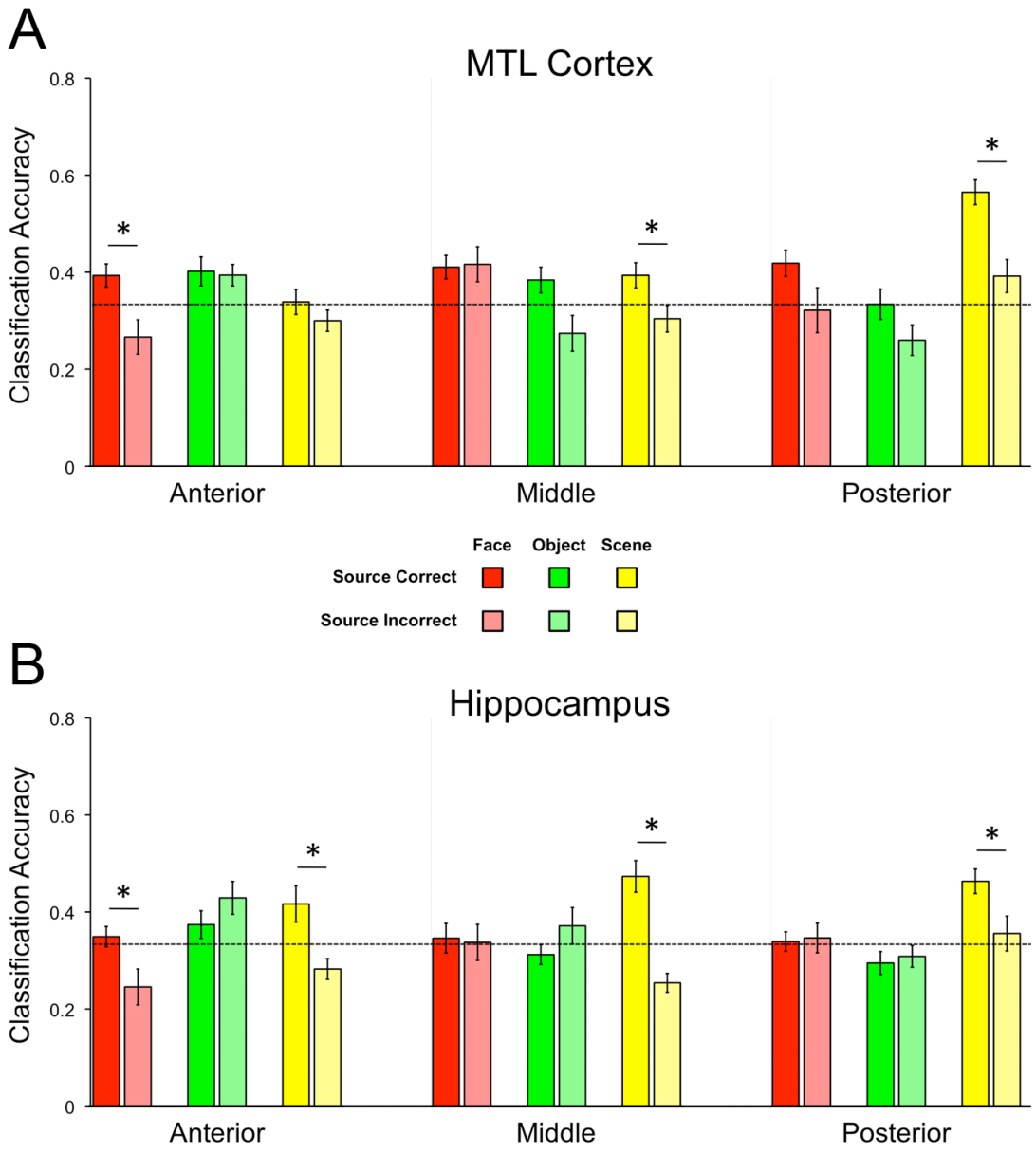
In anterior hippocampus, one-tailed t-tests comparing reinstatement associated with correct source identification versus incorrect source identification revealed greater reinstatement during correct versus incorrect source identification of faces ($t(14) = 3.024$, $p = 0.004$) and scenes ($t(14) = 2.725$, $p = 0.008$), but not objects ($t(14) = -0.125$, $p = 0.903$). Subsequent two-tailed t-tests revealed that the classifier predicted the imagery task at above chance level during correct source identification of scenes ($t(14) = 2.249$, $p = 0.041$) but not faces ($t(14) = 0.730$, $p = 0.477$). By comparison, the classifier did not perform above chance level for incorrect source retrieval of scenes or faces (all $t < -2.315$, all $p > 0.05$).

In the middle segment of hippocampus, one-tailed t-tests revealed greater reinstatement during correct versus incorrect source identification of scenes ($t(14) = 5.221$, $p < 0.001$), but not faces or objects (all $t < 0.154$, all $p > 0.05$). Subsequent two-tailed t-tests revealed that the classifier predicted the imagery task at above chance level during correct source identification of scenes ($t(14) = 4.196$, $p < 0.001$) but not faces or objects (all $t < 0.397$, all $p > 0.05$). The classifier did not perform above chance level for incorrect source retrieval of any of the imagery conditions (all $t < 1.015$, all $p > 0.05$).

One-tailed comparisons in posterior hippocampus revealed greater reinstatement during correct source identification of scenes ($t(14) = 2.675$, $p = 0.009$) but not faces or objects (all $t < -0.228$, all $p > 0.05$). Two-tailed t-tests again showed that the classifier

performed above chance level during correct source identification of scenes ($t(14) = 5.091$, $p < 0.001$), but not faces or objects (all $t < 0.265$, all $p > 0.05$). The classifier did not perform above chance level for incorrect source retrieval of any of the imagery conditions (all $t < 0.635$, all $p > 0.05$).

Figure 3.3: Decoding of retrieval-related activity as a function of content class and source performance in MTL cortex (A) and hippocampus (B). Each bar represents the mean classification accuracy across participants. The error bars represent standard error. The dashed line represents chance level accuracy (33%). Asterisks indicate a significant pairwise difference between the classification accuracy for source correct and source incorrect trials for a given content class.



Distributed patterns in MTL cortex predict participants' source errors

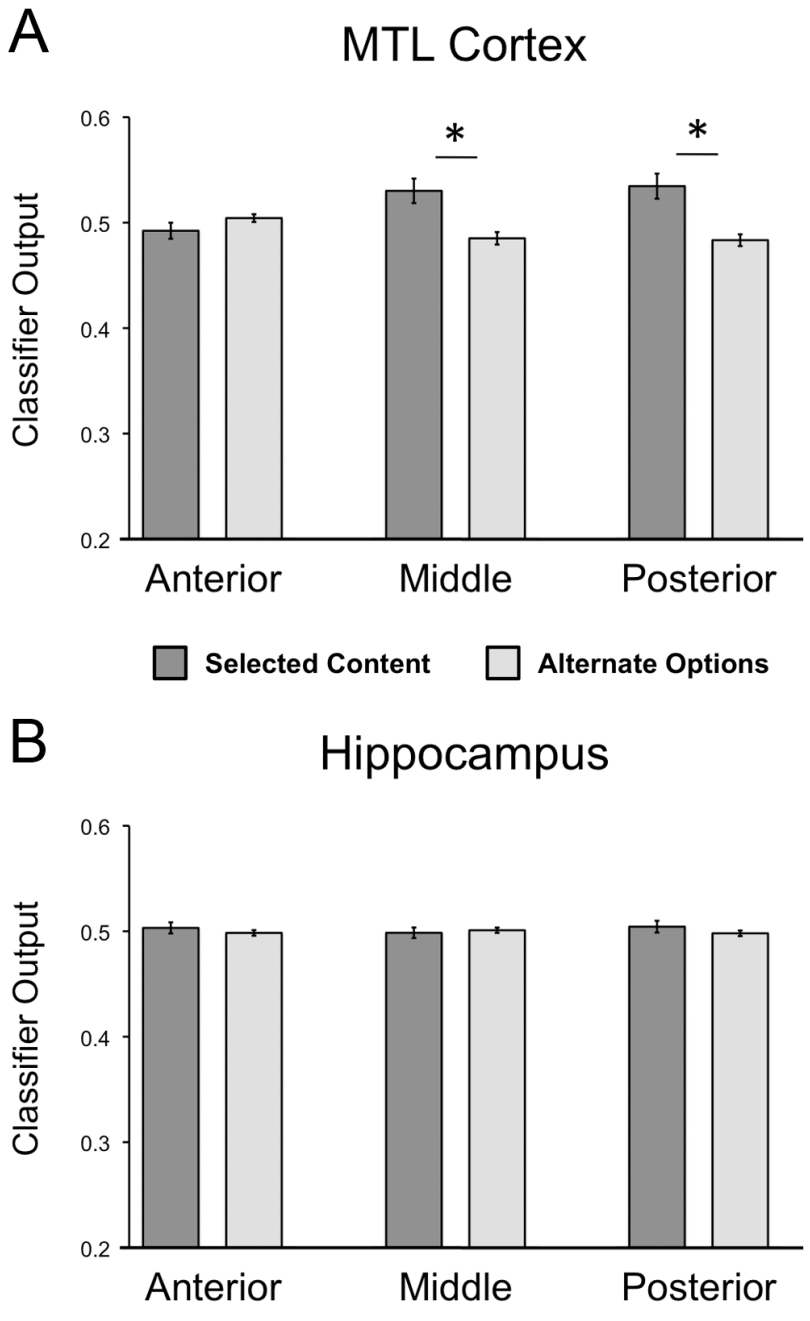
If reinstated patterns are used by hippocampus to guide source judgments, they should be expected to predict participants' actual choices when they make errors. To test this hypothesis, we collapsed all trials where participants made source errors (when they indicated the wrong content class for studied adjectives) or false alarms (when they incorrectly indicated a novel adjective as being studied in one of the imagery tasks). We then compared the mean classifier output for the content selected by the participant to the mean classifier output for the alternative options for each trial.

In the middle and posterior segments of MTL cortex, classifier output for the selected content was significantly greater than the mean output for the alternative options (Figure 3.4a, all $t > 2.580$, all $p < 0.05$). This suggested that distributed activity in these MTL cortical subregions was predictive of participants' choices during errors in source judgments. By contrast, anterior MTL cortex did not demonstrate a significant difference between classifier output for the selected content and the alternative options ($t(14) = -1.062$, $p = 0.847$). Likewise, no hippocampal ROIs demonstrated significant pairwise differences between classifier output for the selected content and the alternative options (Figure 3.4b, all $t < 0.762$, all $p > 0.05$).

While we instructed participants to pay attention to the imagery cues in every encoding trial, it is possible that some adjectives evoked unintentional imagery that diverged from the actual task they were given. In such cases, reinstating this subjective experience would lead the participant to make errors in source judgment. However, we also assessed whether distributed patterns was predictive of participants' choices even for

adjectives for which the participants performed no imagery task. To test this possibility we repeated the analysis but limited it to false alarm (lure) trials. This restricted analysis revealed that, in posterior MTL cortex, classifier output for the selected content was still greater than the output for the alternative options ($t(14) = 2.382$, $p = 0.03$). No other MTL cortical nor hippocampal ROI demonstrated a significant difference in outputs (all $t < 1.750$, all $p > 0.05$).

Figure 3.4: Decoding of participants' decisions during errors in source judgments. The dark bars represent the classification output predicting the participant's choice on each source incorrect and false alarm trial. The light bars represent the classification output for the content classes not chosen by the participant. Asterisks indicate a significant pairwise difference between the outputs for the selected and alternate content classes ($p < 0.05$).



Discussion

The unique functions of medial temporal lobe subregions have been described via many theoretical frameworks, many of which are still under debate. One of the most compelling of these organizing principles is that the different subregions of the MTL are sensitive to certain forms of event content. While recent work using multivariate pattern classification have revealed that MTL cortex and hippocampus can code for multiple forms of content more flexibly than previously thought, the importance of these distributed codes for later retrieval of content-specific details is still unresolved. The current study investigates content-sensitive reinstatement at retrieval across MTL subregions. By training a neural classifier to identify content-specific patterns in the encoding data and testing its generalization performance in retrieval data, we have found the first evidence for reinstatement of multiple forms of event content that predicts later source memory performance. We also provide evidence for distinct functional roles for MTL cortical subregions and hippocampal subregions in coordinating this reinstatement to guide behavior.

Recent papers have examined the role of cortical reinstatement in supporting the retrieval of previous episodic memories. Cortical regions across the whole brain, most notably in occipital cortex but including many others, have consistently demonstrated reinstatement effects that were stronger for accurate versus inaccurate memory judgments (Johnson et al., 2009; Ritchey et al., 2012). However, the convergence of widespread sensory inputs to substructures within the MTL cortex have informed the predominant view of memory processing that these regions are critical for encoding many forms of

event content, and providing this information to hippocampus to guide present behavior. The current report directly interrogates the unique contributions of MTL subregions to the retrieval of episodic details.

In anterior MTL cortex – roughly corresponding to PRc – we observed that reinstatement of face encoding patterns was stronger during accurate face recognition versus forgotten trials. While recent univariate (Hannula et al., 2013) and multivariate (Martin et al., 2013; O'Neil et al., 2013) approaches have implicated PRc in the processing of visual face and object stimuli at retrieval, these studies do not establish that the discrimination between different visual face/object retrieval patterns arise from anything more than the specific goals of the retrieval task being performed by the subjects. Here, our findings clearly link the consistent pattern of activation evoked by successful face retrieval to the brain state present during initial face imagery. Furthermore, we find evidence for content-specific reinstatement that indicates processing in PRc specifically for faces, and not just visual objects or items versus context in general.

In contrast to anterior MTL cortex, the posterior MTL cortical region corresponding to PHc is known to be preferentially active during visuospatial scene encoding (Liang et al., 2012; Staresina and Davachi, 2011) and retrieval (Hannula et al., 2013). In the latter case, middle and posterior parahippocampal gyrus demonstrated higher pattern similarity between patterns evoked during scene recollection and patterns evoked during scene viewing with no corresponding object recollection. While these findings demonstrate consistency in the distributed patterns evoked by contextual scene

retrieval and scene viewing, their design did not permit linking of these patterns to encoding states present during the initial experience. Moreover, it relies on self-reported behavioral measures of recollection versus familiarity. Related findings investigating cortical reinstatement of scene information across the whole brain also employs this self-reported recollection measure (Ritchey et al., 2012). While this study implicated MTL cortex in representing an item level match between encoding and retrieval patterns, their design employed visually presented scenes during both phases of the experiment and unsurprisingly a wide range of visual occipital areas were also identified. Given the strong encoding responses of MTL subregions to visual object and visual spatial input, we controlled the perceptual input at encoding and retrieval by using only visual word cues. Critically, this method ensured that content-based reinstatement effects resulted from purely mnemonic representations evoked in MTL during mental imagery and reactivated during source recognition. This approach revealed a signature of distributed pattern reinstatement in middle and posterior MTL cortex that specifically reflected a objective source recognition measure of scene retrieval. These findings are convergent with a similar recent report from Staresina et al. (2012), which also implicate PHc in reinstatement of scene detail when comparing scene-word and color-word associative encoding and retrieval patterns. While they also argue that this reinstatement is category-specific (relative to reinstatement of associated color information), their design relied on visually presented scene and color stimuli at encoding. Visual scene stimuli have been thoroughly demonstrated to engage distinctive encoding responses in PHc at the level of univariate responses and distributed patterns. Here, we find a category-selective in PHc

despite equivalent visual engagement during encoding and retrieval across all experimental conditions.

Together these findings are broadly complementary to univariate findings of selective encoding in PRc and PHc. However, given the recently demonstrated flexibility of PRc (Liang et al., 2012) and especially PHc (Diana et al., 2008; Liang et al., 2012) in representing multiple forms of event content through distributed encoding patterns, the content-selectivity of the observed reinstatement effects in MTL cortical retrieval responses might seem surprising. In posterior MTL cortex, there was an overall main effect of source memory, indicating a numerical mnemonic difference for reinstatement effects across all content even though subsequent pairwise analyses indicated that it was driven primarily by a scene-specific effect. Furthermore, analyses examining reinstatement but ignoring memory status additionally revealed significant reinstatement of object representations in anterior MTL cortex, and significant reinstatement of face representations in middle and posterior MTL cortex (all $p < 0.05$). While these observations are more parsimonious with previous accounts of distributed representations at encoding by suggesting a capacity for reinstatement of multiple forms of content across MTL cortex, the type of information reinstated to support successful retrieval is certainly more selective. Further research would be needed to understand whether reinstatement that does not differentiate mnemonic behavior reflects another unique function for content representations in MTL cortex.

In hippocampus, we found evidence for scene-specific reinstatement that was predictive of source memory for scenes across its entire anterior-posterior axis. These

findings are consistent with neuropsychological and animal literature that suggests hippocampus critically supports processing of visual scene information, particularly during spatial navigation. While univariate neuroimaging approaches have also described hippocampal engagement during spatial navigation (e.g., Suthana et al., 2009), it has been unclear whether hippocampus makes a specific distinction between scene information and other forms of visual content. Only the recent application of multivariate techniques to brain patterns acquired during scene viewing has revealed a specific distributed hippocampal code for scene information at the category level (Liang et al., 2012) and at the level of decoding individual scenes (Bonnici et al., 2012). The current findings provide the first crucial evidence that distributed patterns evoked to represent scene information at encoding are also reinstated to support their retrieval.

Having established that content-specific reinstatement in MTL subregions was greater during accurate versus inaccurate source judgments, we wondered whether the magnitude of reinstatement was diagnostic of how much episodic detail was retrieved. In our task we observed large individual differences in source recognition across the three forms of content (d-prime range: 1.0-2.8), and we predicted that the fidelity of reinstatement (measured as the magnitude of classifier evidence across all correct trials) reflected the likelihood that a participant would retrieve sufficient source detail to guide accurate source judgments. In posterior MTL cortex, we found that the magnitude of face and scene reinstatement predicted individual face and scene d-primes respectively (Table 3.1). This further supports a role for PHc in supplying contextual detail for multiple forms of content, and provides one possible reason why face reinstatement was

numerically but not significantly different in this region: the full range of source correct trials would include occasions where even the most minor reinstatement effect was sufficient to guide behavior.

In hippocampus we found that the magnitude of scene reinstatement in the middle and posterior segments was correlated with participants' d-prime for scene recognition. While these effects were more spatially specific than the reinstatement effects were observed by comparing memory status within individual subjects, they are consistent with recent findings that distributed visuospatial representations are most distinct in posterior hippocampus (Liang et al., 2012). They also join other reports that differentiate the function of anterior and posterior hippocampus, where the former supports general retrieval of contextual information while the latter supports detailed space representations (Fanselow and Dong, 2010). Notably, we found that complementary to our reinstatement findings in middle and posterior hippocampus, we also found a cluster of voxels in anterior hippocampus that was more engaged during successful content-general source recognition, whose activation predicted reinstatement magnitude in MTL cortex on a trial-by-trial basis.

More broadly, the finding that anterior hippocampal voxels tracked reinstatement magnitude is consistent with models of MTL function that suggest hippocampus leverages content representations stored in MTL cortex during episodic retrieval. However, the directionality of these interactions is still unclear. Recent findings focusing on cortex across the entire brain (Ritchey et al., 2012) and PHc specifically (Staresina et al., 2012) also found correlations between hippocampal activity and cortical

reinstatement, proposing that they reflect hippocampal pattern completion processes that drive reactivation of cortical patterns. However, an alternative possibility suggests an essentially reversed direction to these hippocampal-cortical interactions. According to this alternate view, reinstated content representations would convey critical information to the hippocampus, which uses this information to discriminate between choice behaviors.

To investigate these alternative hypotheses we focused on reinstatement effects during trials where participants made source errors. According to the view that pattern completion drives reinstatement, the strength of reinstatement would be especially driven by those trials where pattern completion was successful (source correct trials) while source errors would be reflected by below threshold influence from hippocampus. This modulation across different memory states is indeed what is observed in recent findings (Staresina et al., 2012; Ritchey et al., 2012) including our aforementioned results. However, if reinstatement instead reflects a signal that is read out by hippocampus to guide choice behavior, we would expect two additional things to be true: 1) cortical reinstatement would be predictive of participants' choices when they make errors in source judgments and 2) hippocampal engagement would additionally track reinstatement of participant choices across error trials. In this study we found that reinstatement in middle and posterior MTL cortex predicted participants' erroneous choices across all forms of content. To ensure that this effect could not be explained by the possibility that participants were "pattern completing" to the wrong form of content due to their idiosyncratic responses to the adjectives during encoding, we repeated the analyses using

only false alarm trials where the word being presented was novel but participants still responded with “face”, “object”, or “scene”. Critically, there was no perceptual overlap between false alarm trials and what they encountered at encoding that would facilitate “erroneous” pattern completion. Nonetheless, reinstatement continued to be predictive of participants’ choices during source errors in middle and posterior MTL cortex. Together these findings provide the crucial first insight into the nature of information that is evoked in the MTL not only when we successfully remember, but also when we forget.

Table 3.1

Table 1
Correlations between mean classifier output for the predicted content class with content-specific d-prime across participants

MTL Cortex

Segment	*Face Output*	*Object Output*	*Scene Output*
Anterior	0.42	-0.03	-0.05
Middle	0.42	-0.03	0.39
Posterior	**0.68****	0.24	**0.73****

Hippocampus

Segment	*Face Output*	*Object Output*	*Scene Output*
Anterior	0.25	-0.14	0.48
Middle	0.14	0.10	**0.59***
Posterior	-0.15	0.20	**0.65****

Single asterisk indicates significant correlations at p < 0.05
Double asterisks indicate significant correlations at p < 0.01

Chapter 4. Disambiguation of Overlapping Experiences

DISTRIBUTED MEDIAL TEMPORAL LOBE REPRESENTATIONS REFLECT DISAMBIGUATION OF OVERLAPPING EVENTS

The majority of text and figures in this section were prepared as a manuscript for submission to *Journal of Neuroscience.*

Co-author contributions: A. R. Preston is my P.I. and helped write the article.

Abstract

A fundamental challenge for memory representation is that while many events are unique, individual events share many overlapping features. Pattern separation in the hippocampus is thought to be an important mechanism for disambiguating similar events in memory. By encoding overlapping events as separate non-overlapping representations, pattern separation ensures that our memories are more distinct than their perceptual similarities would predict. As evidence for pattern separation, research has focused on changes in mean signal in hippocampus during disambiguation of stimuli with slight visual differences or identical stimuli presented within different sequences. However, this approach does not directly measure how hippocampal representations for the same stimulus presented in two different episodic contexts are made distinct through pattern separation. In a high-resolution fMRI experiment, we used multi-voxel pattern analysis (MVPA) to index the distinctiveness of neural representations for events requiring

disambiguation. Participants were scanned while they viewed well-learned four element sequences comprised of face (F), object (O), and scene (S) stimuli. These sequences began and ended in a face or scene (FOOF, FOOS, SOOF, SOOS), and were either overlapping (OL) or non-overlapping (NOL). OL sequences always showed items in the same order and shared the object in the third position with one other OL sequence. By contrast, NOL sequences did not share any elements with other sequences. To measure pattern separation, we used MVPA to compare classifier accuracy for the overlapping objects in OL sequence pairs with classifier accuracy for the two different objects in the corresponding third position of NOL sequence pairs. Critically, if pattern separation supports sequence disambiguation by encoding unique representations for the overlapping objects, classification of OL objects should be better than NOL objects even though the OL objects were physically identical. We found better decoding of OL objects relative to NOL objects in PHC and posterior hippocampus. Moreover, we found that patterns in ventral temporal cortex tracked the identity of upcoming stimuli more accurately during OL objects relative to NOL objects. Finally, we found that posterior hippocampus was more functionally connected with ventral temporal cortex particularly during overlaps. Together, these findings suggest that distinct event representations for overlapping objects emerge through pattern separation in hippocampus, and that these representations may index perceptual features in neocortex that are important for predicting upcoming events.

Introduction

Despite the richness of our senses, much of our experiences are composed of repeated exposure to highly similar or even identical stimuli. Leading memory theories posit that representational changes occur within the medial temporal lobe (MTL) to help resolve ambiguity. Specifically, pattern separation suggests that overlapping events are embodied by neural populations that are more distinct from one another than would be expected from their perceptual similarity. In tandem with pattern completion, which allows sensory input to evoke reactivation of related memory representations, these mechanisms prevent confusion in the face of ambiguous sensory cues and make accurate predictions of expected outcomes. Early rodent data established the necessity of hippocampus to disambiguation of odor sequences (Agster et al., 2002), and later work refined this view by finding that neurons in hippocampus exhibit unique firing patterns in response to the same odor when presented in different sequence contexts (Ginther et al., 2011). However, current understanding of MTL circuitry suggested that distinct subregions might make different contributions to disambiguation. Electrophysiological work in rodents showed that hippocampal subfields played complementary roles in representing changes in a spatial environment, where minor and major shifts were reflected by changes in the firing dynamics of dentate gyrus and CA3 respectively. In humans the contributions of MTL subregions are less defined. Multiple fMRI studies have indicated a central role for hippocampus in learning overlapping sequences of faces and successfully navigating overlapping mazes (Kumaran and Maguire, 2006; Brown et al., 2010). However, high-resolution studies have asked whether hippocampal subfields

respond differentially to highly similar objects and found results consistent with pattern separation across multiple subfields (Bakker et al., 2008; Kirwan and Stark, 2007; Lacy et al., 2010). Moreover, most studies to date have employed univariate analyses while the animal literature would suggest that disambiguation is more likely to be reflected by diverging spatial patterns of neurons within the same region (e.g. Ginther et al., 2011; Leutgeb et al., 2007). To date, one study has used a multivariate pattern-based approach with human fMRI and found that posterior hippocampus could accurately discriminate between objects that appeared in two different sequence contexts. However, the dimensions by which neural pattern diverge during overlapping events are unknown. Across several memory domains, neural reinstatement in the MTL is a mechanism for retrieving past memories that are similar or related to current sensory input. Here we test the hypothesis that human MTL subregions perform complementary functions in support of disambiguation, and provide a link to memory-based prediction through the reinstatement of upcoming content.

Materials and Methods

Subjects

28 right-handed adults (aged 18-35) were recruited from the community around the University of Texas at Austin. Four participants were excluded during fMRI preprocessing for excessive motion during scanning, and one participant was excluded for not completing all phases of the experiment. 23 total subjects were used for behavioral and fMRI analyses.

Apparatus

The experiment was programmed in MATLAB, using the Psychophysics Toolbox for stimulus presentation. The order of trials were generated using OptSeq, an algorithm optimized for maximum efficiency of fMRI contrasts (Dale, 1999). All fMRI analyses were conducted on parallel compute clusters at the Texas Advanced Computing Center. fMRI preprocessing steps and general linear models were conducted with Advanced Normalization Tools (ANTS) and fsltools. Multivariate pattern classification analyses were performed with the PyMVPA toolbox.

Behavioral Paradigm

Stimuli were grayscale images of 12 faces belonging to famous celebrities, 12 famous landmarks (scenes), and 22 commonly encountered objects. These stimuli were organized into 12 four-item sequences in the following configurations: face-object-object-face (FOOF), face-object-object-scene (FOOS), scene-object-object-face (SOOF), and scene-object-object-scene (SOOS) (Figure 4.1a). Four sequences assigned to an overlapping (OL) condition, so that they shared stimuli; specifically, the FOOF and FOOS sequences shared the same object in the third position, and the SOOF and SOOS sequences also shared an object in the third position. Another set of sequences did not share any stimuli and comprised a non-overlapping (NOL) condition. A set of Random sequences did not share stimuli, and the positions of each item (apart from the third position) were shuffled across Random sequences in each run, so that no particular

sequence was ever observed more than once. The assignments of face, object, and scene stimuli to each sequence were randomized across 6 different configurations, counterbalanced across participants to ensure that they did not all view the exact same set of sequences.

In each run of the experiment, participants viewed all 12 sequences one time. The order of OL, NOL, and Random sequences were calculated to maximize the efficiency of estimating contrasts between the conditions. Trials consisted of a study phase and a probe phase. In the study phase, each item of a sequence were presented for 0.5 seconds, followed by 7.5 seconds of a central fixation cross. Immediately after the study phase, one of the items is presented again as a probe stimulus for 4 seconds. Participants were instructed to rapidly indicate via button press which position the image had appeared in during the study phase.

Figure 4.1a: Experimental design. For each sequence type, four different sequences were constructed: one face-object-object-face (FOOF), one face-object-object-scene (FOOS), one scene-object-object-face (SOOF), and one scene-object-object-scene (SOOS) sequence. Each 0.5 s stimulus presentation was followed by a 7.5 s fixation cross. At the end of each sequence, a probe image was shown to test participants' memory for item position. As shown in the figure, in the Overlapping (OL) sequence condition the FOOF and FOOS sequences were paired so that they shared the 3rd position object. Likewise, the SOOF and SOOS sequences were paired so that they shared the 3rd position object. Although the Non-overlapping (shown here) and Random sequences did not share any stimuli, the FOOF/FOOS and SOOF/SOOS sequences were still treated as paired for neuroimaging analyses.

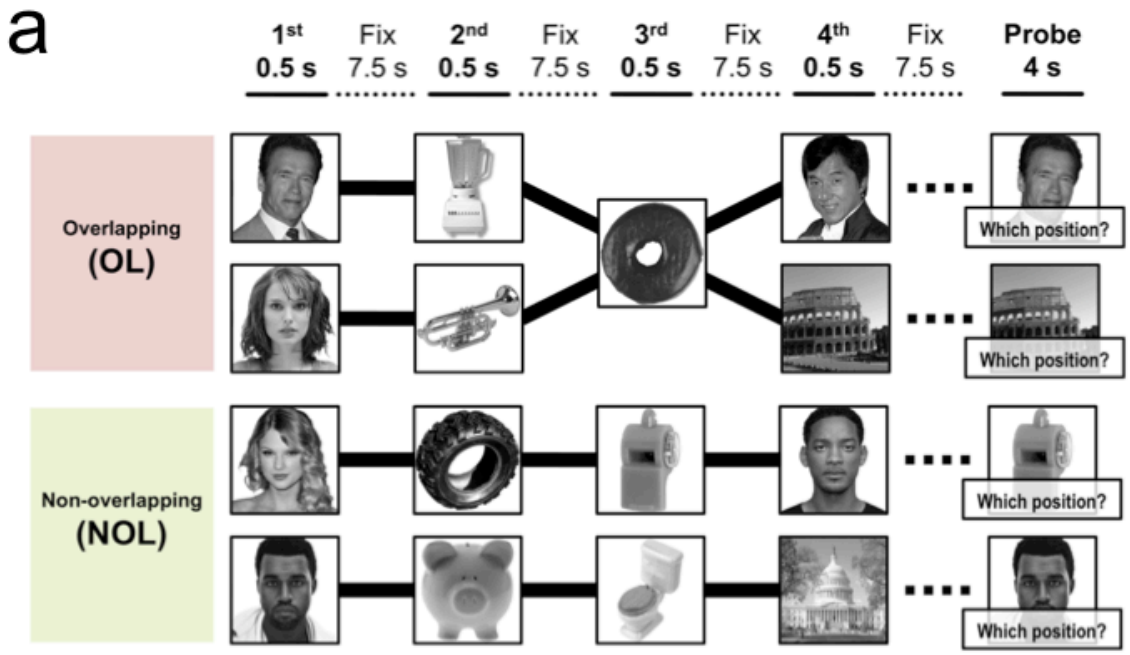
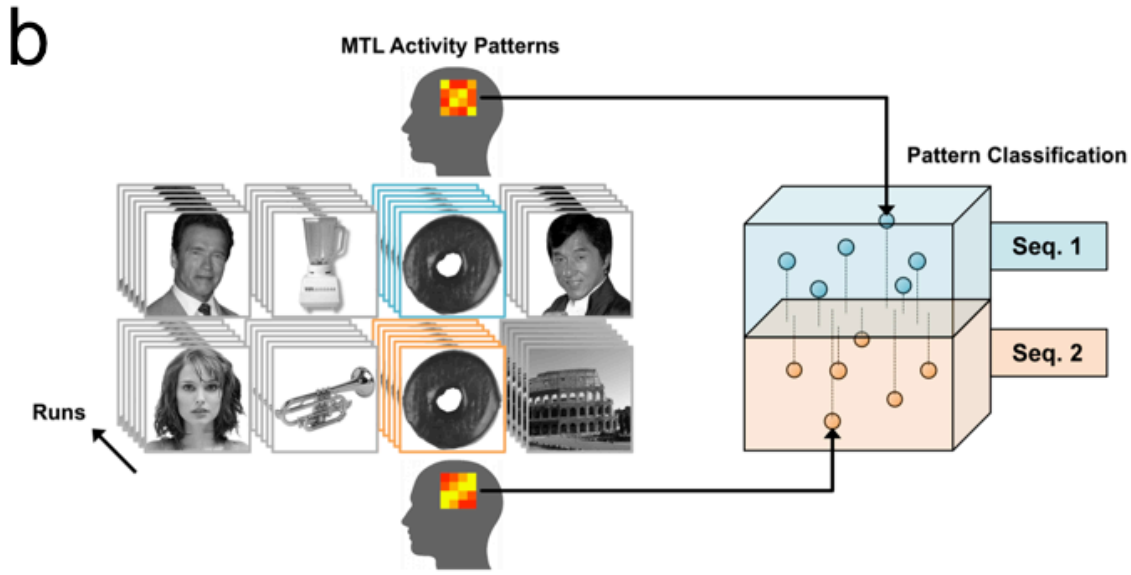


Figure 4.1b: Multivariate classification scheme. Multivariate decoding of sequence context was conducted using a cross-validation scheme. For each sequence pair, the activation patterns evoked by the third position objects are labeled according to whether they belong in sequence 1 (highlighted in blue) or in sequence 2 (highlighted in orange). In each fold of the cross-validation, 3rd position objects from 5 of the 6 runs were used for classifier training, and the objects from the remaining run were used for generalization testing. Across the 6 cross-validation folds, each run would be used for testing. The classification accuracies across all 6 folds were averaged to obtain the decoding accuracy for that sequence pair. For each sequence type, we averaged the classification accuracies of the FOOF/FOOS and SOOF/SOOS sequence pairs to obtain the decoding accuracies for OL, NOL, and Random sequences.



Behavioral Pre-training and Testing

In a session outside of the scanner, participants were pre-trained on an accelerated version of the sequence viewing paradigm. In this version, items are separated by a 2.5 second fixation cross (rather than 7.5 s). Across six runs, participants viewed each of the 12 sequences six times.

Inside the scanner, participants were trained on an additional six runs of the sequence viewing paradigm using the normal timing (7.5 second fixations). After the scanned session, we tested participants' memory by asking them to place scrambled versions of each sequence in the correct order. Following the experiment, participants also filled out a questionnaire to assess their awareness of how each sequence type differed from one another.

fMRI Acquisition

Participants' anatomical and functional brain images were acquired using a 32-channel head-coil on a Siemens 3.0 T Skyra MRI scanner housed at the University of Texas at Austin. Head stabilization was provided by foam padding. A T1-weighted MPRAGE structural image was acquired in 192 1-mm thick sagittal slices, encompassing the entire brain (1 x 1 mm in-plane resolution). To allow high-resolution visualization of hippocampal subfields and MTL cortices, we acquired two T2-weighted structural images (0.43 x 0.43 mm in-plane resolution) as 60 1.5-mm thick oblique coronal images oriented perpendicular to the main axis of hippocampus. These T2-weighted coronal images would later be averaged in preparation for manual demarcation of MTL ROIs. A

final T2-weighted structural image (0.43 x 0.43 mm in-plane resolution) was acquired as 42 1.5-mm thick oblique axial images oriented parallel to the main axis of the hippocampus, allowing high-resolution visualization of all structures from which functional data would be taken.

Functional brain images were acquired using a multi-band EPI sequence (TR = 2000 ms; TE = 32 ms; flip angle = 73°; field of view = 21.6 cm; 1.5 x 1.5 x 1.5 mm resolution). To maintain a 2000 ms TR at such high resolution, the number of slices was restricted to 38 but oriented parallel to the main axis of the hippocampus to ensure complete imaging of the medial temporal lobe and ventral temporal cortex. A total of 1,464 functional volumes were acquired for each participant over 6 scanning runs. To allow for T1 stabilization, the first trial of each run did not commence until 8 seconds (4 TRs) after the beginning of the functional acquisition. These 4 dummy volumes prior to the first trial were discarded and were not used in subsequent analyses.

Data Analysis

fMRI Preprocessing. Functional volumes were first realigned to the first volume in each time series using MCFLIRT to correct for motion within runs. To account for motion across the experiment, each run's time series was coregistered to the first volume of the first run (ie. functional reference image) using ANTS. To ensure that all structural images were properly aligned to each other, the sagittal MPRAGE and two oblique coronal T2-weighted images were coregistered to the oblique axial T2-weighted image using ANTS. A final ANTS step coregistered the oblique axial T2-weighted image to the

functional reference image, and the resulting parameters were applied to each of the other structural images.

Group Template Generation. In preparation for generation of a high-resolution template image, the two oblique coronal T2-weighted images were averaged to reduce noise. We inspected each participant's mean coronal T2-weighted image and selected the 10 that were the most typical with respect to the rest of the group and also contained minimal artifacts and noise. The selected images were then combined into a single high-resolution group template through the use of non-affine transformations calculated by ANTS. This technique provided a common group structural image that does not have the usual limitation of low spatial resolution; hippocampal subfields were able to be easily differentiated on this image.

Regions-of-interest. Hippocampal subfields (CA1, CA2/3/DG, and subiculum) and subregions of MTL cortex (perirhinal cortex, entorhinal cortex, and parahippocampal cortex) were manually demarcated on the high-resolution group template. Each hippocampal subfield was divided into anterior and posterior segments, with the uncal apex marking the transition between the two segments. Perirhinal cortex and entorhinal cortex were combined into a single region representing the anterior MTL cortex. The point located 4.5 mm posterior to the uncal apex in hippocampus was chosen as the standard landmark indicating the transition from anterior MTL cortex to parahippocampal cortex. In sum, this procedure resulted in 8 MTL ROIs: anterior and posterior CA1, CA2/3/DG, and subiculum, anterior MTL cortex, and parahippocampal cortex.

Two additional ROIs (fusiform gyrus and inferior temporal cortex) were generated by applying a FreeSurfer-based algorithm to the high-resolution T1 structural image. They were combined to form a single region representing ventral temporal cortex. In total, the manual and automatic ROI-generation procedures resulted in nine bilateral ROIs.

ANTS warping of functional time series. To enable group-level inferences on brain images, each participant's functional time-series was warped to match the high-resolution T2 group template. The first step in this process was to manually draw hippocampus and MTL cortex on each participant's mean T2 coronal anatomy. Hippocampus and MTL cortex were then divided into anterior and posterior segments according to the same landmarks described in Regions-of-interest. Using ANTS, a label-guided non-affine transformation was calculated from the participant's T2 anatomy to the group template that respected the boundaries of anterior and posterior hippocampus and MTL cortex as demarcated in each anatomy. The parameters of this transformation were then applied to the participant's time series images.

MVPA Classification. In preparation for MVPA analyses, we first performed a spatial smoothing of the time series using a 3-mm FWHM Gaussian kernel. We then applied a high-pass filter to the time series to remove low frequency drift (longer than 100 s). Finally, we z-scored the voxelwise time series along the time dimension. Each time point of the fully processed time series was labeled according to the particular sequence and item position that was presented at the time. In particular we were interested in whether different MTL regions could disambiguate between the 3rd position

items of each sequence pair. For a given pair of sequences, we labeled the peak volumes (TRs 3 and 4) of the 3rd item position according to its pair membership across all experimental runs (Figure 4.1b). A neural classifier using a linear support vector machine was trained and tested on these timepoints, masked by a ROI, and their labels in a leave-one-run-out cross-validation scheme. The classification accuracies from each of the 6 cross-validation folds were averaged to obtain the disambiguation score for one sequence pair. The disambiguation scores for the two sequence pairs of a given sequence type were averaged for each ROI and each participant.

First, we calculated two-tailed Student's t-tests to determine whether classification accuracy for any of the sequence types were significantly different from chance level (50%). To test whether the sequence discrimination differed according to the degree of information overlap, we subjected the disambiguation scores from each ROI to a repeated measures ANOVA with sequence type as a 3-level factor (OL, NOL, and Random) and participants as a random factor. For ROIs that demonstrated a significant main effect of sequence type, we calculated pairwise comparisons to determine which sequence types contributed most to sequence disambiguation.

Pattern Similarity Analysis. We predicted that anticipatory reactivation of upcoming content during overlapping items would be characterized by increased neural pattern similarity between a 3rd position object and the 4th position item belonging to the same sequence, relative to the 4th position item belonging to the other member of the sequence pair. Therefore, we calculated the Pearson's correlation between the linearized voxel patterns evoked by the 3rd position object and those evoked by the 4th position

faces and scenes belonging to each member of the sequence pair. To ensure that pattern correlations were not confounded by temporal autocorrelation between 3rd and 4th position items within the same run, we only calculated correlations between 3rd position objects in a given run (e.g. run 1) with 4th position items in every other run (e.g. runs 2-6). The reinstatement score for a given 3rd position object was then calculated as the difference between its similarity to 4th position items within the same sequence and its similarity to 4th position items in the other sequence, averaged across runs. Reinstatement scores were averaged across all 3rd position items within sequence type for each ROI and each participant.

First, we calculated two-tailed Student's t-tests to determine whether reinstatement scores for any of the sequence types were significantly different from zero. To test whether reinstatement differed according to the degree of information overlap, we subjected the disambiguation scores from each ROI to a repeated measures ANOVA with sequence type as a 3-level factor (OL, NOL, and Random) and participants as a random factor. For ROIs that demonstrated a significant main effect of sequence type, we calculated pairwise comparisons to determine which which sequence types demonstrated the greatest degree of anticipatory reinstatement.

Psychological-physiological interaction modeling. We predicted that MTL regions encoding overlapping episodes would especially interact with neocortical regions encoding the salient perceptual features of upcoming information to support disambiguation. Therefore, we calculated two psychological-physiological interaction (PPI) models to determine whether voxelwise activation within MTL was increasingly

correlated with activation of ventral temporal cortex during the 3rd item of 1) Ordered sequences relative to Random sequences or 2) Overlapping relative to Non-overlapping sequences.

The PPI model for Ordered versus Random sequences was calculated as a general linear model (GLM) for each voxelwise time series within the MTL. The GLM contained psychological regressors, physiological regressors, and PPI regressors. The first regressor was positively weighted (+1) for the 3rd item of OL and NOL (Ordered) sequences and negatively weighted (-1) for the 3rd item of Random sequences. The second regressor was positively weighted for the 3rd item of all sequences. Additional regressors were included that independently contained positive weights for each of the other item positions within each sequence type (e.g. 1st position in OL, 2nd position in OL, 4th position in OL, 1st positions in NOL, 2nd positions in NOL, etc.) One regressor captured all of the probe phases. The physiological regressor was represented by the mean time series of the ventral temporal cortex seed ROI. The critical PPI regressor was calculated as the interaction between the first regressor representing 3rd items in Ordered versus Random sequences and the seed ROI. Finally, additional confound regressors were included to account for the participant's head motion in the scanner.

At the first level, this PPI model was calculated independently in each run of a participant's time series data. A contrast was calculated to determine the parameter estimate for the critical PPI regressor (the Ordered > Random x seed ROI interaction) for each voxel in MTL. A second level model, including the above contrast, was calculated with each of a participant's first level models as inputs with runs as a fixed effect. Finally,

a group model was calculated with each participant's second-level PPI contrast to determine which MTL voxels demonstrated a significant interaction effect across the group.

Our second PPI model was constructed similarly to the first, with the following regressors: 1) positive weights for the 3rd item of OL sequences and negative weights for the 3rd item of NOL sequences, 2) positive weights for the 3rd item of both OL and NOL sequences, and 3) positive weights for each item position of all sequences that were not accounted for by the first two regressors. The critical PPI regressor for this model was therefore calculated as the interaction between the first regressor representing 3rd items in OL versus NOL sequences and the seed ROI. Regressors for the probe phase and head motion were included just as they were in the first model.

To ensure proper small volume correction for MTL voxels showing significant interaction effects at the group level, we employed a program called AlphaSim to estimate the probability of finding each cluster of significant voxels. Across 10,000 Monte Carlo simulations, we generated a null distribution of voxel cluster sizes within the MTL mask given a per-voxel p-value of 0.01. For each PPI model, every cluster of contiguous voxels of p-value less than 0.01 was tested against this null distribution, and was eliminated if it fell outside the top 5% of randomly simulated clusters.

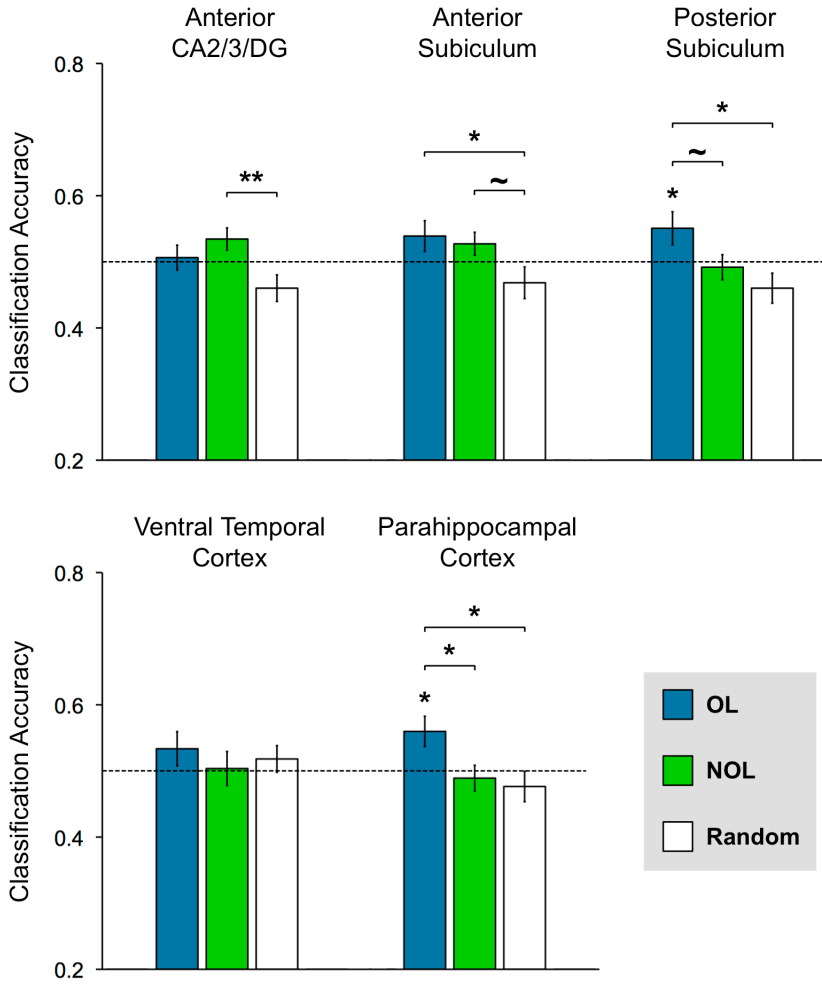
Results

Disambiguation of sequence context

First we interrogated which regions of MTL and surrounding cortex disambiguated between sequence contexts. A repeated measures ANOVA identified multiple regions wherein classification accuracy of sequence contexts actually differed by sequence type: parahippocampal cortex, anterior CA2,3/DG, anterior subiculum, and posterior subiculum (Figure 4.2, all $F(2,22) > 3.27$, all $p < 0.047$). We predicted that objects in the third position would be by disambiguated by either the perceptual features of the objects themselves or their unique sequence contexts. Pairwise comparisons revealed that anterior CA2,3/DG and anterior subiculum demonstrated higher classification accuracy for NOL sequences -- where the objects and sequence contexts are distinct and consistent -- relative to Random sequences -- where the objects are distinct but the sequence contexts are not (CA2,3/DG: $t(22) = 2.86$, $p = 0.009$, subiculum: $t(22) = 1.89$, $p = 0.072$). Critically, we also predicted that MTL would be especially involved in heightening the distinctiveness of sequence representations in the face of overlapping content. In such regions, we expected that classification accuracy of OL sequences -- where objects are overlapping and the sequence contexts must be made more distinct -- would be greater than that of Random sequences. Moreover, NOL sequences -- despite having distinct sequence contexts -- would not be treated differently from Random sequences because there are no overlapping elements to drive separation of sequence representations across learning. Indeed, parahippocampal cortex and posterior subiculum demonstrated a pattern of results most consistent with disambiguation specifically during

overlapping sequences, where classification accuracy for OL sequences was significantly greater than chance (all $t(22) > 2.12$, all $p < 0.045$) while classification accuracies for NOL and Random sequences were not (all $t(22) < -0.43$, all $p > 0.1$). Furthermore, pairwise comparisons revealed no significant difference between NOL and Random sequences for either region (all $t(22) < 1.29$, all $p > 0.21$), while classification accuracy for OL sequence was significantly greater than Random sequences (all $t(22) > 2.37$, all $p < 0.027$) and even NOL sequences (parahippocampal cortex: $t(22) = 2.59$, $p = 0.017$, posterior subiculum: $t(22) = 1.78$, $p = 0.089$).

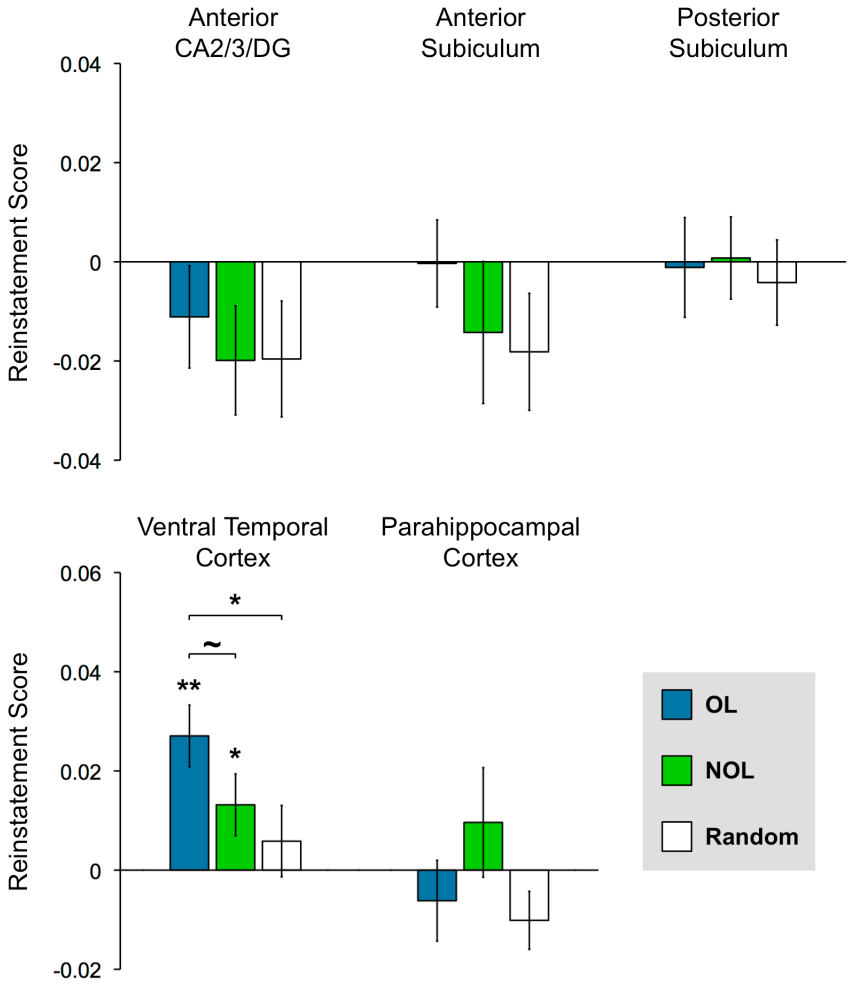
Figure 4.2: Disambiguation of sequence context from 3rd position objects. Classification accuracy for Overlapping object pairs (blue bars), Non-overlapping object pairs (green bars), and Random object pairs (white bars) are shown for each region-of-interest. The height of each bar represents the mean classification accuracy across participants, and the error bar represents the standard error. The dashed line represents the chance level accuracy of 0.5. Asterisks centered above a bar represent significantly greater or less than chance classification accuracy ($p < 0.05$). Asterisks with a bracket represent a significant pairwise difference between sequence types (single asterisk: $p < 0.05$, double asterisk: $p < 0.01$).



Reinstatement of upcoming content

Secondly, we asked whether MTL and surrounding cortex represented information about upcoming content to support disambiguation during sequence overlaps. For each ROI, we used PSA to calculate a reinstatement score reflecting how much each object in the third position evoked reactivation of the fourth item from the same sequence, relative to how much it evoked reactivation of the fourth item from the other sequence. We used a repeated measures ANOVA to determine whether the reinstatement scores in each ROI differed by sequence type. Out of all the regions, only ventral temporal cortex demonstrated a trend for a main effect of sequence type (Figure 4.3, $F(2,22) = 2.77$, $p = 0.074$). We suspected that the reinstatement of the correct future outcome would be strongest in the face of perceptual ambiguity. Pairwise comparisons showed that while the reinstatement scores in ventral temporal cortex were significantly greater than zero for both OL and NOL sequences (all $t(22) > 2.11$, $p < 0.047$), only OL sequences demonstrated significantly higher reinstatement scores relative to Random sequences (OL vs. Random: $t(22) = 2.28$, $p = 0.033$, NOL vs. Random: $t(22) = 0.74$, $p = 0.47$) and a trend for higher reinstatement scores relative to NOL sequences ($t(22) = 1.69$, $p = 0.1$).

Figure 4.3: Reinstatement of content-specific details during the presentation of 3rd position objects. Reinstatement scores (see Materials and Methods) for Overlapping object pairs (blue bars), Non-overlapping object pairs (green bars), and Random object pairs (white bars) are shown for each region-of-interest. The height of each bar represents the mean classification accuracy across participants, and the error bar represents the standard error. Asterisks centered above a bar represent a reinstatement score significantly different from 0 ($p < 0.05$). Asterisks with a bracket represent a significant pairwise difference between sequence types (single asterisk: $p < 0.05$, double asterisk: $p < 0.01$).

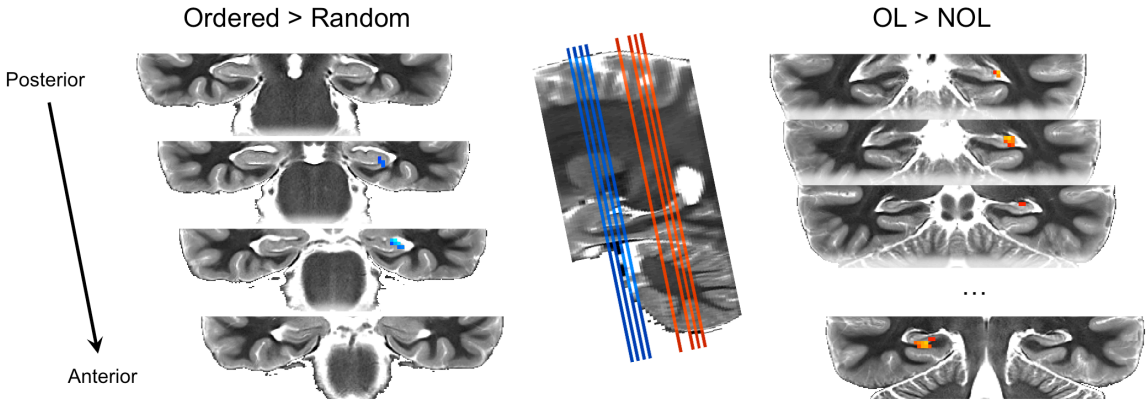


Functional connectivity between ventral temporal cortex and MTL

Having identified regions that independently disambiguated overlapping sequences or reinstated upcoming content, we hypothesized that these regions work as a cohesive network to resolve ambiguity in our experiences. Reactivated information may alternately serve as crucial cues that help identify the current sequence context, or may reflect a decision signal resulting from the disambiguation process as a way to guide behavior. To test the hypothesis that connectivity between regions that perform disambiguation and reinstatement might increase specifically to resolve ambiguous input, we employed psychogical-physiological interaction (PPI) models to assess whether activity in any region of the MTL tracked activity in ventral temporal cortex during the crucial third position objects.

Our first PPI model found one cluster in the anterior hippocampus whose activity was significantly more correlated with activity in ventral temporal cortex during third position objects in OL and NOL (Ordered) sequences relative to Random sequences (Figure 4.4). This cluster included voxels in CA2,3/DG and CA1. Importantly, we also calculated a second PPI model focused on identifying clusters whose activity was significantly more correlated with activity in ventral temporal cortex during third position objects in OL sequences relative to NOL sequences. This analysis revealed two significant clusters located entirely within posterior hippocampus that included voxels in left CA2,3/DG, right CA1, and right subiculum. There was no overlap between the clusters identified in these two models.

Figure 4.4: Functional connectivity between ventral temporal cortex and MTL during the presentation of 3rd position objects. Left, Voxels in MTL significantly more correlated with ventral temporal cortex during Ordered sequences relative to Random sequences. Permutation testing revealed significant clusters restricted to anterior hippocampus (all cluster-level $p < 0.05$). Right, Voxels in MTL significantly more correlated with ventral temporal cortex during OL sequences relative to NOL sequences. Permutation testing revealed significant clusters restricted to posterior hippocampus (all cluster-level $p < 0.05$). Middle, Sagittal section shows the positions of each coronal slice included in the left and right panels.



Discussion

In the present study, we used high-resolution fMRI and multivariate approaches to test the hypothesis that sequence disambiguation is linked to diverging activation patterns in MTL and critically supports memory-based prediction under ambiguous circumstances. Using a paradigm inspired by rodent work with overlapping odor sequences, we used sets of overlapping, non-overlapping, and random sequence pairs to observe what happens during two identical objects, where visual input is exactly matched but the related context and outcomes are completely different. We extend beyond previous findings to show that activity in different MTL subregions make unique contributions to disambiguation, and that the underlying representational changes arising from repeated encounters with overlapping events can be important signals for memory-based prediction.

PHC and hippocampus disambiguate sequence context

Using multivariate decoding techniques, we found evidence that posterior subiculum and parahippocampal cortex can accurately decode two visually identical objects based on their sequence context. Moreover, their decoding accuracies for the two visually distinct objects in the NOL and Random conditions did not rise above chance level, and did not differ according to whether the objects were linked to consistent sequence contexts. This pattern of results indicate that it is the presence of overlapping content in the OL condition that ultimately enhances the distinctiveness of the representations underlying each sequence. These findings are consistent with other

reports that have indicated greater hippocampal and parahippocampal activity during sequence learning and specifically during overlapping events (Kumaran and Maguire, 2006; Brown et al., 2011). Here we provide novel evidence that not only are they involved in processing overlapping content, but they also represent the critical sequence contexts with unique distributed patterns. The distinctiveness of contextually related patterns increases during overlaps, which is consistent with pattern separation.

Parahippocampal cortex has been well established as a region that responds to context or predicts context memory across many domains (Staresina and Davachi, 2008; Davachi, 2006). Pattern similarity approaches have found that different visuospatial contexts can be represented by distinct patterns of activity (Staresina et al., 2012; Libby et al., 2014). More specific to temporal or sequence memory, multiple studies have found that PHC activity predicted memory for word triplets (Tubridy and Davachi, 2011) and sequences of objects respectively (Jenkins and Ranganath, 2010). However, this study is the first demonstration of a consistent code in PHC for different temporal sequences. Furthermore, we show that PHC patterns become even more distinct, in effect sharpening the sequence representations, in response to overlapping objects. The exact purpose of this sharpening is still an open question; however, based on previous work we suspect that they reflect the emphasis of salient perceptual features important to each sequence context, while hippocampus may be more necessary in forming the crucial link between overlapping objects and the correct sequence representations in PHC.

The importance of hippocampus to disambiguation is well-supported across a variety of tasks. For example, using a virtual reality task wherein participants navigate a

series of overlapping and non-overlapping mazes, Brown et al. (2011) found that hippocampal activity was enhanced during the critical overlapping juncture between two mazes compared to non-overlapping mazes. Likewise, using a non-navigational paradigm where participants viewed sequences of faces, Kumaran and Maguire (2006) found that activity in posterior hippocampus uniquely predicted rate of learning of overlapping but not non-overlapping sequences. Multivariate decoding schemes have found that distributed patterns in hippocampus could discriminate between multiple movie clips despite significant overlaps in content (Chadwick et al., 2011). Our results are convergent with these previous findings. However, the specificity of disambiguation effects to posterior subiculum may be somewhat surprising, given that reports of disambiguation and pattern separation in rodent work primarily implicate CA3 or dentate gyrus (e.g., Leutgeb et al., 2007; Ginther et al., 2011). The majority of findings using human fMRI also focus on CA3 and dentate gyrus; some of the earliest research on pattern separation found that voxels in CA3 activated to objects that were new but highly similar to old objects as though they were novel (Bakker et al., 2008; Lacy et al., 2010). More recently, Chadwick et al. (2014) replicated their earlier findings within individual subfields, and found that the decoding of overlapping movie clips was specific to CA3. We note, however, that much of the other human research on pattern separation have either treated hippocampus as a single homogeneous region, or omitted subiculum from a priori testing entirely. Instead, we can look to some recent work by Suthana et al. (2009, 2011), where participants were scanned as they navigated between storefronts in a virtual city block. fMRI analyses from this period yielded encoding-related activation in CA1 and

CA2,3/DG, which contrasts with a second phase in which participants were asked to retrieve the correct route to a storefront from either a single starting point or multiple starting points. Interestingly, the latter condition provided a naturalistic scenario in which disambiguation of overlapping routes was needed to perform the task. fMRI analyses from retrieval revealed that subiculum was more engaged during multiple versus single starting points. Given the extensive pre-training performed by participants in our own task, we suspect that subiculum may also be performing a retrieval-related function. While we cannot rule out the involvement of the CA subfields or dentate gyrus, we have shown here that the diverging representations arising from pattern separation extend to subiculum. Subiculum may play a central role in disambiguation by retrieving the specific content that can make each overlapping event unique.

We devote special attention to a recent result by Hsieh et al. (2014), which suggested that posterior hippocampal patterns evoked by overlapping sequences of objects were more anti-correlated when the overlaps were preceded by visually distinct objects relative to identical objects. This is indeed evidence that hippocampus can disambiguate sequence context based on memory even when visual input is matched, and dovetails nicely with prior work that showed distributed patterns in hippocampus can decode context in general. However, we argue that a key signature of pattern separation is that the representations of overlapping events should be more distinct than expected by chance. While the visual inputs are matched in the Hsieh et al. study, the sequence contexts are not. We have found results consistent with pattern separation by showing

that visually matched inputs can be disambiguated more reliably than two distinct objects even when their respective sequence contexts are matched.

Reinstatement of upcoming content

By design, one of the most distinguishing features of each sequence pair is the identity of the fourth item. We predicted that this information would be enhanced in memory as a consequence of pattern separation, and later be reinstated with even greater fidelity. We found that in ventral temporal cortex, the identity of upcoming content was reinstated through distributed activity during the third position of ordered but not random sequences. Critically, this reinstatement was greatest during overlapping sequences. Given that parahippocampal cortex had already demonstrated disambiguation-related activity and has been noted to be sensitive to visual categories such as faces, objects, and scenes (Brown et al., 2010, 2014; Diana et al., 2008; Liang et al., 2013), we had predicted that PHC might also demonstrate this pattern of results. On the other hand, the fourth position items are not only differentiated by category (face versus scene) but also by its identity as a famous face or landmark. It is possible that item-specific details would be reinstated in an earlier area such as ventral temporal cortex, which contains fusiform gyrus and inferior temporal cortex. Both of these regions have shown sensitivity to morphs of face or object stimuli (e.g., Gilaie-Dotan et al., 2010; Tootell et al., 2008), suggesting that they may encode detailed perceptual information rather than category identity. Regardless of which type of information is predominantly reinstated in this

paradigm, we have shown that crucial episodic details are enhanced to help us resolve ambiguity.

A hypothesis that emerged from the data thus far is that MTL subregions supporting disambiguation may facilitate or leverage reinstatement of upcoming content in neocortex. We found evidence in support of this hypothesis; PPI models showed that activity in voxel clusters within posterior hippocampus, including those within subiculum, were more correlated with activity in ventral temporal cortex specifically during the third position objects of overlapping sequences. Taken together, our findings support a model of disambiguation wherein distinct patterns of activity link overlapping memories in MTL with the details of predicted outcomes in the surrounding cortex.

Chapter 5. Discussion and Future Directions

The core interest of my research program is in understanding how different subregions of the MTL function in service of memory. Leading memory theories proposed that anatomically distinct MTL subregions were functionally differentiated on the basis of event content. And yet, research in humans has yet to come to a consensus on precisely how they represent the elements of our experiences. This section highlights and integrates the novel findings of my research program, which investigated how content is represented by distributed patterns in MTL during three common scenarios that are critical to the mnemonic experience: encoding, retrieval, and during ambiguous perceptual overlaps.

Resolving debates concerning content representation in MTL during encoding.

Animal research suggests neuronal populations in MTL are selectively responsive to stimulus material. Evidence from human neuroimaging, however, is mixed; object-sensitive perirhinal cortex (PRC) also responds to location information and scene-sensitive parahippocampal cortex (PHC) responds to highly contextual objects. Likewise, theories informed by animal research that the hippocampus plays a critical role in processing visual scene content have yet to be supported by human neuroimaging, which predominantly finds hippocampal responses that are content-general. A gap between the two literatures is that animal electrophysiology measures single neurons within a region

of interest, while standard univariate methods in human neuroimaging relies upon measuring the peak response across an entire subregion containing many millions of neurons. This univariate approach precludes the possibility that different stimulus materials may be encoded by a distributed network of neurons within a subregion, and that even sub-maximal responses may convey important event information in such a network. In order to bridge this gap in the conventional approach to content representation, I applied multivariate techniques (Norman et al. 2006) to a high-resolution fMRI dataset to characterize the distributed activity evoked by each content class with greater precision than previously possible.

My novel multivariate approach revealed several surprising findings. First, contrary to predominant theories that the responses in PRC and PHC would be selective for visual objects and visuospatial scene stimuli respectively, pattern classifiers applied to multivoxel patterns from both subregions were able to accurately discriminate between many classes of stimulus content (Liang, Wagner, and Preston 2013). I also used representational similarity analysis (RSA) (Kriegeskorte et al., 2008) to obtain in greater detail which stimulus classes evoked the most distinct patterns in each subregion. While I had predicted that PRC would most strongly discriminate faces from other stimuli, PRC in fact discriminated between faces and scenes from each other and from other content classes. PHC strongly discriminated faces and scenes from each other and from other content classes, but also exhibited discrimination between visual and auditory stimuli in general. Finally, I found that multivoxel patterns in posterior hippocampus discriminated

scenes from other stimuli, providing some of the first evidence for visuospatial scene representation in hippocampus outside of a navigational context.

Taken together, the findings of my first study demonstrated that the traditional view pitting content-specialized MTL cortex with content-general hippocampus was too simplistic. While PRC and PHC behaved similarly in that both regions could distinguish multiple types of stimulus material, they are likely to encode different aspects of any given face or scene. For example, PRC may be responsive to the particular perceptual features that distinguish one face from another, PHC may take that information to distinguish a face from other types of content. One way to test this hypothesis would be to compare the distributed activity of each region to individual face exemplars (which is unfortunately not possible in my block-design dataset); voxel patterns evoked by different faces in PRC may reflect item-level encoding and be anticorrelated from one another relative to PHC, where event content may be distinguished primarily at the category level. The finding that posterior hippocampus contains a distinct code for scene content converges with animal research where dorsal hippocampus is critical for encoding local information about space. This scene coding does not preclude other content-general hippocampal representations, however. Hippocampus is thought to encode episodic memories by integrating information across a wide variety of stimulus domains. In this study, univariate techniques revealed that hippocampus was more active during novel stimulus presentations relative to repeated ones regardless of content. We are presently testing the hypothesis that distinct hippocampal networks signal the novelty

of face and scene information by their functional connectivity with PRC and PHC respectively.

Relationship between memory reinstatement and mnemonic decision-making.

The theory of memory reinstatement is that a key signature of successful remembering is the reactivation of cognitive processes that were present during the initial learning experience (Johnson et al. 2009). While my first study demonstrated that content representations in MTL could distinguish multiple stimulus materials, I had yet to show that they have any bearing on later memory. My second study scanned participants during encoding and retrieval of different content-specific details to look for evidence of reinstatement in MTL and determine whether it is related to memory performance. In the memory encoding session, participants were prompted to mentally visualize a face, object, or scene that best matched trial-unique adjectives. In the subsequent memory retrieval session, participants were shown each of the previously viewed adjectives and performed source judgments on whether they had originally imagined a face, object, or scene.

I trained a pattern classifier on patterns evoked by each adjective during encoding according to the stimulus category the participants were instructed to visualize. When I tested this classifier with activation patterns evoked by adjectives during retrieval, I found that its accuracy in identifying the originally visualized stimulus category was higher for patterns evoked during successful source judgments. This finding broadly supported my hypothesis that successful remembering is reflected by reinstatement of

category information originally present at encoding. However, this effect differed by content class across MTL subregions, where the pattern classifier only exhibited performance increases between accurate versus inaccurate source judgments for faces in PRC, and scenes in PHC and hippocampus. Interestingly, a closer look at the classification unit outputs during errors in source judgment indicated that PHC activity accurately predicted the content class actually selected by participants. A speculative interpretation of this finding implicates a role for PHC in providing choice signals that are read out by hippocampus to guide behavior.

Given the multiple content representations found in MTL cortex in my first study, the selectivity of reinstatement effects observed here might seem surprising at first. However, we are looking at fundamentally different processes. My first study only scanned participants during incidental encoding. Cross-validated MVPA on the encoding data in this experiment confirmed that PRC and PHC accurately discriminated between face, object, and scene imagery trials (Figure 3.2). What these findings revealed is that the selectivity of reinstatement in MTL arose uniquely from retrieval- related processes. Extending the hippocampal scene selectivity observed in my first study, I found the first evidence that reinstatement in hippocampus is also specific to visuospatial scene information. Interestingly, these effects were not restricted to the posterior segment of hippocampus, and indicated that the entirety of hippocampus is brought to bear during retrieval of scene-related episodic details.

Content representations help resolve ambiguity and aid decision making.

Often we are confronted with options that are difficult to discriminate. After meeting with your friend, you come back to the parking garage and notice that a lot of the cars are the same make, model, and even color as your own. In the face of high perceptual similarity, we must rely on our memory to disambiguate overlapping items and avoid confusing situations like trying to drive off in a stranger's vehicle! Animal research and the existing anatomical framework suggest that hippocampus is crucial for maintaining the context in which an item was encountered. Furthermore, content representations throughout MTL may serve as important contextual signals if other event content were strongly associated with overlapping items. Importantly, the representations of overlapping items may themselves become more separated from one another in response to the demand of disambiguation. In this case, re-exposure of these items when given sufficient contextual signals will evoke more distinct neural responses than would be expected by chance. Reliance on univariate statistical approaches means that two perceptually overlapping objects studied in different contexts will always evoke the same object-related peak activation. However, distributed activity may better reflect divergent representations within a single region. I used MVPA to identify regions that accurately discriminate between two identical objects experienced within unique sequence contexts. In my third project, participants were scanned while they studied well-learned sequences of faces, objects, and scenes. Unbeknownst to the participants, some of these sequences shared overlapping objects despite having an otherwise unique sequence of stimuli.

Multivariate decoding of MTL activity evoked during sequence overlaps found that the same object in different sequence contexts were differentiated by distributed patterns in PHC and posterior hippocampus. Critically, neural discrimination of context was actually better for overlapping sequences than non-overlapping sequences, suggesting that MTL maintains more distinct sequence representations in response to perceptual ambiguity. Furthermore, each sequence ended in either a unique face or scene. Using pattern similarity analyses, I examined whether reinstatement of category-level information predicted upcoming stimuli immediately following the presentation of ambiguous objects. I found that while MTL did not demonstrate this pattern of response, ventral temporal cortex demonstrated stronger reinstatement during the overlapping sequences relative to the other conditions. Moreover, ventral temporal cortex and posterior subiculum showed the stronger functional connectivity during sequence overlaps relative to non-overlapping sequences. These novel findings highlight the interaction between the representations of current events being disambiguated in hippocampus and the content-specific representations of anticipated events in neocortex.

Future Directions

Collectively, my experiments reveal how MTL subregions transform rich sensory inputs into memory representations. My core research goal as a graduate student has been to link these content representations with memory-guided behavior during simple decisions (i.e., source memory judgments) and more complicated ones (i.e., discrimination of overlapping sequences). Importantly, the mechanisms that support

reinstatement and disambiguation are applicable in domains outside of episodic memory that interest me as a future post-doctoral fellow. I have shown that MTL resolves ambiguity in well-learned overlapping sequences through pattern separation of internal object representations. This type of representational change can have tremendous impact on the way we perceive novel information. For example, you are trading in your old car for this year's latest model. Upon arriving at the dealership you are confronted with a wide range of vehicles, but your brain rapidly compares them to stored memory representations and identifies one of them as being the same model car as your own. And yet, it simultaneously highlights design and styling features in this year's model that makes it perceptually distinct from your car. Memory can fundamentally alter perception, and the MTL is a critical node in the brain network that flexibly encodes new information based on our existing knowledge.

REFERENCES

Agster KL, Fortin NJ, Eichenbaum H. 2002. The hippocampus and disambiguation of overlapping sequences. J Neurosci. 22:5760-8.

Amaral DG, Insausti R, Cowan WM. 1983. Evidence for a direct projection from the superior temporal gyrus to the entorhinal cortex in the monkey. Brain Res. 275:263-277.

Amaral DG, Insausti R. 1990. Hippocampal formation. In Paxinos G (Ed.), The Human Nervous System (pp. 711-755). San Diego: Academic Press.

Aminoff E, Gronau N, Bar M. 2007. The parahippocampal cortex mediates spatial and nonspatial associations. Cereb Cortex. 17:1493-1503.

Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage. 54:2033-2044.

Awipi T, Davachi L. 2008. Content-specific source encoding in the human medial temporal lobe. J Exp Psychol Learn Mem Cogn. 34:769-779.

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

Bar M, Aminoff E. 2003. Cortical analysis of visual context. Neuron. 38:347-358.

Bar M, Aminoff E, Ishai A. 2008. Famous faces activate contextual associations in the parahippocampal cortex. Cereb Cortex. 18:1233-1238.

Barense MD, Bussey TJ, Lee AC, Rogers TT, Davies RR, Saksida LM, Murray EA, Graham KS. 2005. Functional specialization in the human medial temporal lobe. J Neurosci. 25:10239-10246.

Barense MD, Gaffan D, Graham KS. 2007. The human medial temporal lobe processes online representations of complex objects. Neuropsychologia. 45:2963-2974.

Bellgowan PS, Bandettini PA, van Gelderen P, Martin A, Bodurka J. 2006. Improved BOLD detection in the medial temporal region using parallel imaging and voxel volume reduction. Neuroimage. 29:1244-1251.

Bird CM, Burgess N. 2008. The hippocampus and memory: insights from spatial processing. Nat Rev Neurosci. 9:182-194.

Bird CM, Shallice T, Cipolotti L. 2007. Fractionation of memory in medial temporal lobe amnesia. Neuropsychologia. 45:1160-1171.

Bird CM, Vargha-Khadem F, Burgess N. 2008. Impaired memory for scenes but not faces in developmental hippocampal amnesia: a case study. Neuropsychologia. 46:1050-1059.

Bishop CM. 2006. Pattern recognition and machine learning. In: Berlin: Springer p 209.

Bohbot VD, Kalina M, Stepankova K, Spackova N, Petrides M, Nadel L. 1998. Spatial memory deficits in patients with lesions to the right hippocampus and to the right parahippocampal cortex. Neuropsychologia. 36:1217-1238.

Bonnici HM, Chadwick MJ, Kumaran D, Hassabis D, Weiskopf N, Maguire EA. 2012. Multi-voxel pattern analysis in human hippocampal subfields. Frontiers in Human Neuroscience. 6:290.

Bonnici HM, Kumaran D, Chadwick MJ, Weiskopf N, Hassabis D, Maguire EA. 2011. Decoding representations of scenes in the medial temporal lobes. Hippocampus. 22:1143-1153.

Brown TI, Ross RS, Keller JB, Hasselmo ME, Stern CE. 2010. Which way was I going? Contextual retrieval supports the disambiguation of well learned overlapping navigational routes. J Neurosci. 30:7414-22.

Brown TI, Hasselmo ME, Stern CE. 2014. A High-resolution study of hippocampal and medial temporal lobe correlates of spatial context and prospective overlapping route memory. Hippocampus. 24:819-39.

Buffalo EA, Bellgowan PS, Martin A. 2006. Distinct roles for medial temporal lobe structures in memory for objects and their locations. Learn Mem. 13:638-643.

Burwell RD. 2000. The parahippocampal region: corticocortical connectivity. Ann N Y Acad Sci. 911:25-42.

Canto CB, Wouterlood FG, Witter MP. 2008. What does the anatomical organization of the entorhinal cortex tell us? Neural Plast. 381243.

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

Carr VA, Viskontas IV, Engel SA, Knowlton BJ. 2010. Neural activity in the hippocampus and perirhinal cortex during encoding is associated with the durability of episodic memory. J Cogn Neurosci. 22:2652-2662.

Chadwick MJ, Hassabis D, Maguire EA. 2011. Decoding overlapping memories in the medial temporal lobes using high-resolution fMRI. Learn Mem. 18:742-746.

Chadwick MJ, Hassabis D, Weiskopf N, Maguire EA. 2010. Decoding individual episodic memory traces in the human hippocampus. Curr Biol. 20:544-547.

Chadwick MJ, Bonnici HM, Maguire EA. 2014. CA3 size predicts the precision of memory recall. Proc Natl Acad Sci U S A. 111:10720-5.

Chen J, Olsen RK, Preston AR, Glover GH, Wagner AD. 2011. Associative retrieval processes in the human medial temporal lobe: hippocampal retrieval success and CA1 mismatch detection. Learn Mem. 18:523-528.

Chua EF, Schacter DL, Rand-Giovannetti E, Sperling RA. 2007. Evidence for a specific role of the anterior hippocampal region in successful associative encoding. Hippocampus. 17:1071-1080.

Cipolotti L, Bird C, Good T, Macmanus D, Rudge P, Shallice T. 2006. Recollection and familiarity in dense hippocampal amnesia: a case study. Neuropsychologia. 44:489-506.

Cohen J, MacWhinney B, Flatt M, Provost J. 1993. Psyscope: An interactive graphical system for designing and controlling experiments in the Psychology laboratory using Macintosh computers. Behavioral Research Methods Instruments Computers. 25:257-271.

Colgin LL, Denninger T, Fyhn M, Hafting T, Bonnevie T, Jensen O, et al. 2009. Frequency of gamma oscillations routes flow of information in the hippocampus. Nature. 462:353-357.

Dale AM. 1999. Optimal experimental design for event-related fMRI. Hum Brain Mapp. 8:109-14.

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

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

Diana RA, Yonelinas AP, Ranganath C. 2008. High-resolution multi-voxel pattern analysis of category selectivity in the medial temporal lobes. Hippocampus. 18:536-541.

Diana RA, Yonelinas AP, Ranganath C. 2010. Medial temporal lobe activity during source retrieval reflects information type, not memory strength. J Cogn Neurosci. 22:1808-1818.

Dudukovic NM, Preston AR, Archie JJ, Glover GH, Wagner AD. 2010. High-resolution fMRI Reveals Match Enhancement and Attentional Modulation in the Human Medial-temporal Lobe. J Cogn Neurosci.

Dudukovic NM, Wagner AD. 2007. Goal-dependent modulation of declarative memory: neural correlates of temporal recency decisions and novelty detection. Neuropsychologia, 45:2608-2620.

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.

Duncan K, Ketz N, Inati SJ, Davachi L. 2012. Evidence for area CA1 as a match/mismatch detector: a high-resolution fMRI study of the human hippocampus. Hippocampus. 22:389-398.

Duvernoy HM. 1998. The Human Hippocampus. New York: Springer.

Edelman S. 1998. Representation is representation of similarities. Behav Brain Sci. 21:449-467; discussion 467-498.

Eichenbaum H, Cohen NJ. 2001. From Conditioning to Conscious Recollection: Memory Systems of the Brain. New York: Oxford University Press.

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

Ekstrom AD, Bazih AJ, Suthana NA, Al-Hakim R, Ogura K, Zeineh M, et al. 2009. Advances in high-resolution imaging and computational unfolding of the human hippocampus. Neuroimage. 47:42-49.

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.

Epstein R, DeYoe EA, Press DZ, Rosen AC, Kanwisher N. 2001. Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. Cognitive Neuropsychology. 18:481-508.

Epstein R, Kanwisher N. 1998. A cortical representation of the local visual environment. Nature. 392:598-601.

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

Fernandez G, Tendolkar I. 2006. The rhinal cortex: 'gatekeeper' of the declarative memory system. Trends Cogn Sci. 10:358-362.

Gabrieli JD. 1998. Cognitive neuroscience of human memory. Annu Rev Psychol. 49:87-115.

Gage FH, Thompson RG. 1980. Differential distribution of norepinephrine and serotonin along the dorsal-ventral axis of the hippocampal formation. Brain Res Bull. 5:771-773.

Gaisler-Salomon I, Schobel SA, Small SA, Rayport S. 2009. How high-resolution basal-state functional imaging can guide the development of new pharmacotherapies for schizophrenia. Schizophr Bull. 35:1037-1044.

Gilaie-Dotan S, Gelbard-Sagiv H, Malach R. 2010. Perceptual shape sensitivity to upright and inverted faces is reflected in neuronal adaptation. Neuroimage. 50:383-95.

Ginther MR, Walsh DF, Ramus SJ. 2011. Hippocampal neurons encode different episodes in an overlapping sequence of odors task. J Neurosci. 31:2706-11.

Glover GH, Lai S. 1998. Self-navigated spiral fMRI: Interleaved versus single-shot. Magnetic Resonance In Medicine. 39:361-368.

Glover GH, Law CS. 2001. Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. Magnetic Resonance in Medicine. 46:515-522.

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

Hannula DE, Libby LA, Yonelinas AP, Ranganath C. 2013. Medial temporal lobe contributions to cued retrieval of items and contexts. Neuropsychologia. In press.

Harrison SA, Tong F. 2009. Decoding reveals the contents of visual working memory in early visual areas. Nature. 458:632-635.

Haskins AL, Yonelinas AP, Quamme JR, Ranganath C. 2008. Perirhinal cortex supports encoding and familiarity-based recognition of novel associations. Neuron. 59:554-560.

Hassabis D, Chu C, Rees G, Weiskopf N, Molyneux PD, Maguire EA. 2009. Decoding neuronal ensembles in the human hippocampus. Curr Biol. 19:546-554.

Hasselmo ME, Schnell E. 1994. Laminar selectivity of the cholinergic suppression of synaptic transmission in rat hippocampal region CA1: 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.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 293:2425-2430.

Haynes JD, Rees G. 2006. Decoding mental states from brain activity in humans. Nat Rev Neurosci. 7:523-34.

Hsieh LT, Gruber MJ, Jenkins LJ, Ranganath C. 2014. Hippocampal activity patterns carry information about objects in temporal context. Neuron. 81:1165-78.

Insausti R, Insausti AM, Sobreviela MT, Salinas A, Martinez-Penuela JM. 1998. Human medial temporal lobe in aging: anatomical basis of memory preservation. Microsc Res Tech. 43:8-15.

Insausti R, Amaral DG. 2008. Entorhinal cortex of the monkey: IV. Topographical and laminar organization of cortical afferents. J Comp Neurol. 509:608-641.

Jenkins LJ, Ranganath C. 2010. Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. J Neurosci. 30:15558-65.

Johnson JD, Muftuler LT, Rugg MD. 2008. Multiple repetitions reveal functionally and anatomically distinct patterns of hippocampal activity during continuous recognition memory. Hippocampus. 18:975-980.

Johnson JD, McDuff SGR, Rugg MD, Norman KA. 2009. Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. Neuron. 63:697-708.

Jones EG, Powell TP. 1970. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. Brain. 93:793-820.

Jung MW, Wiener SI, McNaughton BL. 1994. Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. J Neurosci. 14:7347-7356.

Kim DH, Adalsteinsson E, Glover GH, Spielman DM. 2002. Regularized higher-order in vivo shimming. Magn Reson Med. 48:715-722.

Kirchhoff BA, Wagner AD, Maril A, Stern CE. 2000. Prefrontal-temporal circuitry for novelty encoding and subsequent memory. Journal of Neuroscience. 20:6173-6180.

Kirwan CB, Stark CE. 2007. Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. Learn Mem. 14:625-633.

Kloosterman F, Witter MP, Van Haeften T. 2003. Topographical and laminar organization of subicular projections to the parahippocampal region of the rat. J Comp Neurol. 455:156-171.

Knierim JJ, Lee I, Hargreaves EL. 2006. Hippocampal place cells: parallel input streams, subregional processing, and implications for episodic memory. Hippocampus. 16:755-764.

Kriegeskorte N, Bandettini P. 2007. Combining the tools: activation- and information-based fMRI analysis. Neuroimage. 38:666-668.

Kriegeskorte N, Formisano E, Sorger B, Goebel R. 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc Natl Acad Sci U S A. 104:20600-20605.

Kriegeskorte N, Mur M, Bandettini P. 2008. Representational similarity analysis - connecting the branches of systems neuroscience. Front Syst Neurosci. 2:4.

Kuhl BA, Rissman J, Wagner AD. 2011. Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. Neuropsychologia.

Kumaran D, Maguire EA. 2005. The human hippocampus: cognitive maps or relational memory? J Neurosci. 25:7254-7259.

Kumaran D, Maguire EA. 2006. The dynamics of hippocampal activation during encoding of overlapping sequences. Neuron. 49:617-29.

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

Kumaran D, Maguire EA. 2009. Novelty signals: a window into hippocampal information processing. Trends Cogn Sci. 13:47-54.

Lacy JW, Yassa MA, Stark SM, Muftuler LT, Stark CE. 2011. Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. Learn Mem. 18:15-18.

Lee I, Rao G, Knierim JJ. 2004. A double dissociation between hippocampal subfields: differential time course of CA3 and CA1 place cells for processing changed environments. Neuron. 42:803-815.

Lee AC, Buckley MJ, Pegman SJ, Spiers H, Scahill VL, Gaffan D, Bussey TJ, Davies RR, Kapur N, Hodges JR, Graham KS. 2005. Specialization in the medial temporal lobe for processing of objects and scenes. Hippocampus. 15:782-797.

Lee AC, Bussey TJ, Murray EA, Saksida LM, Epstein RA, Kapur N, Hodges JR, Graham KS. 2005. Perceptual deficits in amnesia: challenging the medial temporal lobe 'mnemonic' view. Neuropsychologia. 43:1-11.

Lee AC, Scahill VL, Graham KS. 2008. Activating the medial temporal lobe during oddity judgment for faces and scenes. Cereb Cortex. 18:683-696.

Leutgeb JK, Leutgeb S, Moser MB, Moser EI. 2007. Pattern separation in the dentate gyrus and CA3 of the hippocampus. Science. 315:961-966.

Leutgeb S, Leutgeb JK, Treves A, Moser MB, Moser EI. 2004. Distinct Ensemble Codes in Hippocampal Areas CA3 and CA1. Science.

Liang JC, Wagner AD, Preston AR. 2013. Content Representation in the Human Medial Temporal Lobe. Cereb Cortex. 23:80-96.

Libby LA, Ekstrom AD, Ragland JD, Ranganath C. 2012. Differential connectivity of perirhinal and parahippocampal cortices within human hippocampal subregions revealed by high-resolution functional imaging. J Neurosci. 32:6550-6560.

Libby LA, Hannula DE, Ranganath C. Medial Temporal Lobe Coding of Item and Spatial Information during Relational Binding in Working Memory. J Neurosci. 34:14233-42.

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.

Litman L, Awipi T, Davachi L. 2009. Category-specificity in the human medial temporal lobe cortex. Hippocampus. 19:308-319.

Macevoy SP, Epstein RA. 2009. Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. Curr Biol. 19:943-947.

Macevoy SP, Epstein RA. 2011. Constructing scenes from objects in human occipitotemporal cortex. Nat Neurosci. 14:1323-1329.

Manns JR, Eichenbaum H. 2006. Evolution of declarative memory. Hippocampus. 16:795-808.

Martin CB, McLean DA, O'Neil EB, Kohler S. 2013. Distinct familiarity-based response patterns for faces and buildings in perirhinal and parahippocampal cortex. J Neuro. 33:10915-10923.

McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychological Review. 102:419-457.

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

Moeller S, Yacoub E, Olman CA, Auerbach E, Strupp J, Harel N, et al. 2010. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magn Reson Med. 63:1144-1153.

Morris RG, Moser EI, Riedel G, Martin SJ, Sandin J, Day M, O'Carroll C. 2003. Elements of a neurobiological theory of the hippocampus: the role of activity-dependent synaptic plasticity in memory. Philos Trans R Soc Lond B Biol Sci. 358:773-786.

Moser E, Moser MB, Andersen P. 1993. Spatial learning impairment parallels the magnitude of dorsal hippocampal lesions, but is hardly present following ventral lesions. J Neurosci. 13:3916-3925.

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

Moser MB, Moser EI, Forrest E, Andersen P, Morris RG. 1995. Spatial learning with a minislab in the dorsal hippocampus. Proc Natl Acad Sci U S A. 92:9697-9701.

Munoz-Lopez MM, Mohedano-Moriano A, Insausti R. 2010. Anatomical pathways for auditory memory in primates. Front Neuroanat. 4:129.

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

Norman KA, Polyn SM, Detre GJ, Haxby JV. 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci. 10:424-430.

O'Neil EB, Barkley VA, Kohler S. 2013. Representational demands modulate involvement of perirhinal cortex in face processing. Hippocampus. 23:592-605.

O'Reilly RC, Rudy JW. 2001. Conjunctive representations in learning and memory: principles of cortical and hippocampal function. Psychol Rev. 108:311-345.

Olsen RK, Nichols EA, Chen J, Hunt JF, Glover GH, Gabrieli JD, Wagner AD. 2009. Performance-related sustained and anticipatory activity in human medial temporal lobe during delayed match-to-sample. J Neurosci. 29:11880-11890.

Pfeuffer J, Van de Moortele PF, Ugurbil K, Hu X, Glover GH. 2002. Correction of physiologically induced global off-resonance effects in dynamic echo-planar and spiral functional imaging. Magn Reson Med. 47:344-353.

Pihlajamaki M, Tanila H, Kononen M, Hanninen T, Hamalainen A, Soininen H, Aronen HJ. 2004. Visual presentation of novel objects and new spatial arrangements of objects differentially activates the medial temporal lobe subareas in humans. Eur J Neurosci. 19:1939-1949.

Poldrack RA. 2006. Can cognitive processes be inferred from neuroimaging data? Trends Cogn Sci. 10:59-63.

Poppenk J, McIntosh AR, Craik FI, Moscovitch M. 2010. Past experience modulates the neural mechanisms of episodic memory formation. J Neurosci. 30:4707-4716.

Poppenk J, Evensmoen H.R, Moscovitch M, Nadel L. 2013. Long-axis specialization of the human hippocampus. Trends Cogn Sci. 17:230-40.

Preston AR, Bornstein AM, Hutchinson JB, Gaare ME, Glover GH, Wagner AD. 2010. High-resolution fMRI of Content-sensitive Subsequent Memory Responses in Human Medial Temporal Lobe. J Cogn Neurosci. 22:156-173.

Preston AR, Thomason ME, Ochsner KN, Cooper JC, Glover GH. 2004. Comparison of spiral-in/out and spiral-out BOLD fMRI at 1.5 and 3 T. NeuroImage. 21:291-301.

Preston AR, Wagner AD. 2007. The medial temporal lobe and memory. In: Kesner RP, Martinez JL, eds. Neurobiology of Learning and Memory, Second Edition Oxford: Elsevier p 305-337.

Prince SE, Daselaar SM, Cabeza R. 2005. Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations. J Neurosci. 25:1203-1210.

Pruessner JC, Kohler S, Crane J, Pruessner M, Lord C, Byrne A, Kabani N, Collins DL, Evans AC. 2002. Volumetry of temporopolar, perirhinal, entorhinal and parahippocampal cortex from high-resolution MR images: considering the variability of the collateral sulcus. Cereb Cortex. 12:1342-1353.

Pruessner JC, Li LM, Serles W, Pruessner M, Collins DL, Kabani N, Lupien S, Evans AC. 2000. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. Cereb Cortex. 10:433-442.

Ritchey M, Wing EA, LaBar KS, Cabeza R. 2012. Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. Cereb Cortex. doi:10.1093/cercor/bhs258.

Quian Quiroga R, Kraskov A, Koch C, Fried I. 2009. Explicit encoding of multimodal percepts by single neurons in the human brain. Curr Biol. 19:1308-1313.

Ranganath C, Rainer G. 2003. Neural mechanisms for detecting and remembering novel events. Nat Rev Neurosci. 4:193-202.

Rissman J, Greely HT, Wagner AD. 2010. Detecting individual memories through the neural decoding of memory states and past experience. Proc Natl Acad Sci U S A. 107:9849-9854.

Rissman J, Wagner AD. In press. Distributed Representations in Memory: Insights from Functional Brain Imaging. Annu Rev Psychol.

Schobel SA, Lewandowski NM, Corcoran CM, Moore H, Brown T, Malaspina D, et al. 2009. Differential targeting of the CA1 subfield of the hippocampal formation by schizophrenia and related psychotic disorders. Arch Gen Psychiatry. 66:938-946.

Schultz H, Sommer T, Peters J. 2012. Direct evidence for domain-sensitive functional subregions in human entorhinal cortex. J Neurosci. 32:4716-4723.

Serences JT, Ester EF, Vogel EK, Awh E. 2009. Stimulus-specific delay activity in human primary visual cortex. Psychol Sci. 20:207-214.

Small SA, Nava AS, Perera GM, Delapaz R, Stern Y. 2000. Evaluating the function of hippocampal subregions with high-resolution MRI in Alzheimer's disease and aging. Microsc Res Tech. 51:101-108.

Small SA, Wu EX, Bartsch D, Perera GM, Lacefield CO, DeLaPaz R, et al. 2000. Imaging physiologic dysfunction of individual hippocampal subregions in humans and genetically modified mice. Neuron. 28:653-664.

Sommer T, Rose M, Glascher J, Wolbers T, Buchel C. 2005. Dissociable contributions within the medial temporal lobe to encoding of object-location associations. Learning & Memory. 12:343-351.

Squire LR, Stark CE, Clark RE. 2004. The medial temporal lobe. Annu Rev Neurosci. 27:279-306.

Staresina BP, Davachi L. 2006. Differential encoding mechanisms for subsequent associative recognition and free recall. J Neurosci. 26:9162-9172.

Staresina BP, Davachi L. 2008. Selective and shared contributions of the hippocampus and perirhinal cortex to episodic item and associative encoding. J Cogn Neurosci. 20:1478-1489.

Staresina BP, Duncan KD, Davachi L. 2011. Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. J Neurosci. 31:8739-8747.

Staresina BP, Henson RNA, Kriegeskorte N, Alink A. 2012. Episodic reinstatement in the medial temporal lobe. J Neuro. 32:18510-18156.

Strange BA, Hurlemann R, Duggins A, Heinze HJ, Dolan RJ. 2005. Dissociating intentional learning from relative novelty responses in the medial temporal lobe. Neuroimage. 25:51-62.

Suthana N, Ekstrom A, Moshirvaziri S, Knowlton B, Bookheimer S. 2011. Dissociations within human hippocampal subregions during encoding and retrieval of spatial information. Hippocampus. 21:694-701.

Suthana NA, Ekstrom AD, Moshirvaziri S, Knowlton B, Bookheimer SY. 2009. Human hippocampal CA1 involvement during allocentric encoding of spatial information. J Neurosci. 29:10512-10519.

Suzuki M, Johnson JD, Rugg MD. 2011. Recollection-related hippocampal activity during continuous recognition: a high-resolution fMRI study. Hippocampus. 21:575-583.

Suzuki WA. 2009. Comparative analysis of the cortical afferents, intrinsic projections and interconnections of the parahippocampal region in monkeys and rats. In: Gazzaniga MS, ed. The Cognitive Neurosciences IV Cambridge, MA: MIT Press p 659-674.

Suzuki WA, Amaral DG. 1994. Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. Journal Of Comparative Neurology. 350:497-533.

Swanson LW, Cowan WM. 1977. An autoradiographic study of the organization of the efferent connections of the hippocampal formation in the rat. J Comp Neurol. 172:49-84.

Taylor KJ, Henson RN, Graham KS. 2007. Recognition memory for faces and scenes in amnesia: dissociable roles of medial temporal lobe structures. Neuropsychologia. 45:2428-2438.

Tootell RB, Devaney KJ, Young JC, Postelnicu G, Rajimehr R, Ungerleider LG. 2008. fMRI mapping of a morphed continuum of 3D shapes within inferior temporal cortex. Proc Natl Acad Sci U S A. 105:3605-9.

Tubridy S, Davachi L. Medial temporal lobe contributions to episodic sequence encoding. Cereb Cortex. 21:272-80.

Van Hoesen G, Pandya DN. 1975. Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey. I. Temporal lobe afferents. Brain Res. 95:1-24.

Van Hoesen G, Pandya DN, Butters N. 1975. Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey. II. Frontal lobe afferents. Brain Res. 95:25-38.

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

Vercauteren T, Pennec X, Perchant A, Ayache N. 2009. Diffeomorphic demons: efficient non-parametric image registration. Neuroimage. 45:S61-72.

Verney C, Baulac M, Berger B, Alvarez C, Vigny A, Helle KB. 1985. Morphological evidence for a dopaminergic terminal field in the hippocampal formation of young and adult rat. Neuroscience. 14:1039-1052.

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.

Witter MP, Naber PA, van Haeften T, Machielsen WC, Rombouts SA, Barkhof F, et al. 2000. Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways. Hippocampus. 10:398-410.

Witter MP, Van Hoesen GW, Amaral DG. 1989. Topographical organization of the entorhinal projection to the dentate gyrus of the monkey. J Neurosci. 9:216-228.

Wittmann BC, Bunzeck N, Dolan RJ, Duzel E. 2007. Anticipation of novelty recruits reward system and hippocampus while promoting recollection. Neuroimage. 38:194-202.

Wolosin SM, Zeithamova D, Preston AR. 2012. Reward modulation of hippocampal subfield activation during successful associative encoding and retrieval. J Cogn Neurosci. 24:1532-1547.

Worsley KJ, Friston KJ. 1995. Analysis of fMRI time-series revisited--again. Neuroimage. 2:173-181.

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.

Yassa MA, Stark SM, Bakker A, Albert MS, Gallagher M, Stark CE. 2010. High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnestic Mild Cognitive Impairment. Neuroimage. 51:1242-1252.

Zeineh MM, Engel SA, Bookheimer SY. 2000. Application of cortical unfolding techniques to functional MRI of the human hippocampal region. Neuroimage. 11:668-683.

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.